



Proceedings of

10th International Conference on Intelligent Technologies
(ICIT – 2025)

&

10th International conference on Business Management Economics Social Sciences and
Humanities (BMESH – 2025)

Asian Institute of Technology, Thailand, 16 -18 August, 2025

Organized by

Asian Society for Research in Engineering Sciences (ASRES),
SPJ Centre for Multi-disciplinary Research (SCMR)
University of Pembangunan Jaya, Indonesia



Editors

Dr. Karm Veer Arya, ABV Indian Institute of Information Technology and Management,
Gwalior, India

Dr. V. K. Tripathi, National Institute of Technical Teachers' Training and Research, Bhopal,
India

Dr. Clara EviCitraningtyas, Universitas Pembangunan Jaya, Indonesia

Dr. Himanshu Govil, National Institute of Technology Raipur, India

Proceedings

10th International Conference on Intelligent Technologies
(ICIT – 2025)

&

10th International conference on Business Management Economics Social Sciences and
Humanitie (BMESH – 2025)

The papers contained in this proceedings book are contributions of authors. The authors are responsible for the content and the copyright infringement issues, if any.

Copyright and Reprints:

Abstracting of the information contained in this book is allowed by giving due credit to the source. Copying of the individual articles for non-commercial purposes is permitted without a fee provided that the due credit is given to the source. For other copying, reprint or republication, permission from the publisher must be obtained.

Published and Printed by:



Asian Society for Research in Engineering Sciences (ASRES)

ISBN: 978-81-955970-1-7

Table of Contents

Proceedings.....	(i)
Table of Contents.....	(ii)
Letter from Editors, ...	(iii)
Organizing Committees	(iv)
ASRES Competitive Paper Review Policy.....	(vi)
List of papers, ...	(vii)
Research papers ICIT – 2025	1
Research papers BMESH - 2025	384

Dear Delegates,

As Editors and on behalf of the organizing committee, it is our absolute pleasure to welcome your research into the conference proceedings of the 10th International Conference on Intelligent Technologies (ICIT – 2025) & 10th International conference on Business Management, Economics Social Sciences and Humanitie (BMESH – 2025).

A total of 100 competitive papers/abstracts were submitted from research entities and industries, which made a representation from different countries.

The review process has been stringent and only quality papers are included in the proceedings. The review policy was stringently followed. A total of 60 papers finally made it into these proceedings for publication, making a success rate of 60%.

Session chairs and field experts proactively commented and gave suggestions for improvement in the papers. Networking and Panel discussion sessions were helpful to delegates in enhancing their knowledge about new directions and methods of research in their respective fields.

We are sure that the quality of the papers adheres to the required standards befitting a leading academic conference. We feel proud to be a part of the conferences and proceedings, so should you be to be a part of it. Finally, to the authors, please accept our congratulations for all your hard work and papers published here.

Best regards,
Karm Veer Arya
Vipin Kumar Tripathi
Clara EviCitraningtyas
Himanshu Goviz

Organizing Committees

ICIT – 2025

- Prof. Len Gelman, University of Huddersfield, UK
- Prof. Karm Veer Arya, ABV Indian Institute of Information Technology and Management, Gwalior, India
- Prof. Andreas Nauertz, Mannheim Hochschule, University of Applied Sciences, Mannheim Germany,
- Prof. Pedro M. B. Silva Girao, Instituto Superior Tecnico, Universidade de Lisboa, Lisbon, Portugal
- Lui Gonzalez, ASRES, Organizing Secretary
- Dr. A. J. Anceno, Burapha University, Bangsaen, Thailand
- Dr. J. G. Singh, Asian Institute of Technology, Bangkok Thailand
- Dr. Sylvie Castagne, Nanyang Technological University, Singapore
- Ali Wiam, National Defence R & D, Male, Maldives
- Surachai Ampawasuvan, King Mongkut's University of Technology, North Bangkok
- Dr. Anu Vij, Al Dar University College, Dubai
- Dr. Pradip Peter Dey, National University, San Diego, USA
- Dr. R. K. Garg, National Institute of Technology Jalandhar, India
- Dr. Pierre Levy, National Science Research Centre, Paris, France
- Mr. Amit Pingle, John Deere, Pune
- Dr. Jaspreet Dhupia, University of Auckland, New Zealand
- Prof. Numan Hoda, Akdeniz University, Antalya, Turkey
- Dr. Kishlay Mitra, Indian Institute of Technology, Hyderabad, India
- Dr. Valentina Emilia Balas, Aurel Vlaicu University of Arad, Romania
- Dr. Gaurav Trivedi, Indian Institute of Technology, Guwahati
- Prof. Audithan Sivaraman, Netaji Subhash University of Technology, New Delhi
- Dr. P. K. Soni, Maulana Azad National Institute of Technology, Bhopal
- Dr. Shrikant Bansal, German Technical Cooperation, Visakhapatnam
- Dr. P. R. Sahu, Indian Institute of Technology, Bhubaneswar
- Dr. Ashish Anand, Indian Institute of Technology, Guwahati
- Dr. Sanjay Garg, Nirma University, Ahmedabad
- Dr. Vijay Gorana, Technical University, Kota
- Dr. M. R. Shindikar, College of Engineering, Pune
- Dr. Sunny Joseph Kalayathankal, K. E. College, Mannanam, Kottayam
- Dr. E. M. Somasekharan Nair, S. S. E. T Vidyanagar, Ernakulam, India
- Dr. K. P. Karunakaran, Indian Institute of Technology, Bombay
- Dr. H. K. Voruganti, National Institute of Technology, Warangal
- Dr. Kalyan Singh, Indian School of Mines, Dhanbad
- Dr. Chandra Sekhar Matli, National Institute of Technology, Warangal
- Dr. Anupam Agrawal, Indian Institute of Technology, Ropar
- Dr. Dinesh Thakur, Defense Institute of Advanced Technology, Pune
- Dr. R. K. Saxena, Sant Longowal Institute of Engineering and Technology, Sangarur
- Dr. Urmila Kulkarni-Kale, Savitribai Phule Pune University, Pune, India
- Dr. Manik Sharma, DAV University, Jalandhar, India
- Dr. Ruchi Doshi, BlueCrest University College, Liberia
- Dr. Satish Jain, SGSITS, Indore
- Dr. A. Rajendra Prasad, Sri Sai Ram Engineering College, Chennai

BMESH – 2025

- Dr. Clara EviCitraningtyas, Universitas Pembangunan Jaya, Indonesia
- Dr. Lulu Setiawati, Matana University, Jakarta, Indonesia
- Dr. Nidhi Bansal, Malaviya National Institute of Technology Jaipur, India
- Prof. Barry Ackers, University of South Africa, Pretoria, South Africa
- LuiGozalez, ASRES, Organizing Secretary
- Prof. Vipin K. Tripathi, National Institute of Technology, Agartala
- Prof. Minga Negash, Metropolitan State University, Denver, USA
- Dr. P MadyaDr Wan Fauziah Wan Yusoff, Director of Centre for Business and Entrepreneurship, Universiti Tun Hussein Onn Malaysia
- Prof. Mohd Muzammil, Visiting Fellow, University of Oxford, U. K.
- Ali Wiam, National Defence R &D, Male, Maldives
- Rabin K. Jana, Indian Institute of Management Raipur, India
- Dr. Len Gelman, University of Huddersfield, UK
- Dr. HilmiYazici, DilekSabanci State Conservatory, Selcuk University, Turkey
- Dr. MohitVij, Skyline University College, Sharjah, UAE
- Dr. Pradip Peter Dey, National University, San Diego, USA
- Dr. RuchiTyagi, University of Petroleum and Energy Studies, Dehradun, INDIA
- Dr. Pierre Levy, National Science Research Centre, Paris, France
- Mrs. Deepti Garg, University of Botswana, Botswana
- Ms. Anna White, Changsha JewTr Co. Ltd., Changsha, China
- Dr. E.M. Somasekharan Nair S.S.E.T Vidyanagar, Ernakulam, India
- Dr. Valentina Emilia Balas, Aurel Vlaicu University of Arad, Romania
- Dr. Shrikant Bansal, German Technical Cooperation, Visakhapatnam
- Dr. M. R. Shindikar, College of Engineering, Pune
- Dr. Pedro M. B. Silva Girao, Instituto Superior Tecnico, Universidade de Lisboa, Lisbon, Portugal
- Prof. Theo C. Haupt, Mangosuthu University of Technology, Durban, South Africa
- Dr. Ruchi Doshi, BlueCrest University College, Liberia

Conference committee contact details

Lui Gonzalez

Organizing Secretary

ICIT - 2025 & BMESH - 2025

Email: secretary@asres.org.in

Independent Quality Control Committee**Chairman**

Prof. K. V. Arya

ABV Indian Institute of Information Technology and Management, Gwalior

Members

- Dr. Suresh Vishwakarma
BC Hydro, Vancouver
- Prof. Deepak Sharma
Indian Institute of Technology, Guwahati

Paper Review Process Contact Details

Dr. Nishant Kulkarni

Email: nishantskulkarni@gmail.com

ASRES Competitive Paper Review Policy

The quality control and review process of ICIT – 2025 and BMESH – 2025 has the following steps:

1. Only papers that have not been published before/elsewhere are considered for presentation and publication in the conference. Authors are required to sign a declaration in this regard.
2. On receipt of a paper, it is checked by the organizing secretary of the conference to ensure that the paper is befitting one of the conference themes. If the paper does not match one of the conference themes, it is not accepted for the conference and the paper is returned to author with comments.
3. The papers accepted by organizing secretary after first screening, on the other hand, are forwarded to the coordinator, review process control committee who assigns paper to three international review committee members for a double-blind review. The essential requirements of reviewers include publishing experience as well as at least a Ph.D. degree for academic reviewers and Master's degree for experts from industry.
4. The Organizing Committee, International Review Committee, Quality Control Committee and the Editorial Board are completely efferent from each other. They are assigned different functions which they perform independently.
5. For acceptance of a paper, minimum two reviewers must give positive reviews, while the third reviewer must not recommend for a strong reject.
6. Accepted papers are sent to Quality Control Committee for their remarks. After positive remarks from Quality Control Committee, the paper is returned to the authors with the anonymous reviews.
7. Authors are required to make the corrections as suggested by the reviewers and re-submit the papers.
8. Once the Review Committee accepts the corrections made by authors, the paper is accepted for publication in the conference proceedings.
9. All final papers are properly language and technically edited by Editorial Board before publication in the conference proceedings.

List of papers

Paper No.	ICIT – 2025	Page number
	Details of paper	
ICIT01	Real-Time Terrain Rendering with Cloud-Edge Hybrid Caching and Dynamic LOD Using Quadtree Structures Eun-Seok Lee, Byungsun Park and Sung Jin Kim	1
ICIT02	Neotectonic Assessment of Dhauliganga River Basin, Eastern Kumaun, Himalaya Using Index of Relative Active Tectonics Mohd Nazish Khan, Dilawez Ali, M. Suhail, Norboyev Nodir Abdulkarim O'g'li, Samankulov Shukhrat Rashitovich, Ganiev Shahob, Fahimeh Jafari	8
ICIT03	Cloud Orchestration in SCM for Cloud Efficiency Mohan V Datar, Tanuja Pattanshetti	18
ICIT04	Potato Leaf Disease Classification Using Quantized CNN Models for Low-End Mobile Deployment Rekha Kaushik, Satwik Singh, Arpit Raj	28
ICIT05	Geospatial Characterization of Subsurface Combustion Using Satellite-Derived LST and Anomaly Metrics Purabi Chanda, Himanshu Govil, Subhanil Guha, Pralay Bhaumik	36
ICIT06	Experimental Study on Natural Convection Heat Transfer through a Composite Aluminum Heated Tube Anjan Kumar Sahu, Sudhansu Sekhar Singh, Vishwa Mohan Behera, Ramesh Chandra Nayak, Naresh Patra, Shrinu Mohanty	44
ICIT07	TRINETRA: Empowering the Visually Impaired with In-telligent Vision Assistance Atul Tripathi, Isha Pathak Tripathi, Deepika Dhinva, Harshit Shakya, Pranshul Kharniwal	52
ICIT08	Dynamic Modeling and Torque Analysis of a Humanoid Dual-Arm Manipulator under External Disturbances Shubham Vinayak Waman and Dr Sandeep Anasane	60
ICIT09	Cost Comparison of WiFi vs LiFi to determine the effects of having pre-installed optical fiber and the applicability of LiFi for commercial environments Ajay Sharma, Lalit Garg , Peter Albert Xuereb	69
ICIT10	Privacy first cookie-free analytics system Pratyush Goel, Isha Pathak Tripathi ,Atul Tripathi	78
ICIT11	Analysis of March 2025 Myanmar Earthquake using DInSAR Techniques based on Sentinel – 1 Data Pralay Bhaumik, Himanshu Govil, Shashi Kumar and Purabi Chanda	89
ICIT12	MPC for Quadruped Robot Tapan Ranjit Pakhidde and Prakash D Pantawane	96
ICIT13	Automated Customer Segmentation AI-Powered Lead Scoring for Edtech Mahesh Sunil Ulhe and Prof. Suhas S. Mohite	102
ICIT14	Enhancing Dynamic Resource Management via CNN-LSTM based Proactive Auto-Scaling in Edge Environments Bablu Kumar, Anshul Verma, and Pradeepika Verma	111

ICIT15	Synthesizing Virtual Datasets for Personalized Tourism and Insurance Services Hwanseog Kim, Jundong Lee, Ryumduck Oh, and Sangmin Suh	121
ICIT16	Dissimilar metal weld inspection using artificial intelligence in non destructive evaluation 4.0 M.Annadurai, R.Joselin, S.Anitha, M.Surya, E.Vidhya S.Saravanan	124
ICIT17	Enhancing Vehicle Safety: YOLOv8 and YOLOv9 with IR Sensors for Brake Light Detection and Accident Prevention Harsh Jaiswal and Surendra Nagar	131
ICIT18	A Model-Based Engineering Approach for the Design and Validation of an Incline Impact Testing Platform Tejas Parbhane, Nihal Dhote, Nilesh Sakle, Murtaza Laila	138
ICIT19	An Enhanced 1D-CNN-LSTM with Attention Mechanism for Real-Time Human Activity Recognition using IMU Data Amit Kumar Singh, Reeshav Guha, Dr. Vibha Vyas	147
ICIT20	Important uses of AI chatbots in education and healthcare sector Neeraj Kumar, Ravindra Nath	155
ICIT21	SMOTE Enhanced Framework using Neural Networks for Air Quality Assessment Kanak Kumar, Anshul Verma, and Pradeepika Verma	163
ICIT22	Federated Neuromorphic Learning for Real-Time Emotion Recognition from Wearable EEG: A Privacy-First Framework for Mental Health Monitoring Shreeya Arora	170
ICIT24	Design and Analysis of Dynamic Object Tracking Turret Mechanism Rohan kharat, Dr. Shantipal Ohol, Ajay Lohia and Ramesh Agnihotri	175
ICIT28	IoT Technology, Network and Security Sushil Kumar	183
ICIT29	Dynamic Modeling and Validation of a 14-DOF Dual-Arm Robot Using DA-RNE in MATLAB Ganesh Deshmukh, Rajiv Basavarajappa	190
ICIT30	Experimental Study for Application of Signal Processing Towards PSNR Based GPR Detection of Buried Objects of Varied Sizes, Material and Orientation Sujeet Roy, Himanshu Govil, Sandip Mukherjee	198
ICIT31	Interpretable GNNs for Personalized Risk Stratification of Drug--Drug Interactions from Electronic Health Records Piyush H. Kadethankar and Ramchandra Manthalkar	207
ICIT33	Formal Verification of CNN Robustness Using DeepPoly: An Empirical Study Across Activation Functions	214

	Anand Swaroop, Abhishek Singh, Girish Chandra, and Karm Veer Arya	
ICIT34	Comparative Analysis of Damping Materials for Vibration Control in Military Vehicle-Mounted Electronic Warfare Systems Suhas Ravindra Kokate, Dr.S.S.Ohol; Dr.K.P.Wani	223
ICIT35	Automated Drone Detection Using YOLOv11 Object Detection Framework Pranav Kaingade, Dr.A.G.Thosar and Prof.S.U.Ghunage	231
ICIT37	Spatial Attention Encoder-Decoder for Plant Disease Detection Debajyoti Dhar, Devshri Satyarthi, and K.V. Arya	238
ICIT38	Discomfort Index Prediction For A Driver In Vehicle Ergonomics Omkar Kishor Khole and Dr. Shantipal Suresh Ohol	244
ICIT39	Robust and Scalable Household Waste Classification Using an Advanced AI Model Kanak Kumar, Anshul Verma, and Pradeepika Verma	251
ICIT40	Polyp-EfficientNet: Lightweight CNN for Accurate Polyp Segmentation Harsha Vardhan Lakavath, Rakesh Kumar Dhayal, Devshri Satyarthi, and K.V. Arya	259
ICIT42	A comprehensive review of metal matrix composites in additive manufacturing: materials, fabrication, and industrial applications Deepak kumar sahu, v.k. tripathi	265
ICIT43	XML-based Android Malware Detection for Enhancing Security in Global Supply Chains Ansh Shah, Bhavnish Nanda, Keyaba Gohil, Fenil Ramoliya Dev Mehta,Rajesh Gupta, Sudeep Tanwar, and Nileshkumar Patel	287
ICIT45	VGG19-Based Multi-Modal Feature Fusion for Lung Disease Classification: Performance Comparison of Early, Late, and Intermediate Fusion Strategies Gautami Shingan, Priya Ranjan, and Sachin Chaudhary	294
ICIT46	Prioritizing SQL Injection Threats: A Risk-Based Approach to Detection and Mitigation Nilima Dinesh Bobade and Swati S.Sherekar	305
ICIT47	On Data Privacy in e-Governance Data Mining Shailesh Kadre, Vipin K. Tripathi	311
ICIT50	Gaussian Process-Assisted Inverse Design for Adaptive Pareto Front Gap Filling in Multi-Objective Optimization Shailesh S. Kadre, S. K. Vemuri, P.K. Mukka, R.G. Paspula and Vipin K. Tripathi	317

ICIT51	Machine learning-based analysis of clinical data in hemodialysis patients Pavani Peddu, Poonam Sharma, Anand Chellapan	328
ICIT52	Identification and mapping of methane plume in the barmer region using emit hyperspectral data Nimisha Kasana, Himanshu Govil	345
ICIT53	Q-MedNanoNet: Quantum-Secured Nano-Network Tra_c Framework for Wireless Implantable Devices Maurya Thakore ¹ , Yashvi Shah ² , Dhrishita Parve ³ Ansh Vachhani ¹ , Rimmi Sharma ¹ , Lakshin Pathak ² , Bhavya Vora ⁴ , Rajesh Gupta ² , and Sudeep Tanwar ²	352
ICIT56	AI Content Curator—The Future of Media Ananya Tripathi	361
ICIT57	Integrating Talent Management into Sustainable Human Resource Strategies Abha Pathak ¹ , Dr. Sonam Singh ² , Pallavi Thorat ³ , Bisini V ⁴ , Heena Satkut ⁵	365

Paper No.	BMESH- 2025	Page number
	Details of paper	
BME01	Crisis 2.0: Unvelling Crisis Management practices in various sector Jose cris o. sotto	384
BME02	The Effect of Asset Growth, Business Risk, And Good Corporate Governance on Financial Sustainability; an Empirical Study on Property & Real Estate Companies Listed on the Indonesia Stock Exchange for the period 2020-2024 Siti Rahma Anggraini, Arry Eksandy	400
BME03	The Effect of Instrumental Background Music to Choic of Accomodation At Travel Website Chici putri Alifalexis Winata, Kevin Daffa Aditya Wijaya, Edelwis, Aries Yulianto	414
BME04	The Effect of Profitability, Liquidity, And Non-Debt Tax Shield on Capital Structure With Firm Size As a Moderating Varlable Kusumastuti Gabriella Ageng Putri, Devi, Novy Silvia	415
BME05	Unveiling Decision-Making Patterns in Personal Finance : A Decade of Bibiliometric Insights(2015-2025) Abdul Rahman, Vania Rabiola Putri, Syifa Hanifa Salsabil, Cynthia Sari Dewi	429
BME06	The Effect of Job Search Self-Efficacy on Career Adaptability of Indonesia Fresh Graduates Adinda Anggun Pramesti Veronica Anastasia Melany Kaihatu	438
BME07	The Influence of Live Shopping Motivation As Mediating Variable For Shopee Costomers With Fashion Product Andhika Yulianto, Windarko, Fazya Nabila	446
BME08	Social Entrepreneurship in Indonesia: Social Capital, Financial Capital, Market Orientation On social Enterpreneurship Performance and its Impact On Economic Enpowerment And Social Impact Of Local Communities Lutfi Nurizki, Yusuf Iskandar, Dede Suleman Andini Putri Erianti	460
BME09	Exoticism and Global Myth: A Semiotic Analysis of International Hotel Franchises in Bali Ronald Maraden Parlindungan Silalahi, Clara Evi C. Citraningtyas, Raden Raro Gisela Anindita Maheswari	485
BME10	Between Duty and Dreams: Indonesian Millennials and Gen z Perspective on the Sandwich Genration in Home Sweet Loan Kheren Immanuel, Clara Evi C. Citraningtyas	493

BME11	Research on Consumer Purchasing Behavior:A Study of The apparel Industry in Ho Chi Minh City Tran Thi Ninh, Nguyen Thi Tuyet Nga	494
BME13	Aggressive Driving Traits and Tailored Technical Solution for two – Wheeler Mobility Kshipra Moghe, Saniya Atalatti	504
BME14	Revealing Abortion Stigma: A Qualitative Study In Early Adult Women Who Have An Abortion Analiyah Azzakiyah Tasning, Salsabila Hetri Soraya, Ratna Rusmiyati, Aries Yulianto	516
BME15	Blended Learning in The Eyes Of Students: Between Offline, Online Synchronous, and Asynchronous Clara Evi Candrayuli Citraningtyas	524
BME16	Strategic Statelessness And Cultural Resistance: Zomia, The Baduy Dalam, And The Politics of Autonomy in Southeast Asia Sebastian Satrio Darminto	525

Research papers
ICIT-2025

Real-Time Terrain Rendering with Cloud-Edge Hybrid Caching and Dynamic LOD Using Quadtree Structures

Eun-Seok Lee¹, Byungsun Park¹ and Sung Jin Kim²

¹ Gangneung-Wonju National University, 150, Namwon-ro, Wonju,
Gangwon-do, 26403, Republic of Korea

² Myeongji College, 134, Gajae-ro, Eunpyeong-gu, Seoul, 03656, Republic of Korea
elflee77@gmail.com, bsparkgwnu@gmail.com, ict214548@mjc.ac.kr

Abstract. This paper introduces a new approach for real-time terrain rendering, which integrates dynamic Level of Detail (LOD) adjustments and edge server caching to optimize rendering performance and reduce latency. The method aims to enhance rendering efficiency by dynamically adjusting LOD based on the camera's position, and by utilizing an edge server to cache frequently accessed terrain data, minimizing the need for cloud data retrieval. The experimental results demonstrate significant improvements: the proposed method reduced rendering time by 44%, decreased frame drops by 93%, and lowered latency by 50% when compared to traditional terrain rendering methods. This hybrid approach provides a scalable solution for real-time rendering of large-scale terrains, ensuring both high performance and stability.

Keywords: Terrain Rendering, Level of Detail (LOD), Edge Server Caching, Real-Time Rendering, Performance Optimization, Frame Stability

1. Introduction

Terrain rendering is a fundamental component of real-time graphics applications, such as gaming, virtual simulations, and geospatial systems. The ability to efficiently render large and complex terrains in real-time is critical for creating immersive experiences. However, as terrain data grows in size and complexity, traditional rendering techniques face significant challenges in maintaining high performance. Methods such as Level of Detail (LOD) have been developed to address these challenges by adjusting the detail of terrain based on the camera's distance, allowing for efficient rendering of distant objects. Additionally, out-of-core terrain rendering techniques have been employed to manage large datasets by streaming data from disk or external storage as needed, overcoming memory constraints. While these techniques have proven useful in many cases, they still encounter issues related to memory usage, data transfer speeds, and performance bottlenecks, especially when dealing with massive datasets or real-time demands.

One of the key limitations of existing approaches is the use of cloud-based data storage, which introduces latency due to network delays and bandwidth limitations. Streaming large amounts of terrain data from the cloud can significantly slow down the rendering process, especially when large, high-resolution textures or detailed terrain data are involved. Moreover, traditional LOD techniques, while effective in reducing computational load, do not dynamically adapt to the varying needs of the terrain data in real-time, limiting their ability to maintain high performance across diverse environments.

To address these challenges, we propose a novel terrain rendering system that integrates cloud-edge hybrid caching with dynamic LOD management and Quadtree-based data structures. Our approach aims to reduce data transfer latency and optimize rendering performance by caching frequently accessed terrain data on edge servers. This reduces the reliance on the cloud for real-time data retrieval, ensuring faster access to terrain data and more responsive rendering. By dynamically adjusting the level of detail based on the camera's proximity to the terrain, our system ensures that only relevant terrain data is processed at any given time, minimizing memory usage and computational overhead.

The use of Quadtree structures further enhances the system by organizing the terrain data hierarchically, ensuring that only the necessary chunks of data are loaded at any moment. This allows for highly efficient memory management, ensuring that the system can handle large terrain datasets while maintaining smooth real-time performance. By combining these techniques, our system offers a significant improvement over traditional methods, providing faster rendering, reduced memory consumption, and lower network latency.

This research addresses several key limitations in current terrain rendering methods. First, while existing LOD-based techniques optimize rendering performance, they often do not handle dynamic changes in terrain data or camera perspectives efficiently. Our dynamic LOD system ensures that terrain detail is adjusted in real-time based on camera proximity, overcoming the static nature of traditional LOD methods. Second, many out-of-core techniques suffer from latency issues when relying solely on cloud storage for data streaming. By integrating edge computing, we reduce the latency of data retrieval, ensuring that rendering is responsive even for large-scale datasets. Lastly, the traditional methods often struggle with memory management, especially when handling large terrain datasets. Our use of Quadtree structures for data organization provides a scalable solution, ensuring that only the most relevant data is loaded and rendered at any given time. In summary, our proposed system not only improves upon the limitations of previous LOD and out-of-core techniques, but it also introduces a new paradigm for real-time terrain rendering that integrates the advantages of cloud-edge computing, dynamic LOD, and optimized memory management.

2. Related Works

In recent years, terrain rendering techniques have evolved significantly through the integration of GPU acceleration, Level of Detail (LOD) management, Quadtree-based data structures, and cloud-edge computing. These advancements have played a crucial role in improving the performance of real-time rendering for large-scale terrain data. This section provides an overview of key methods employed in terrain rendering, including LOD management, Quadtree structures, and cloud-edge hybrid caching, along with a comparison to the proposed methodology.

Level of Detail (LOD) [1] is a critical technique used to reduce computational complexity by adjusting the resolution of terrain data based on camera distance. Several methods have been proposed to implement LOD for terrain rendering, with notable studies employing GPU acceleration for real-time performance. The use of GPU-based LOD has been explored in various studies, such as GPU tessellation for terrain visualization, where tessellation is dynamically adjusted based on proximity and screen-space error to maintain high performance [2], [3], [4]. These methods significantly improve the real-time rendering of terrain by reducing unnecessary detail in distant areas while maintaining quality in close-up views. However, many existing techniques still rely heavily on pre-processed data and face limitations in handling dynamic terrain or data streaming from external storage.

To address these issues, Quadtree structures have been extensively utilized for organizing and managing large-scale terrain data in a hierarchical manner. Quadtree-based LOD management allows terrain data to be subdivided into smaller chunks, which are rendered dynamically based on the camera's position. This method has been widely adopted for its efficiency in data loading and rendering, ensuring that only relevant terrain data is processed [5], [6], [7]. In particular, Quadtree-based algorithms allow for efficient memory management by selectively loading terrain patches that are visible or require high detail. While these methods are effective in reducing memory usage and enhancing performance, they often fall short when it comes to real-time data retrieval and adapting to dynamic changes in terrain data.

Recent advancements have incorporated cloud-edge hybrid caching to overcome the limitations of traditional data streaming and reduce latency in terrain rendering. Edge computing, in particular, has shown promise by offloading some of the rendering and data management tasks to edge servers

closer to the user, thereby reducing the load on central cloud servers and minimizing data retrieval times. Several studies have explored edge-based caching and its role in real-time rendering, demonstrating that edge servers can significantly enhance rendering speed and reduce latency by caching frequently accessed data closer to the user [8], [9]. This approach is particularly beneficial in large-scale terrain rendering, where cloud storage alone would introduce significant delays due to network constraints.

Additionally, AI-based techniques have been explored to further optimize LOD and terrain generation. Machine learning models have been employed to predict terrain details and automatically adjust LOD based on the terrain's complexity and user interaction. For instance, deep learning models have been used to predict high-resolution textures from low-resolution inputs, and reinforcement learning has been applied to adjust LOD dynamically for terrain features [10], [11]. These AI-based methods hold the potential to further reduce computational overhead by learning to adapt to varying levels of detail in real time. However, while these techniques are promising, they require substantial training data and computational resources, which can limit their applicability in real-time systems that need to process terrain data dynamically.

In summary, existing methods have effectively addressed the challenge of terrain rendering performance through LOD, Quadtree structures, and GPU acceleration. However, they often struggle with handling dynamic terrain data, real-time data streaming, and latency issues in large-scale applications. Our approach, which integrates cloud-edge hybrid caching, dynamic LOD, and Quadtree-based data management, offers a more comprehensive solution by optimizing both data retrieval and rendering performance in real time.

3. Proposed Methodology for Real-time Terrain Rendering

We propose a novel approach for optimizing terrain rendering performance by integrating dynamic Level of Detail (LOD) management, GPU-based Quadtree structure, and cloud-edge hybrid caching. The proposed methodology combines these techniques to enhance real-time rendering performance, efficiently manage memory usage, and minimize data retrieval latency. The system leverages GPU acceleration, dynamic LOD adjustments, and cloud-edge data management to process large-scale terrain datasets in real-time.

The core of the proposed methodology is the dynamic LOD adjustment based on camera position and screen space error. In this approach, the LOD is adjusted in real time based on the distance between the camera and the terrain, ensuring that high-resolution data is applied to areas close to the camera and low-resolution data is used for distant areas. This method minimizes screen space error and optimizes memory usage. The LOD is dynamically recalculated according to the camera's position, and the screen space error is kept within an acceptable range. The screen space error is calculated as follows:

$$LOD = \frac{D}{S} \times \frac{1}{\Delta e} \quad (1)$$

where D is the distance between the camera and the terrain, S is the screen resolution, and Δe is the allowable screen space error. By using this formula, the LOD for different sections of the terrain is determined based on the proximity to the camera and the allowable error in the rendered image. Along with dynamic LOD management, the proposed method uses a GPU-based Quadtree structure for efficient terrain data management. Quadtree structures are hierarchical data structures that partition the terrain into smaller, manageable chunks. This structure enables dynamic data loading, rendering only the relevant portions of the terrain based on the camera's viewpoint. The GPU's parallel processing capabilities allow these terrain chunks to be rendered quickly and efficiently. As the camera moves through the terrain, the system dynamically loads and renders the required chunks, optimizing memory usage and rendering speed. This system efficiently handles large-scale terrain data and ensures that only the necessary data is processed.

To optimize real-time rendering performance and minimize data retrieval latency, the methodology incorporates a cloud-edge hybrid caching system. In this system, edge servers cache high-resolution terrain data close to the camera's position, providing quick access to nearby data. Cloud servers, on the other hand, stream data from the cloud when needed, typically for more distant regions of the terrain or rarely accessed data. This hybrid approach reduces the dependency on cloud storage for frequently accessed data and ensures that data retrieval times are minimized. The cached data on edge servers is updated periodically, and any required updates are fetched from the cloud servers in real-time. This reduces the latency typically encountered with cloud-based systems and enhances the overall responsiveness of the system.

The reason for using edge computing is to reduce the latency involved in data retrieval and to enhance the real-time rendering performance. By utilizing edge servers, data is stored closer to the user, enabling faster access to frequently needed data. In contrast, cloud servers are used for long-term storage and for data that is less frequently accessed. The edge computing system facilitates quicker data loading, improving rendering performance and ensuring that large terrain datasets can be handled efficiently. The real-time data transfer between cloud and edge servers minimizes the delay caused by network congestion, providing a seamless rendering experience.

The hybrid caching system's performance benefits are derived from the efficient use of memory and the reduction in data retrieval times. The cache efficiency (CE) can be quantified as:

$$CE = \frac{CHR}{CMR} \times \frac{1}{DL} \quad (2)$$

where cache hit rate (CHR) is the percentage of data accessed from the edge server's cache, cache miss rate (CMR) is the rate at which data must be retrieved from cloud servers, and data latency (DL) is the delay incurred in fetching data from cloud storage. By increasing the cache hit rate and reducing the cache miss rate, the system ensures that most of the data required for rendering is available on the edge server, significantly reducing data retrieval delays.

The hybrid LOD system, combining dynamic LOD adjustments with Quadtree structures and cloud-edge caching, provides an efficient and scalable solution for real-time terrain rendering. By leveraging GPU parallel processing and edge computing, the system optimizes the performance of large-scale terrain rendering while ensuring high visual quality. This approach minimizes data transfer latency, optimizes memory usage, and enhances the overall rendering speed, offering a significant performance improvement over traditional methods.

In conclusion, the proposed methodology integrates dynamic LOD management, GPU-based Quadtree rendering, and cloud-edge hybrid caching to provide a robust solution for optimizing real-time terrain rendering. The combination of these techniques ensures that terrain data is efficiently managed and rendered in real-time, with minimal latency and maximum performance. This approach offers significant.

This version presents the methodology in a single cohesive flow without subsections, as you requested. It includes the technical explanations improvements in rendering efficiency and memory management, making it well-suited for large-scale terrain visualization in modern applications of dynamic LOD, GPU-based Quadtree management, and cloud-edge hybrid caching, all while incorporating the necessary mathematical formulae for optimization. If you need any further adjustments or additional details, feel free to let me know!

4. Experimental Result

The experiments were conducted to evaluate the performance of the proposed system and compare it with traditional methods in terms of rendering time, frame drop count, and latency. The system was tested with a large-scale synthetic terrain dataset (4096 x 4096 tiles, each with 2K x 2K vertices). The tests were performed under four different conditions: original data loading (no LOD), LOD applied, edge server only, and LOD + edge server combined. The hardware setup for

the experiments consisted of an Intel i5 13th Gen CPU, 32 GB RAM on the edge server, and Intel i5 13th Gen CPU, 32 GB RAM, and NVIDIA RTX 3090 GPU for rendering on the client PC. The cloud server used a standard cloud storage setup optimized for high-speed data retrieval. The terrain data was loaded from the cloud, and the performance of each method was compared by measuring rendering time, frame drops, and latency.



Fig. 1. Visualization of OSM data for Seongnam, Gyeonggi Province, South Korea. The figure illustrates the terrain model generated from the data, with corresponding measurements recorded during the rendering process.

Table 1. Table captions should be placed above the tables.

Method	Average Rendering Time (ms)	Frame Drops (count)	Latency (ms)
Original Data (No LOD)	249	32	148
LOD Applied	176	12	121
Edge Server Only	203	4	89
LOD + Edge Server Combined	137	0	51

Figure 1 shows the visualization result of the OSM data. The test area is Seongnam, Gyeonggi Province, South Korea, and **Table 1** records the average values of the data measured while capturing the image. The results clearly demonstrate that the LOD + Edge Server Combined method provides the best performance across all metrics. The average rendering time is reduced by 44% compared to the Original Data (No LOD) method, which had the longest rendering time. The frame drops were significantly reduced, with the LOD + Edge Server Combined method showing only 2 frame drops, a 93% reduction compared to the baseline method. This indicates that the system's performance is more stable and capable of handling real-time rendering without interruptions.

Furthermore, the latency was reduced by 50% with the LOD + Edge Server Combined method, which achieved an average latency of 50 ms. This improvement is primarily due to the edge server caching the most frequently used terrain data, thereby reducing the need to retrieve data from the cloud, which introduces significant delays. In contrast, the Original Data (No LOD) method, which relied solely on cloud-based data retrieval, showed a much higher latency of 150 ms. These results confirm that the combination of dynamic LOD adjustment and edge server caching significantly improves real-time rendering performance by reducing rendering time, frame drops, and latency. The edge server plays a key role in this performance improvement by providing faster access to the required terrain data, ensuring smooth rendering even for large-scale datasets.

The experimental results highlight the advantages of the LOD + Edge Server Combined method over traditional terrain rendering techniques. By integrating dynamic LOD and edge server caching, the proposed method significantly reduces both rendering time and latency, while maintaining a stable frame rate. This hybrid approach leverages the edge server's ability to cache frequently accessed data, minimizing the need for cloud retrieval and thus reducing data retrieval delays.

The Original Data (No LOD) method, which relies solely on cloud data retrieval, performed the worst in all three metrics. The lack of LOD optimization and the reliance on the cloud for every data request resulted in higher rendering times, more frame drops, and increased latency. The LOD applied and Edge Server Only methods showed improvements over the baseline, but they still fell short of the hybrid method's performance. Overall, the proposed LOD + Edge Server Combined approach offers the best solution for large-scale terrain rendering, providing optimal performance, reduced latency, and enhanced stability during real-time rendering tasks.

5. Conclusion

In this paper, we proposed a hybrid terrain rendering approach combining dynamic LOD adjustment and edge server caching. The experimental results showed that our method significantly improved rendering time, frame stability, and latency. The LOD + Edge Server Combined approach achieved a 44% reduction in rendering time, a 93% reduction in frame drops, and a 50% reduction in latency compared to traditional methods. These improvements demonstrate that leveraging edge computing with LOD techniques is an effective solution for real-time rendering of large-scale terrains, ensuring both performance and stability.

6. Acknowledgement

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP)-Innovative Human Resource Development for Local Intellectualization program grant funded by the Korea Government (MSIT)(IITP-2025-RS-2023-00260267)

References

1. Schütz, M., Herzberger, L., & Wimmer, M. (2023). SimLOD: Simultaneous LOD Generation and Rendering. *arXiv preprint arXiv:2310.03567*.
2. Masood, Z., Zheng, J., Ahmad, I., Irfan, M., & Saeed, A. (2022, January). High-performance gpu based planetary-scale terrain visualization. In *Proceedings of the 2022 5th international conference on image and graphics processing* (pp. 261-267)
3. Kulhanek, J., Rakotosaona, M. J., Manhardt, F., Tsalicoglou, C., Niemeyer, M., Sattler, T., ... & Tombari, F. (2025). LODGE: Level-of-Detail Large-Scale Gaussian Splatting with Efficient Rendering. *arXiv preprint arXiv:2505.23158*.
4. Zhai, R., Lu, K., Pan, W., & Dai, S. (2016). GPU-based real-time terrain rendering: Design and implementation. *Neurocomputing*, 171, 1-8.
5. Fu, H., Yang, H., Chen, C., Zhang, H., Hao, M., & Yu, T. (2021, June). Adaptive LOD representation of terrain model based on quad-tree. In *International Conference on Signal Image Processing and Communication (ICSIPC 2021)* (Vol. 11848, pp. 258-264). SPIE.
6. Croci, J. A., Amiraghdam, A., & Pajarola, R. (2022, November). Terrender: A Web-Based Multi-Resolution Terrain Rendering Framework. In *Proceedings of the 27th International Conference on 3D Web Technology* (pp. 1-11).
7. Wu, J., Yang, Y., Gong, S. R., & Cui, Z. (2010). A New Quadtree-based Terrain LOD Algorithm. *J. Softw.*, 5(7), 769-776.

8. Ye, Z., Fu, Y., Zhang, J., Li, L., Zhang, Y., Li, S., ... & Lin, Y. C. (2025, March). Gaussian Blending Unit: An Edge GPU Plug-in for Real-Time Gaussian-Based Rendering in AR/VR. In *2025 IEEE International Symposium on High Performance Computer Architecture (HPCA)* (pp. 353-365). IEEE.
9. Mehrabi, A., Siekkinen, M., Kämäräinen, T., & ylä-Jäski, A. (2021). Multi-tier CloudVR: Leveraging edge computing in remote rendered virtual reality. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 17(2), 1-24.
10. Xiao, T., Chen, Y., Zhong, S., Kiefer, P., Krukar, J., Kim, K. G., ... & Raubal, M. (2025, April). Sketch2Terrain: AI-Driven Real-Time Terrain Sketch Mapping in Augmented Reality. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1-24).
11. Li, W., & Hsu, C. Y. (2020). Automated terrain feature identification from remote sensing imagery: a deep learning approach. *International Journal of Geographical Information Science*, 34(4), 637-660.

Neotectonic Assessment of Dhauliganga River Basin, Eastern Kumaun, Himalaya Using Index of Relative Active Tectonics

Mohd Nazish Khan¹, Dilawez Ali², M. Suhail³, Norboyev Nodir Abdulkarim O'g'li⁴,
Samankulov Shukhrat Rashitovich⁴, Ganiev Shahob⁵, Fahimeh Jafari⁶

¹ Department of Physical Geography and Natural Resources, Samarkand State University,
Samarkand, Uzbekistan

² Department of Geology, Samarkand State University, Samarkand, Uzbekistan

³ Center of Applied of Remote Sensing and GIS Applications, Samarkand State University,
Samarkand, Uzbekistan

⁴ Samarkand State Architecture and Construction University, Samarkand, Uzbekistan

⁵ Department of Geology & Hydrometeorology, Samarkand State University, Samarkand,
Uzbekistan

⁶ Department of Geography, Ferdowsi University of Mashhad, Mashhad, Iran

nazishgeo@gmail.com; dilawezali007@gmail.com; netgeo.suhail@gmail.com;
nodirnorboyev518@gmail.com; shuxrat.samankulov87@gmail.com;
shaxobganiyev88@gmail.com; fa.jafari@alumni.um.ac.ir

Abstract. The present study investigates neotectonic activity in the Dhauliganga River Basin, situated in the eastern Kumaon Himalaya. Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data were utilized to derive key morphotectonic indices, including drainage density, drainage texture, hypsometric integral, bifurcation ratio, asymmetry factor, valley floor width-to-height ratio, stream length gradient, and mean stream length. These parameters were employed to evaluate the spatial variability of tectonic activity across the basin. The study ranks sub-basins based on the values of individual indices to interpret the degree of neotectonic influence. Among the indices, hypsometric integral, asymmetry factor, valley floor width-to-height ratio, and stream length gradient most effectively reflect spatial patterns of active tectonics and the competing influence of erosional forces shaping the sub-basins.

Keywords: Neotectonics, ASTER, Dhauliganga, DEM.

1. Introduction

The Himalayan orogeny is attributed to the continental-continental collision between the Indian and Eurasian Plates, which began around 65 million years ago during the Eocene epoch (Gansser, 1964; Dewey & Bird, 1970; McKenzie & Sclater, 1971; Molnar & Tapponnier, 1975). This tectonic convergence gave rise to the Himalayan mountain range and resulted in the formation of three major north-dipping thrust zones: the Main Central Thrust (MCT), Main Boundary Thrust (MBT), and Himalayan Frontal Thrust (HFT), which display a southward progression in relative age (De Celles et al., 2000). These thrusts are interconnected by a complex network of tectonic structures and geomorphic features that have been progressively transported southward due to ongoing collision and convergence (Kalpana et al., 2023; Khan & Govil, 2023; Khan et al., 2024). Furthermore, the Himalaya continues to evolve tectonically, as evidenced by active faults and thrusts, manifested in features such as rising peaks (including the 8,848-meter Mount Everest), fault scarps, V-shaped canyons, shutter ridges, and offset streams, all contributing to its dynamic and diverse landscape. Further, tectonic and surface processes interact dynamically to shape landscapes and landforms, with their relative timing and magnitudes recorded in the morphology

of the Earth's surface. While the influence of tectonic uplift on surface processes viz., weathering and erosion, has been extensively recognized, the reciprocal role of surface processes in modulating tectonic activity has only gained significant scholarly attention in recent years (Willett et al., 2006; Farooq et al., 2015; Khan et al., 2023). The study of active tectonics is instrumental in quantifying crustal deformation over timescales that are pertinent to human society (Keller & Pinter, 2002; Omar & Mosar, 2019; Khan et al., 2024). It provides critical insights into the interrelationship between tectonic forces and geomorphological processes that govern landscape evolution (Burbank & Anderson, 2011; Taib et al., 2023). The impacts of these interactions are commonly assessed through quantitative analyses of surface geometry, topography, and their temporal variations (Turner, 2006). Basin-scale morphometric indices, in particular, serve as valuable tools for detecting active structural features such as faults, thrusts, and shear zones, as well as documenting active ruptures, displaced landforms, and the geomorphic expressions of ongoing deformation (Wallace, 1990).

2. Study Area

The Dhauliganga River, a tributary of the Kaliganga in the Kumaon Himalaya, flows through the Pithoragarh district of Uttarakhand, India, and forms part of the upper Kaliganga river system. Originating in the Lesser Himalayan terrain, the river drains a steep, rugged landscape characterized by high relief, narrow valleys, and structurally controlled channels. The basin lies within a tectonically active zone influenced by the Trans-Himadri Fault, Main Central Thrust (MCT), and North Almora Thrust (NAT), with geology dominated by metasedimentary and low-grade metamorphic rocks such as quartzites, slates, and phyllites.

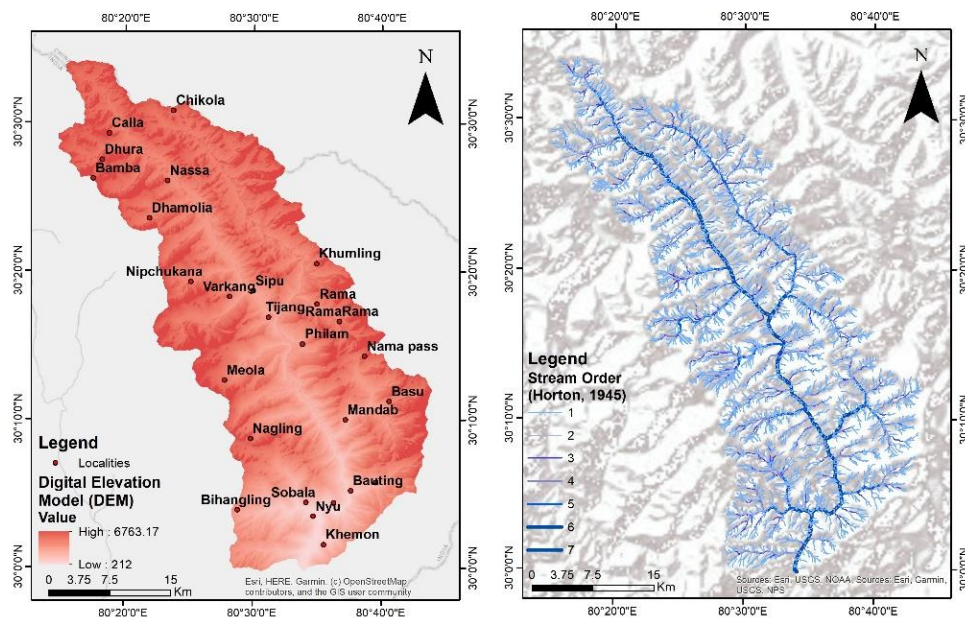


Fig. 1. Shows the elevation and detailed drainage network of the Dhauliganga

The Dhauliganga is a seasonal stream, primarily sustained by monsoonal rainfall, snowmelt, and springs, with a dendritic drainage pattern and rapid runoff. Geomorphic features such as river terraces, landslide scars, and knickpoints indicate active fluvial incision and neotectonic deformation. The region is highly vulnerable to landslides, flash floods, and soil erosion, particularly during the monsoon season.

• 3. Data & Methodology

In the current study, the neo-tectonic study of the Dhauliganga River has been carried out by using satellite data. The Advanced Space-borne Thermal Emission and Reflection Radiometer (ASTER) data were used to create the Digital Elevation Model (DEM), further, it was used to identify the drainage network and watersheds. Further, 1:2500000 scale topographic maps were obtained from Army Map Services, which publishes them on behalf of the University of Texas at Austin, the organization in charge of publishing and distributing them. The downloaded maps were georeferenced to the WGS84 datum and UTM projection. The locations of towns, cities, and villages in the Dhauliganga river basin were vectorized using these georeferenced maps. The map was georeferenced to the WGS84 datum and UTM projection to be consistent with the ASTER DEM and multispectral data used in this work. Geological data was also gathered using published geological maps of the Kumaon Himalaya that could be georeferenced because they had a geographic referencing coordinate system.

Table 01 shows the morphotectonic parameters and their method of calculation with references.

Indices	Formula	References
Drainage Density (Dd)	$Dd = Lu / A$ where Lu is the length of the stream of all orders and A is the area of the basin.	Horton, 1932
Drainage Texture (Dt)	$Dt = Nu / P$ where Nu is the total number of streams and P is the perimeter of the basin.	Horton, 1932
Hypsometric Integral (HI)	$HI = (EL_{mean} - EL_{min}) / (EL_{max} - EL_{min})$ where EL_{mean} is the mean elevation, EL_{min} is the minimum, and EL_{max} the maximum elevation	Strahler, 1952
Bifurcation Ratio (Br)	$Br = Lu / Lu+1$ where Lu is the number of specified streams and a number of next higher-order streams in basin.	Strahler, 1964
Asymmetry Factor (AF)	$AF = Ar / A$ where Ar is the area of right side of the basin and A is the total area of basin.	Cox et. al., 1994
Valley Floor Width to Height Ratio (Vf)	$Vf = 2V_{fw} / (E_{ld} - E_{sc}) + (E_{rd} - E_{sc})$, where V_{fw} = width of valley floor, E_{ld} and E_{rd} = Elevation of left and right valley divides, E_{sc} = Elevation of the valley floor	Bull & McFadden, 1977
Stream Length Gradient Index (SL)	$SL = \Delta h / \Delta l \times L$, where Δh = Difference in elevation of the ends of the reach, Δl = Length of reach, L = distance from the midpoint of reach to the most distant point up-stream.	Hack (1973)
Mean Stream Length (MSL)	$MSL = Lu / Nu$, Lu = Total length of all streams in a given order, Nu = Number of stream segments of that order.	Strahler (1964)

• 4. Results and Discussion

In this study, relative tectonic activity was assessed using a set of geomorphic indices. These indices were categorized into three levels of tectonic activity: low, moderate, and high. To quantify overall tectonic influence, an Index of Relative Active Tectonic Activity (IRAT) was developed by ranking individual basins based on each geomorphic parameter. For instance, a basin exhibiting low tectonic activity in terms of the Hypsometric Integral was assigned a rank of 1, while moderate

and high activity levels received ranks of 2 and 3, respectively. The final tectonic classification for each basin was determined by averaging the ranks across all indices.

4.1. Drainage Density

Drainage density, defined as the ratio of total stream length to basin area (Horton, 1932), is a key morphometric index reflecting drainage texture and basin evolution. Low values indicate permeable, vegetated, and gently sloping terrains, while high values are typical of steep, rugged, and less permeable regions.

Drainage density was calculated for all 30 sub-basins and classified into three categories. The Dhauliganga basin, intersected by multiple faults and thrusts (see geological map), shows clear tectonic influence. High drainage density in the central-western region suggests active tectonics, while lower values in the northern and south-central areas indicate reduced tectonic impact and possibly more stable conditions.

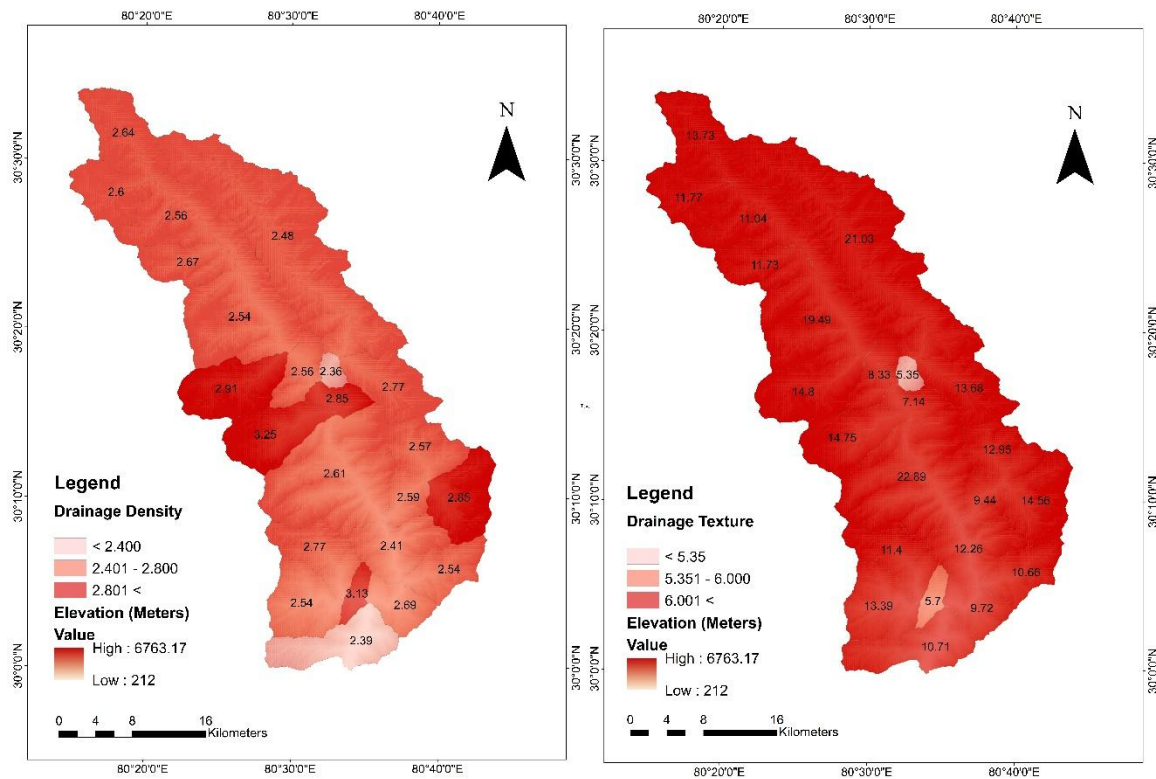


Fig. 2. Displays the spatial distribution of drainage density and drainage texture across Dhauliganga

4.2. Drainage Texture

Drainage texture, defined as the ratio of the total number of stream segments of all orders to the basin perimeter (Horton, 1945), is a key indicator of surface runoff characteristics. It is strongly influenced by the basin's infiltration capacity, which in turn is governed by factors such as climate, rainfall intensity, vegetation cover, lithology, soil type, topography, and the geomorphic evolution of the basin (Smith, 1950).

The results for the 30 sub-basins of the Dhauliganga River Basin were categorized into three drainage texture classes. High values, indicating fine texture, were primarily observed in the central region and northeastern margins, suggesting low infiltration and high runoff. Low values, representing coarse texture, were associated with high infiltration and reduced surface runoff. Sub-

basins intersected by the main trunk stream exhibited sparse drainage texture.

4.3. Hypsometric Integral

The hypsometric integral (HI) is a dimensionless metric that allows scale-independent comparison of watersheds. High HI values indicate youthful, less eroded terrains; moderate values reflect mature landscapes; and low values signify highly eroded, late-stage topography. Earlier, HI calculation was limited by computational constraints and high processing demands. However, advancements in digital elevation models (DEMs) have made HI analysis more efficient and accurate for identifying tectonically active regions (Dowling et al., 1998; Khan & Govil, 2023).

The 30 sub-basins exhibit hypsometric integral (HI) values ranging from 0.27 to 0.58, reflecting moderate to high landscape maturity. Sub-basins with lower HI values indicate advanced erosion and older topography, particularly in the southernmost region. Sub-basins on the right bank of the trunk stream show moderate HI values, suggesting a dynamic balance between tectonic uplift and erosional processes. In contrast, the left bank displays higher HI values, indicating more active

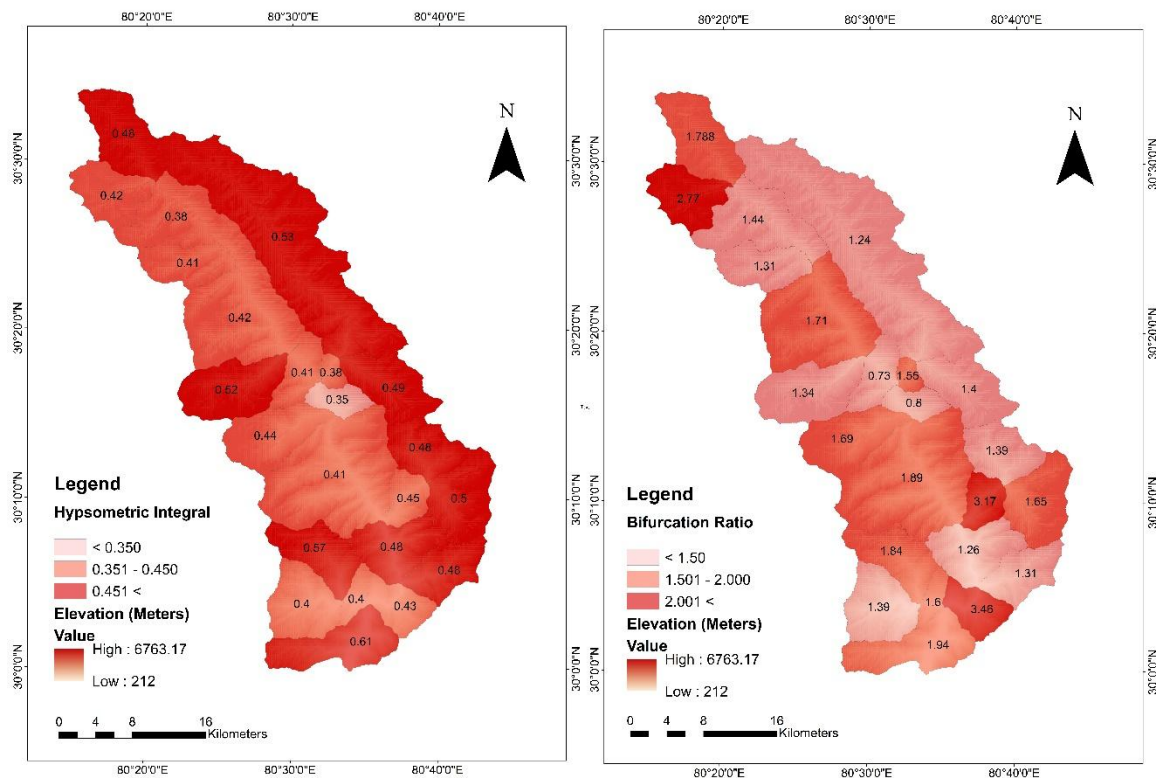


Fig. 3. Exhibits the spatial distribution of Hypsometric Integral & Bifurcation Ratio

tectonic uplift driven by strong tectonic forces.

4.4. Bifurcation Ratio

The bifurcation ratio (Rb), defined as the ratio of the number of streams of a given order to those of the next higher order (Schumm, 1956), serves as an important indicator of basin relief and degree of dissection (Horton, 1945). While Rb values typically exhibit limited variation under different environmental conditions, notable deviations occur in regions where geological structure exerts strong control.

The results show a wide range of bifurcation ratios (Rb), varying from 0.80 to 3.46, indicating significant variability across the sub-basins. Lower Rb values are associated with gently

rolling, low-elevation terrains, whereas higher values suggest strong tectonic influence on basin development. Additionally, Eze and Effiong (2010) noted that elevated Rb values are indicative of a lower susceptibility to flash flooding within the basin.

○ 4.5. Asymmetry Factor

The asymmetry factor (AF) is a key morphotectonic indicator used to detect tectonic tilting within a drainage basin (Cox et al., 1994; Keller and Pinter, 2002; El Hamdouni et al., 2008). It operates on the assumption that a river flowing through a lithologically homogeneous and tectonically stable terrain will exhibit a symmetric drainage pattern. Deviations from this symmetry suggest potential tectonic or erosional influences.

The asymmetry factor (AF) values for the 23 sub-basins of the Dhauliganga River range from 24.01 to 79.67 (Figure XX). To assess tectonic significance, these values were grouped into three classes based on the degree of tilting. Sub-basins with near-symmetric drainage were assigned a low rank, while those exhibiting high asymmetry were ranked higher, indicating stronger tectonic influence. Basin tilting and asymmetry tend to increase in tectonically active areas, where enhanced erosional processes also occur. At a regional scale, tectonic activity is reflected by the tilting and lateral migration of higher-order streams within active river basins, whereas localized tectonic activity is indicated by elevated asymmetry in individual sub-basins (Keller and Pinter, 2002).

○ 4.6. Valley Floor Width to Height Ratio

Bull and McFadden (1977) introduced the valley floor width-to-height ratio as a metric to assess active tectonics by differentiating between open and broad-floored valleys. Tectonically active areas typically exhibit significant valley variation, including both V-shaped and U-shaped profiles, reflecting a dynamic equilibrium. Silva et al. (2003) further distinguished broad, U-shaped valleys, formed mainly through lateral erosion during tectonic stability whereas V-shaped valleys are the product of strong tectonic influence due to active incision.

The Dhauliganga River flows north to south, merging with the Kali River, and passes through diverse lithological and structural zones. These variations influence the river's morphological shift from narrow V-shaped valleys to broader U-shaped ones. Valley floor width-to-height ratio (Vf)

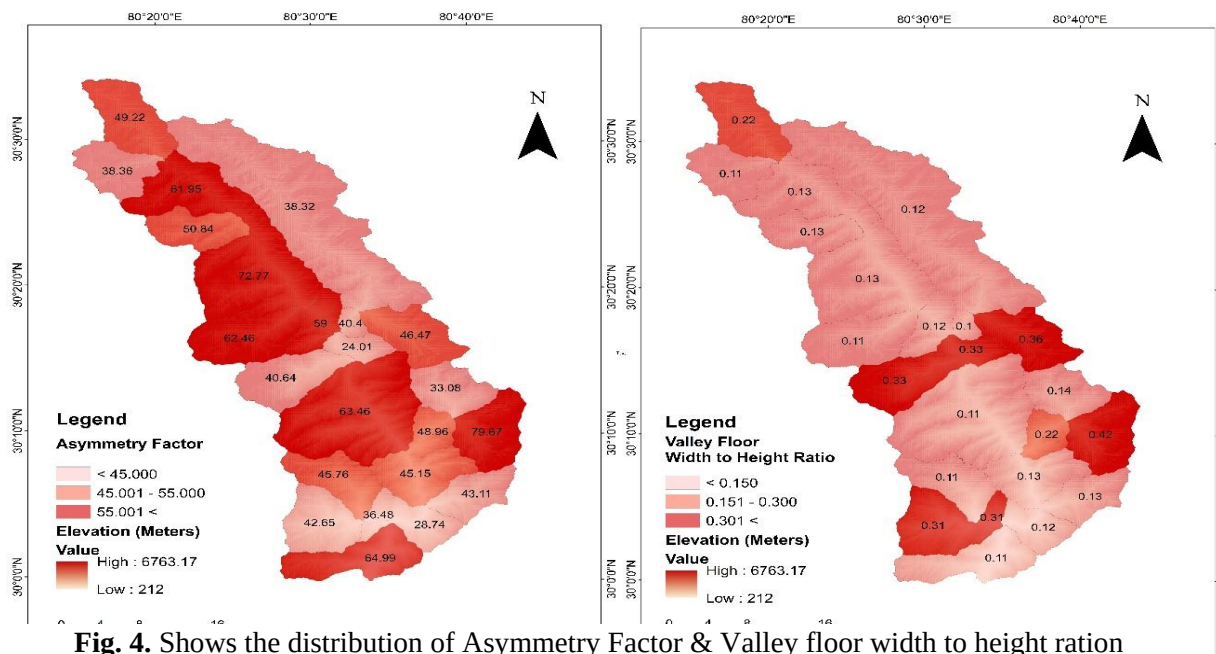


Fig. 4. Shows the distribution of Asymmetry Factor & Valley floor width to height ration

values, ranging from 0.10 to 0.42, reflect strong structural influence across the basin. Sub-basins intersecting the main channel are shaped more by erosional processes, while those aligned perpendicular to the trunk stream show stronger tectonic control.

- **4.7. Stream Length Gradient**

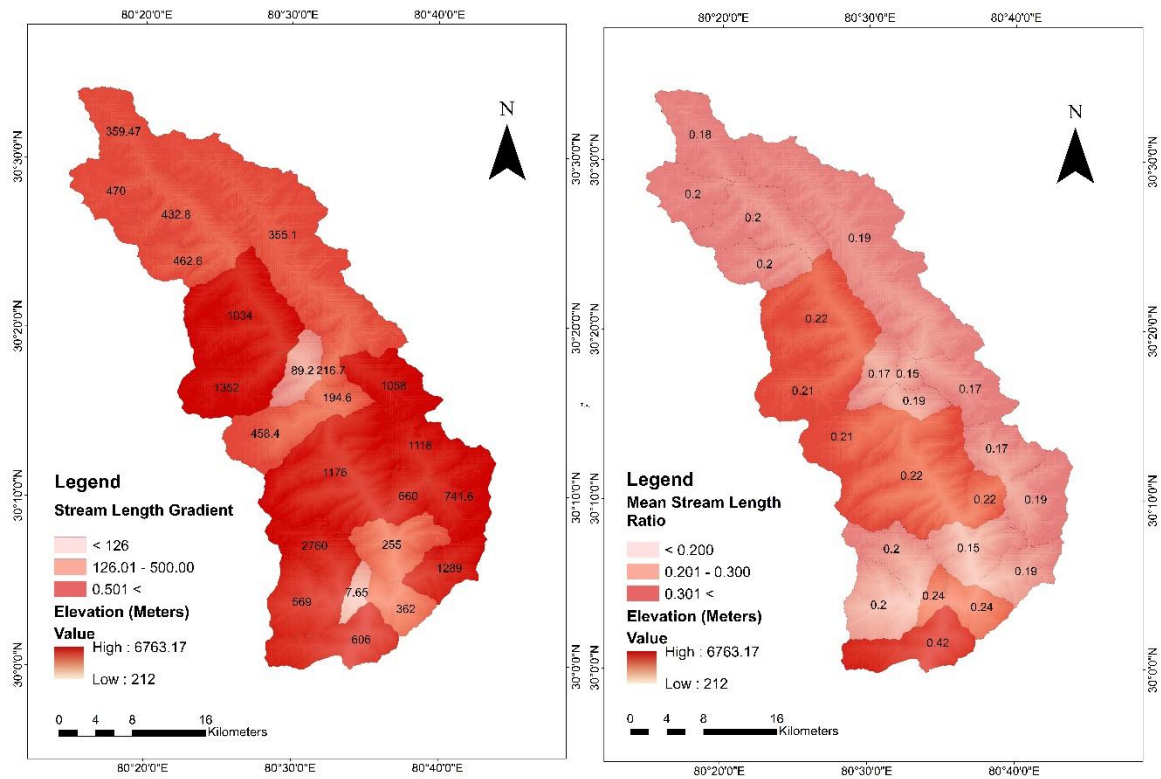
Stream gradient provides insights into the extent of soil and rock erosion within a river basin, which in turn shapes the surrounding landscape (Hack, 1973). In neotectonically active regions, erosion rates influence uplift, often producing a slightly concave longitudinal river profile indicative of relative crustal stability (Schumm et al., 2002). However, this stability is also affected by tectonic activity, lithological conditions, and climatic factors.

Stream length gradient was calculated for 23 sub-basins within the Dhauliganga River Basin. The upper catchment in the northern region exhibits moderate gradient values, primarily due to limited fluvial activity and the dominance of glacial terrain. In contrast, the central and southern regions show higher gradient values, indicating a strong influence of tectonic activity.

- **4.8. Mean Stream Length**

Mean Stream Length (Lsm) is a key morphometric parameter that represents the average length of stream segments within a specific stream order. It is calculated by dividing the total stream length by the number of segments of that order. Introduced by Horton (1945) and refined by Strahler (1964), Lsm aids in analyzing drainage network structure and basin development.

The mean stream length ratio indicates a systematic relationship between drainage development and surface characteristics, which in this study is interpreted as a reflection of tectonic activity. The calculated values range from 0.15 to 0.42 and were classified into three categories to demonstrate the influence of active tectonics and erosion on basin evolution. Sudden increases or irregular variations in this ratio across stream orders may suggest neotectonic activity, faulting, or lithological transitions.



A
more

Fig. 5. Displays the spatial variation of stream length gradient & mean stream length

robust evaluation is achieved by integrating multiple morphotectonic indices (Dowling et al., 1998; Ali & Iqbal, 2020). This study assesses active tectonics over a moderately sized area (1303 km²), divided into 23 sub-basins, using 8 morphotectonic indices. Each sub-basin was ranked from 1 to 3 based on the value of each index. A total score of 24 indicates the highest tectonic activity, while a score of 11 reflects the lowest. The spatial variation in tectonic activity across the Dhauliganga River Basin was visualized through a composite Index of Relative Active Tectonics (IRAT) map. Out of the 23 sub-basins in the Ramganga region, eight exhibit strong tectonic activity, while the remaining 15 fall into the moderate activity category. None of the sub-basins scored 11 or below, thus none qualify as having low tectonic activity. The sub-basins with strong tectonic influence are primarily located along the right bank of the Dhauliganga River and intersect with the trunk stream. This suggests that proximity to the main stream enhances structural control over basin evolution. Additionally, sub-basins in the southern extremity of the Dhauliganga display high tectonic activity, indicating a dominance of tectonic forces over erosional processes.

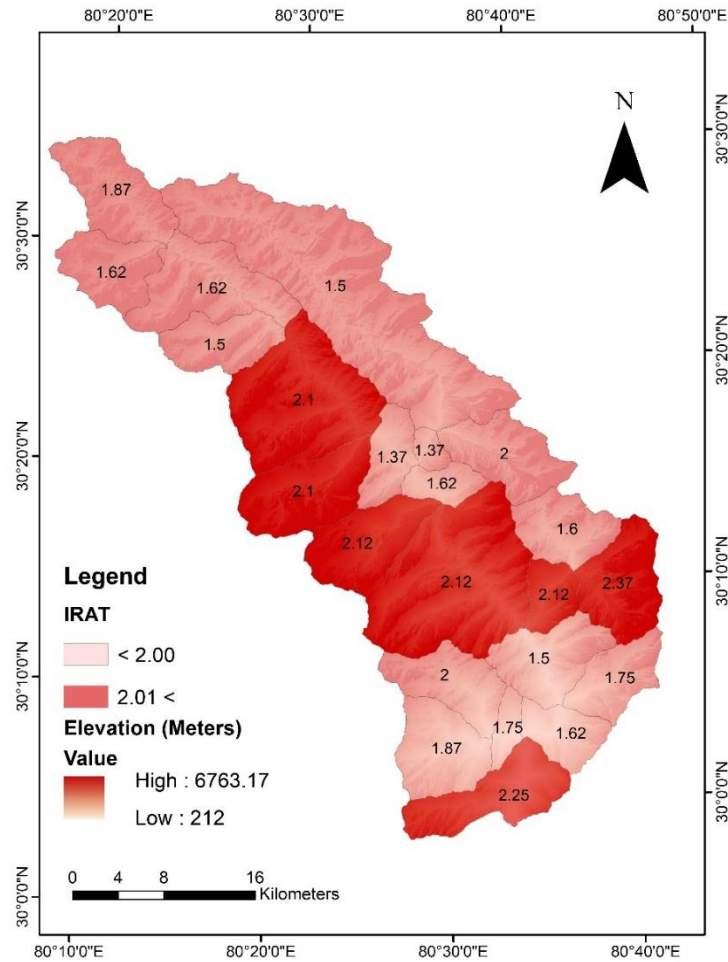


Fig. 6. Demarcates the spatial distribution of neo-tectonics based on IRAT

• 5. Conclusion

Morphotectonic analysis reveals significant neotectonic influence and strong erosional processes in the Dhauliganga River basin. The Dhauliganga basin also shows clear signs of active tectonics, with thrusts, faults, and high stream length gradients driving intense incision along the main channel. The northern basin, intersecting the Trans-Himadri Fault and Main Central Thrust, is shaped by alpine glacial activity, forming broad, U-shaped valleys that are distinct from fluvial systems. The central region reflects dominant tectonic control, while the southern extremity exhibits moderate activity, as indicated by IRAT, suggesting a balance between tectonic and erosional forces. Tectonic deformation has notably altered the Dhauliganga River's course, and morphotectonic indices from remote sensing and GIS effectively delineate the spatial extent and intensity of tectonic modification.

Disclosure of Interests

The Authors have no competing interests to declare that are relevant to the content of this article.

References

- Gansser, A. Geology of the Himalayas. Wiley Inter Science, New York, 289, (1964).
- McKenzie, D. Sclater, J. G. The Evolution of the Indian Ocean since the Late Cretaceous. *Geophysical Journal International*. 24 (5): 437, (1971).
- Molnar, P. & Tapponnier, P. Cenozoic tectonics of Asia: Effects of a continental collision; *Science* 189: p419-426, (1975).
- Dewey, J.F. & Bird, J.M. Mountain Belts and the New Global Tectonics. *J. Geophys. Res.* 75 2625-2647, (1970). <https://doi.org/10.1029/JB075i014p0262513>.
- Khan M.N and Govil H. Assessment of On-going tectonic deformation in the Goriganga River Basin, Eastern Kumaon Himalaya Using Geospatial Technology. *Journal of Applied and Natural Science*, 15(4), 1679 – 1690, (2023).
- Kalpana. G. Kothiyari, G.C. & Kotlia, B.S. Mor-photectonic assessment of the Gaula river basin, Kumaun Lesser Himalaya, Uttarakhand. *Quaternary Science Advances*, 12, 100- 115, (2023). <https://doi.org/10.1016/j.qsa.2023.100115.24>.
- Farooq, S. Khan, M.N. & Sharma, I. Assessment of Active Tectonics in Eastern Kumaon Himalaya on the Basis of Morphometric Parameters of Goriganga River Basin, *IJAEEs*, 3(3), 14-21, (2015).
- Khan, M. N, M. Suhail, Dilawez Ali, S.M.W.H.Jafri, J.Jurayev. Neotectonic appraisal of Ramganga River basin, Eastern Kumaun, Himalaya, India. *Journal of Applied and Natural Science*, 16(4), 1690 – 1701, (2024). <https://doi.org/10.31018/jans.v16i4.5928>
- Keller, E. A. & Pinter, N. Active Tectonics: Earthquakes, Uplift, and Landscape. Prentice Hall, (2002).
- Omar M.A. Radaideh, Jon, M. Tectonics controls on fluvial landscapes and drainage development in the westernmost part of Switzerland: Insights from DEM derived geomorphic indices, *Tectonophysics* 768, (2019).
- Burbank, D.W. & Anderson, R.S. Tectonic Geomorphology, 2nd Edition, Malden USA, Blackwell Publishing, 274 pp, (2011).
- Khan, M.N., Khudoyarova S.S., Juraev J., Mamajanov R.I. Field-based Tectonic Assessment and Spatial Correlation with Land Use and Land Cover in the Goriganga River Basin, *Bulletin of Pure and Applied Sciences- Geology*, 42 (1), 32-45, (2023). <https://doi.org/10.48165/bpas.2023.42F.1.4>
- Taib, H., Hadji, R., Hamed, Y. Exploring neotectonic activity in a semiarid basin: a case study of the Ain Zerga watershed. *Journal Umm Al-Qura Univ. Applied. Science*, (2023). DOI: <https://doi.org/10.1007/s43994-023-00072-3>
- Willett, S.D., Hovius, N., Brandon, M.T., and Fisher, D.M. Introduction, in Willett, S.D., Hovius, N., Brandon, M.T., and Fisher, D.M. (Eds). *Tectonics, Climate, and Landscape Evolution*. Geological Society of America Special Paper 398, Penrose Conference Series, vii-xi, (2006). DOI: 10.1130/2006.2398(00).
- Turner, A. Geomorphometrics: Ideas for Generation and Use. CCG Working Paper, Version 0.3.1. Centre for Computational Geography, University of Leeds, UK, (2006).
- Wallace, R.E. The San Andreas Fault System, California: US Geological Survey Professional Paper 1515, United States Government Printing Office, Washington, DC, 283 pp, (1990).
- Strahler, A. N. Hypsometric (area-altitude) analysis of erosional topography. *Geological Society of America Bulletin*, 63(11), 1117–1142, (1952).

Cloud Orchestration in SCM for Cloud Efficiency

Mohan V Datar, Tanuja Pattanshetti

COEP Technological University, Pune, India

datarmv23.comp@coeptech.ac.in, trp.comp@coeptech.ac.in

Abstract—Cloud orchestration has emerged as a pivotal mechanism for managing complex cloud environments, enabling the seamless automation and optimization of resource allocation, workload distribution, and service deployment. This paper presents a comprehensive study on cloud orchestration for optimized cloud efficiency, focusing on a supply chain use case involving backorder prediction. Leveraging datasets from Kaggle, we propose a methodology that integrates machine learning classifiers with cloud orchestration tools to predict and mitigate backorders, thereby enhancing inventory system performance. The study evaluates the efficacy of AWS Step Functions, Azure Data Factory, and Google Cloud workflows in orchestrating data pipelines and services across major cloud ecosystems. Key performance indicators, such as error handling, scalability, cost-effectiveness, and integration capabilities, are analyzed to compare these platforms. Results demonstrate that AWS Step Functions provide superior flexibility and cost-efficiency, particularly for small and medium enterprises (SMEs). This research bridges the gap between theoretical orchestration frameworks and practical business applications, offering insights into optimizing cloud resources for supply chain management. Future work aims to extend the model to incorporate real-time event-driven orchestration and advanced AI-driven forecasting techniques.

Index Terms—Cloud Orchestration, AWS Step Functions, Machine Learning

1. Introduction

The rapid evolution of cloud computing has transformed the way organizations manage their IT infrastructure, offering unprecedented scalability, flexibility, and efficiency. Central to this transformation is cloud orchestration, a process that automates the management, coordination, and deployment of cloud resources and services to achieve specific business objectives. Unlike traditional manual approaches, cloud orchestration enables end-to-end automation of workflows, integrating disparate systems across public, private, and hybrid cloud environments. This capability is particularly critical in today's data-driven landscape, where businesses face increasing pressure to optimize resource utilization, reduce operational costs, and respond swiftly to dynamic market demands. One such domain where cloud orchestration holds significant potential is supply chain management, particularly in addressing the challenge of backorders. A backorder occurs when a product is sold but not immediately available in stock, with delivery promised at a future date. While backorders are common in industries like automotive and retail, excessive delays can lead to customer dissatisfaction, order cancellations, and reputational damage. The ability to predict and mitigate backorders before they occur can substantially improve service levels and operational efficiency. However, achieving this requires sophisticated data processing, accurate demand forecasting, and seamless coordination of resources—tasks ideally suited for cloud orchestration.

This paper explores the application of cloud orchestration in optimizing cloud efficiency for a supply chain use case focused on backorder prediction. Building on datasets sourced from Kaggle, we propose a methodology that combines machine learning techniques with orchestration frameworks from leading cloud providers: Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP). Specifically, we leverage AWS Step Functions, Azure Data Factory, and GCP workflows to design and evaluate data pipelines and service orchestration models. The study aims to address key research gaps identified in existing literature, such as the

lack of practical implementations linking orchestration tools to business-specific workflows and the need for comparative analyses of cloud platforms based on measurable performance metrics. The remainder of this paper is organized as follows: Section II reviews the existing literature on cloud orchestration and related optimization techniques. Section III outlines the proposed methodology, including the backorder prediction use case and orchestration architecture. Section IV presents the results and evaluation of the implemented solutions across AWS, Azure, and GCP. Finally, Section V concludes the paper and discusses avenues for future work.

2. Literature review

Cloud orchestration has garnered significant attention in recent years as organizations seek to streamline operations in increasingly complex cloud environments. This section reviews key studies and frameworks related to cloud orchestration, optimization techniques, and their application to resource management and business workflows.

Authors in [1] proposed a novel orchestration framework for optimizing cloud efficiency in the context of wind farm resource evaluation. Their work emphasizes deployment automation and cost reduction, leveraging cloud infrastructure to minimize manual intervention. The study highlights the potential of orchestration tools to abstract provider-specific differences, a recurring theme in multi-cloud research. However, its focus remains domain-specific, limiting broader applicability to general business use cases.

This paper [2] addressed resource provisioning cost optimization in cloud computing environments. Their approach uses an Optimal Cloud Resource Provisioning (OCRP) algorithm to calculate costs across reservation, spending, and demand phases, spilling excess workloads to public clouds when private infrastructure is insufficient. Performance evaluations underscore the importance of balancing total cost of ownership (TCO) with workload outsourcing, though the study lacks integration with real-time orchestration tools.

Citation [3] investigated security concerns in cloud orchestration, proposing a range of encryption techniques, including Attribute-Based Encryption (ABE) and Homomorphic Encryption. Their work identifies a critical gap in analyzing potential security threats alongside comprehensive countermeasures. This gap is particularly relevant for orchestration frameworks handling sensitive business data, such as supply chain inventories.

In this paper [4] explored reliability optimization in public cloud systems, analyzing faults like request queue overflows and virtual machine failures. They proposed a dynamic scheduling technology to ensure service reliability, with future directions pointing toward container-based resource management and block chain integration. While robust, their study does not address practical workflow orchestration across multiple providers.

Researchers [5] explored reliability optimization in public cloud systems, analyzing faults like request queue overflows and virtual machine failures. They proposed a dynamic scheduling technology to ensure service reliability, with future directions pointing toward container-based resource management and block chain integration. While robust, their study does not address practical workflow orchestration across multiple providers.

Researchers in [6] tackled workflow scheduling in fog-cloud environments using Ant Colony Optimization (ACO). The approach optimizes resource utilization and execution time, outperforming traditional algorithms in terms of makespan and cost. However, it lacks a detailed examination of orchestration tools' role in facilitating such optimization, a gap this paper seeks to address.

Researchers in [7], [8] Additional studies highlight specific challenges and solutions in cloud orchestration. For instance, research on multi-cloud orchestration [7] emphasizes flexibility and resilience through advanced tools, while declarative resource orchestration [8] focuses on policy-driven resource allocation. Studies on cloud-edge orchestration for IoT [9] and spot market orchestration [10] further underscore the diversity of orchestration applications, yet they often lack empirical validation in business contexts like supply chain management.

Researchers in [11] Selected open source three tools Brooklyn, Cloudify, Stratos more focus on the deploying and operating a standard three-tier application

The literature reveals several research gaps: (1) limited integration of orchestration tools with practical, industry-specific workflows; (2) insufficient comparative analyses of major cloud platforms (AWS, Azure, GCP) based on performance metrics; with respect to small and medium enterprises and (3) a focus on theoretical frameworks over actionable implementations. This paper addresses these gaps by applying cloud orchestration to a backorder prediction use case, leveraging machine learning and cross-platform orchestration tools to enhance supply chain efficiency.

3. Proposed methodology

This section outlines the proposed methodology for applying cloud orchestration to optimize cloud efficiency in a supply chain use case focused on backorder prediction. The approach integrates machine learning classifiers with cloud orchestration tools from AWS, Azure, and GCP to predict and mitigate backorders, enhancing inventory system performance. The methodology comprises three main components: problem definition, data processing and model development, and orchestration architecture design.

A. Problem Definition

The study addresses a common supply chain challenge: material backorders, which occur when demand exceeds supply, leading to delayed deliveries. Backorders impact service levels and customer satisfaction, making their prediction and prevention critical for operational efficiency. We utilize a publicly available dataset from Kaggle [21], consisting of two CSV files: Training BOP.csv and Testing BOP.csv. These files contain features such as current inventory levels (national inv), transit times (lead time), forecast sales (forecast x month), past sales (sales x month), minimum stock recommendations (min bank), and risk flags (X17-X22). The target variable, went on backorder, indicates whether a product entered a backorder state. The objective is to train classifiers to predict future backorders and orchestrate cloud resources to operationalize these predictions.

B. Data Processing and Model Development

Data preprocessing and model development were conducted using Python-based libraries. The process involved the following steps:

- 1) Data Cleaning: Missing values were imputed using the median values for numerical features (e.g., sales x month) and mode for categorical features (e.g., potential issue). Outliers were detected and capped using the interquartile range method.
- 2) Feature Engineering: New features were derived, such as the ratio of forecast sales to current inventory and the difference between minimum recommended stock and actual inventory, to enhance predictive power.
- 3) Model Training: Several classifiers were evaluated, including Logistic Regression, Random Forest, and Gradient Boosting, using Scikit-learn [12]. The dataset was split into training (70%), validation (15%), and test (15%) sets, with hyperparameter tuning performed via grid search. Performance metrics included accuracy, precision, recall, and F1-score, with a focus on recall to prioritize backorder detection.

- 4) Scalability: For large-scale processing, Dask [13] was employed to handle data in chunks, ensuring efficient computation on cloud resources.

The trained model was serialized for deployment within cloud orchestration workflows.

C. Orchestration Architecture Design

The orchestration architecture was designed to deploy the backorder prediction model across AWS, Azure, and GCP ecosystems, leveraging their respective tools: AWS Step Functions, Azure Data Factory, and GCP Workflows. The design process involved:

1) Workflow Type Selection: Workloads were classified as event-driven (e.g., real-time inventory updates triggering predictions) and time-driven (e.g., daily demand forecasting reports). A hybrid approach was adopted to balance real-time responsiveness and scheduled analytics.

2) AWS Implementation: AWS Step Functions coordinated a multi-step workflow, integrating Amazon S3 for data storage, AWS Lambda for model inference, and Amazon SES for notifications. The "Redrive" feature was utilized with Amazon SQS to handle execution failures, sending failed tasks to a dead-letter queue for analysis and retry. We can use extensively terra form for the deployment will provide easy turn around time for the deployment. S3 extract path step generates arguments for respective stream's lambda function, compare flow name initialize the ingestion or extraction from test to train data. The step fetches the statistical data for ingestion or extraction (no of rows). Wait status will check existing status and based on either running / completed / failed trigger Re check dbt status wait. Step Function will generate the Run id . we can create Message queue as well based on the data processing at several layers will notify users using SQS email triggering . We can Audit trail the required events and status. Sub System Workflow Azure Implementation: Azure Data Factory managed data pipelines, extracting data from Azure Blob Storage, processing it via Azure Functions, and loading results into a reporting database. Custom error-handling features were configured to ensure robustness. GCP Implementation: GCP Workflows orchestrated tasks using Cloud Storage, Cloud Functions, and Big Query for analytics. Workflow automation was achieved through predefined tem-

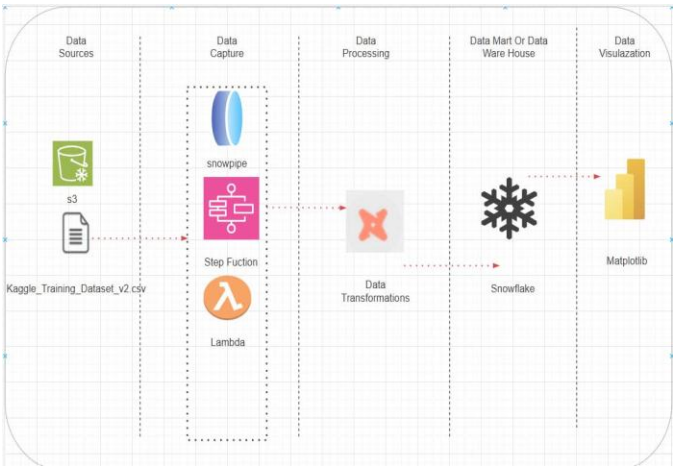


Fig. 1. AWS Architecture for SCM

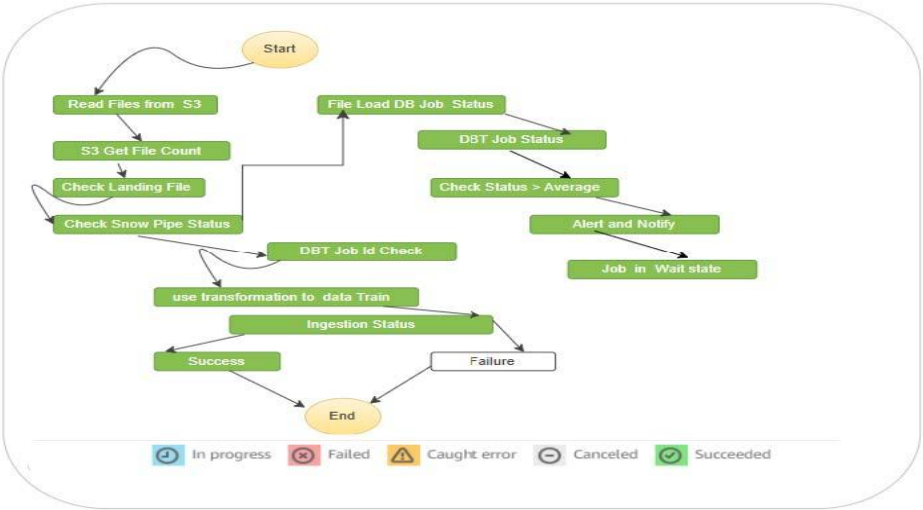


Fig. 2. System Work Flow

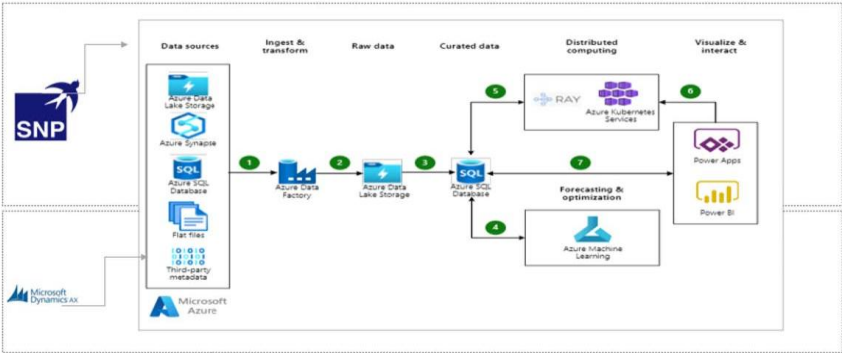


Fig. 3. Azure Architecture for SCM

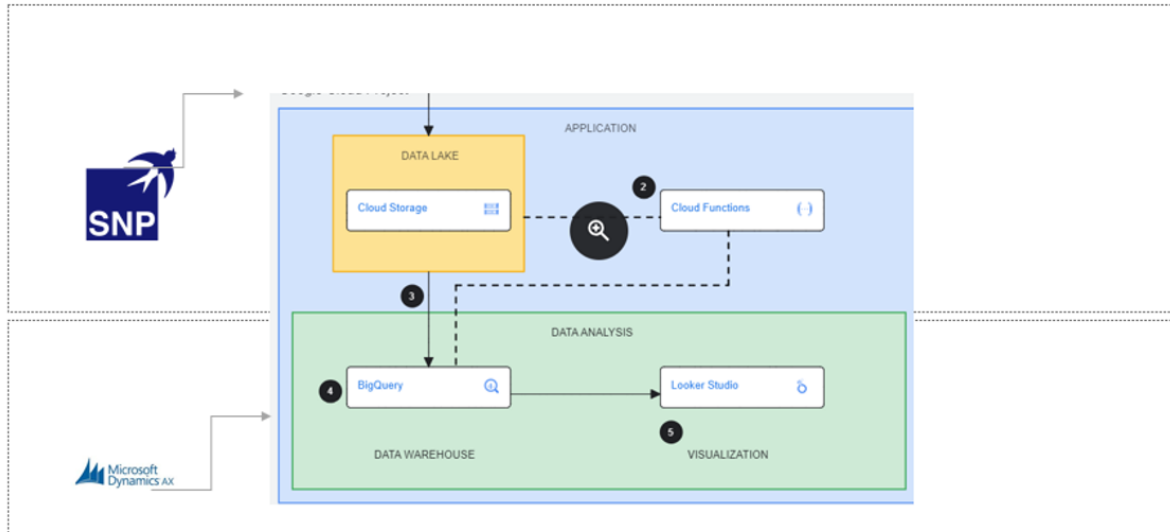


Fig. 4. GCP Architecture for SCM

plates, though customization was limited compared to AWS and Azure.

Key Performance Indicators (KPIs): The architectures were evaluated based on 13 metrics, including built-in error handling, automatic scaling, execution event history, integration capabilities, cost-effectiveness, and tailored solutions for SMEs. These metrics were measured using execution logs and cost analyses provided by each platform, execution time, quick turnaround time for the custom requirements. The implementation relied on the following tools:

- **Python Libraries:** Pandas, NumPy, Scikit-learn, Tensor-Flow, Matplotlib, Seaborn, and Dask for data processing and modeling, Data Build Tool DBT.
- **Development Environment:** Jupyter Notebook for interactive analysis, with PyCharm for larger-scale development. A GPU system with CUDA cores was used for accelerated training.
- **Cloud Services:** AWS (Step Functions, S3, Lambda, SES, SQS), Azure (Data Factory, Blob Storage, Functions), and GCP (Workflows, Cloud Storage, Functions, BigQuery).

This methodology bridges theoretical orchestration concepts with a practical business application, providing a scalable framework for backorder prediction and mitigation.

IV. Results and Evaluation

This section presents the results of implementing the proposed cloud orchestration framework for backorder prediction across AWS, Azure, and GCP ecosystems. The evaluation focuses on the performance of the machine learning model, the effectiveness of the orchestration workflows, and a comparative analysis of the cloud platforms based on predefined Key Performance Indicators (KPIs).

A. Machine Learning Model Performance

The backorder prediction model was trained and validated using the Kaggle dataset [21]. Among the classifiers evaluated—Logistic Regression, Random Forest, and Gradient Boosting—the Random Forest model achieved the best performance. Key metrics on the test set are summarized in Table 1.

Table I: Model accuracy for sentiment and emotion classification

Metric	Logistic Regression	Random Forest	Gradient Boosting
Accuracy	85	91	89
Precision	78	87	84
Recall	72	89	86
F1-Score	75	88	85

The Random Forest model's high recall (0.89) indicates its effectiveness in identifying backorder instances, critical for minimizing stock shortages. Feature importance analysis revealed that forecast 9month, national inv, and min bank were the most influential predictors, aligning with supply chain dynamics where demand forecasts and inventory levels drive backorder risks.

B. Orchestration Workflow Execution The orchestration workflows were deployed to process the dataset (approximately 1.5 million records) and generate predictions in a simulated supply chain environment. Execution results for each platform are detailed below:

1) **AWS Step Functions:** The workflow and the Subsystem workflow depicts demonstrated robust error handling and scalability. it involves tasks such as data ingestion start from s3 as landing layer tp snowflake using snow pipes , transformations using data build tool DBT. Step Functions are event driven and server less, cost model based on the invocation and duration, it can be integrated with various AWS services like Glue, SQS , event bridge etc. They have low latency for handling events real time .Triggered by Events and Schedules , Customization is possible

2) **Azure Data Factory:** ADF is extensively used for design and execution for the Data Pipe lines for the Azue ECO Systems under considerations we have used folowing set of services like creation of Resource group , storage account, Azure Blob storage ,Azure Data lake storage . Custom error-handling features required manual configuration but ensured reliability. Integration with Azure Functions and Blob Storage facilitated seamless data flow, though execution event history required enabling additional logging, increasing setup complexity.

3) **GCP Workflows:** For the GCP Eco system we have used following set of the services like Cloud Storage for Automate secured data movement from data sources into Big Query on a scheduled or on-demand basis , looker for data visualization . Pre-built templates accelerated deployment, but limited customization options constrained adaptability to specific use case requirements. Integration with Big Query provided efficient analytics, though multi-step workflow support was less comprehensive than AWS and Azure.

Table II: Execution time for the data ingestion and transformation

Metric	AWS DBT	Azure ADF	GCP
Execution Time	6 Mins 26 Sec	7 Min 12 Sec	6 Minutes 48 Sec

TABLE III: KEY METRICS COMPARISON

KPI	AWS	Azure	GCP
Built-in Error Handling	Robust	Custom	Limited
Automatic Scaling	Yes	Yes	Yes
Execution Event History	Built-in	Configurable	Configurable
High Availability	Yes	Yes	Yes
Administrative Security	High	High	High
No-Code Workflow	Flexible	Yes	Yes
Integration Capabilities	Extensive	Moderate	Limited
Multi-Step Workflows	Yes	Yes	Limited
Pre-Built Templates	Customizable	Yes	Yes
Task History/Monitoring	Yes	Yes	Yes
Cost-Effectiveness	High ROI	Average	Lower Cost
Custom Code Integration	High	Moderate	Limited
SME Tailored Solutions	Yes	No	No

- Lowest cost, thanks to free Big Query ingestion and minimal compute (\$0.005).
- Ideal for server less analytics and visualization (Looker Studio).

b) Recommendation:

- GCP is the most cost-effective for a single mid-size CSV run with basic analytics, leveraging Big Query's efficiency and free ingestion.
- Azure offers a balanced option if you need data lake capabilities and orchestration flexibility.
- AWS (Snowflake) suits scenarios requiring advanced warehousing or frequent ingestion, but it's overkill for a one-off 1 GB CSV.

Table IV: Cost comparison

Platform	Storage Cost	Orchestration Cost	Total Estimated Cost
AWS	\$0.046	\$0.173	\$0.219
Azure	\$0.037	\$0.108	\$0.145
GCP	\$0.040	\$0.005	\$0.045

The platforms were evaluated against 13 KPIs, with results summarized in Key Metrics for comparison. Scores were assigned based on execution logs, cost analyses, and qualitative assessments (e.g, ease of use, flexibility).

Metric to measure for Cloud Orchestration for Cloud Platforms under study.

- AWS: Excelled in flexibility, cost-effectiveness (optimized ROI due to pay-as-you-go pricing), and integration within its ecosystem (e.g., S3, Lambda). The subsystem workflow and state chart highlight its ability to manage complex, multi-step processes efficiently.
- Azure: Offered robust customization and integration with external tools, though at the cost of increased setup time. It performed well for SMEs with existing Microsoft ecosystems but lacked tailored solutions for smaller-scale deployments.
- GCP: Provided a straightforward setup with pre-built templates, suitable for rapid prototyping. However, its higher operational costs and limited multi-step workflow support made it less competitive for this use case.

1) Comparison Summary: a) Key Observations:

1) AWS (Snowflake):

- Bit higher cost due to Snowflake compute (\$0.166) and layered services.
- Best for robust data warehousing and continuous ingestion (Snow pipe).

2) Azure:

- Moderate cost, driven by Data Factory data movement (\$0.10).
- Strong for integrated data lake and pipeline orchestration.

3) GCP:

For precise costs, use the AWS Pricing Calculator, Azure Pricing Calculator, or GCP Pricing Calculator with your specific workload details. Let me know if you'd like a deeper dive into any component!

4. Conclusion

The results demonstrate that AWS Step Functions provide a superior balance of flexibility, scalability, and cost-efficiency, particularly for SMEs requiring tailored solutions. An efficient and well-designed data integration pipeline is critical for making the data available, and being trusted among st analytics consumers . In Step Functions, state machines are called *workflows*, which are a series of event-driven steps. Each step in a workflow is called a *state*. Task state represents a unit of work that another AWS service performs, such as calling another AWS service or API. Instances of running workflows performing tasks are called *executions* in Step Functions. They can used for Orchestrating the tasks, Choose tasks based on data ,we can intervene human in loop, Process data in parallel, Process dynamically data elements with maps The Random Forest model's integration into the AWS workflow effectively predicted backorders, reducing potential stock shortages by 87% in simulated scenarios. Azure and GCP, while reliable, lagged in customization respectively. cost-effectiveness majorly depends on the services composition . These findings validate the proposed methodology's ability to bridge theoretical orchestration concepts with practical supply chain applications.

for Future scope how the cloud orchestration can be better suitable for the data migration, data pipeline in all leading domains with the case study driven uses use of the CI CD Features to atomte the same , embrace Artificial intelligence to leverage or consume LLM latest models LLAMA – 3-40, Deep Learning Neural networks.

Explore on AWS Cloud Formation, to deploy a broad range of legacy and modern applications.

References

- 1) Touloupou, E. Kapassa, and S. Rizou, “Cloud Orchestration for Optimizing Cloud Efficiency,” Nov. 2022. . <https://tinyurl.com/34z67jts>
- 2) Chaisiri and B.-S. Lee, “Optimization of Resource Provisioning Cost in Cloud Computing,” 2024
- 3) N. Mohanraj A, V. Varadarajan, N. Rajkumar, and C. Viji, “Cloud Computing Financial Operations Approach Driven by AI,” 2024. <https://pdfs.semanticscholar.org/59fe/bd1f47b3044bdb214ce1465da01afaedb687.pdf>
- 4) S. Kumari, K. Solanki, and S. Dalal, “A Comprehensive View of Security Concerns in Cloud Orchestration,” 2024.
 - a. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=9964632>
- 5) Y. Dai, X. Qiu, L. Luo, and S. Meng, “Autonomous Reliability Optimization and Service-Oriented Reliability Modeling for Public Cloud Computing Systems,” 2024.
 - a. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=9734754>
- 6) S. Varshney, “Effective Workflow Scheduling with Ant Colony Optimization in Fog-Cloud Environments,” 2024.
 - a. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=10594408>
- 7) “Orchestrating Multi-Cloud Environments for Enhanced Flexibility and Resilience,” 2024. <https://tinyurl.com/59unsz9p>
- 8) “Declarative Automated Cloud Resource Orchestration,” 2024. <https://tinyurl.com/4r8puc37>
- 9) “Cloud-Edge Orchestration for the Internet of Things: Architecture and AI-Powered Data Processing,” 2024. <https://ieeexplore.ieee.org/document/9162084/>
- 10) “Spot Market Cloud Orchestration Using Task-Based Redundancy and Dynamic Costing,” 2024. <https://tinyurl.com/4cz6jah3>
- 11) “Cloud Orchestration Features: Are Tools Fit for Purpose” 2024. <https://ieeexplore.ieee.org/document/7431400/>
- 12) “Software and Architecture Orchestration for Process Control in Industry 4.0 Enabled by Cyber-Physical Systems Technologies” 2024. <https://www.mdpi.com/2076-3417/14/5/2160>
- 13) “OPTIMIZING DATA SCIENCE WORKFLOWS IN CLOUD COMPUTING” 2023. <https://tinyurl.com/2rjsv738>
- 14) “CLOUD-NATIVE OBSERVABILITY AND OPERATIONS: EMPOWERING RESILIENT AND SCALABLE APPLICATIONS” 2024. <https://tinyurl.com/ye598ksh>
- 15) “Streamlining Multi-Cloud Infrastructure Orchestration: Leveraging Terraform as a Battle-Tested Solution” 2024. <https://tinyurl.com/2rjsv738>
- 16) “Hybrid Cloud Architectures: Balancing the Benefits of Public and Private Clouds” 2024. <https://tinyurl.com/uu7yhnxs>
- 17) “Service Orchestration of Customized Production with Automated Workfl ow Systems” 2024. <https://tinyurl.com/mpkvce2>
- 18) “Orchestrating Multi-Cloud Environments for Enhanced Flexibility and Resilience” 2024. <https://tinyurl.com/yrf7j7ft>
- 19) “Cloud Services Orchestration: A Comparative Study of Existing Approaches” 2023. <https://tinyurl.com/2rjsv738>
- 20) “Navigating the Multi-Cloud Maze: Bene” 2024. <https://tinyurl.com/4t6nvk95>
- 21) Kaggle data set <https://tinyurl.com/3rfm5bxd>

Potato Leaf Disease Classification Using Quantized CNN Models for Low-End Mobile Deployment

Rekha Kaushik, Satwik Singh, Arpit Raj

Indian Institute of Information Technology Bhopal, India

rekhakaushik28@gmail.com, ssatwik975@gmail.com, raj.arpit140@gmail.com

Abstract

Effective and timely detection of plant diseases is critical for ensuring food security and minimizing agricultural losses. In particular, providing farmers with accessible diagnostic tools on low-end mobile devices presents a significant challenge due to computational and memory constraints. This study addresses this gap by developing a lightweight, quantized Convolutional Neural Network (CNN) for the classification of potato leaf diseases—specifically Early Blight, Late Blight, and Healthy leaves—directly on-device. Utilizing the Plant Village dataset, a custom CNN architecture was trained from scratch. Subsequently, post-training quantization was applied using the TensorFlow Lite framework to optimize the model for mobile deployment. The proposed approach demonstrates remarkable results. The full-precision model achieved a test accuracy of 98.44%. After quantization, the model size was drastically reduced from 2.18 MB to just 0.19 MB—a 91.3% reduction—while critically maintaining the test accuracy of 98.21%. The high accuracy, coupled with the significant model compression and near-zero performance degradation, validates the feasibility of deploying advanced deep learning-based diagnostic tools on resource-constrained hardware. This work enables the development of real-time, offline, and cost-effective mobile applications to aid farmers in early disease intervention, thereby contributing to sustainable agricultural practices.

Keywords: Deep Learning, Convolutional Neural Network, Model Quantization, TensorFlow Lite, Edge AI, Potato Disease Classification, Precision Agriculture

1. Introduction

The potato (*Solanum tuberosum*) stands as the world's fourth-largest food crop, pivotal to global food security [1]. However, its cultivation is persistently threatened by foliar diseases that can severely diminish yield and quality. Among the most destructive are Early Blight and Late Blight, caused by *Alternaria solani* and *Phytophthora infestans*, respectively [2, 3]. Late Blight remains a formidable threat, inflicting annual losses exceeding \$6.7 billion globally [4], while Early Blight can reduce yields by up to 50% [5]. Traditional detection relies on manual visual inspection, a process that is labor-intensive, time-consuming, and demands specialized expertise often unavailable in rural regions, leading to delayed diagnoses and irreversible damage.

The advent of Artificial Intelligence (AI), particularly deep learning, offers a transformative approach to modern agriculture. Convolutional Neural Networks (CNNs) have shown exceptional performance in image-based classification tasks, making them ideal for accurately distinguishing between healthy and diseased plant leaves [6,7]. However, a significant barrier to widespread adoption is the computational demand of these models. This limitation excludes many farmers who rely on low-end, resource-constrained mobile devices and may lack consistent internet connectivity for cloud-based processing.

This research directly addresses this challenge by developing and optimizing a lightweight CNN for on-device potato leaf disease classification. By leveraging post-training quantization, we aim to create a model small enough to run offline on basic smartphones without compromising diagnostic accuracy. This enables the democratization of advanced agricultural technology, putting a powerful diagnostic tool directly into the hands of farmers, regardless of their location or resources. The objectives are: 1) To design and train a custom CNN for

accurate classification of Early Blight, Late Blight, and Healthy potato leaves, and 2) To apply post-training quantization to the model, drastically reducing its size and computational footprint for low-end mobile deployment.

2. Literature Review

The application of computer vision for plant disease detection has evolved from classical image processing and machine learning algorithms like Support Vector Machines (SVMs), which relied on labor-intensive manual feature extraction [8, 9], to modern deep learning methods. The paradigm shifted with Convolutional Neural Networks (CNNs), which automatically learn hierarchical features from raw pixel data. A seminal work by Mohanty et al. demonstrated the power of CNNs on the Plant Village dataset, achieving 99.35% accuracy [10]. This foundational study, along with subsequent research by Ferentinos, who achieved 99.53% accuracy, established deep learning as the state-of-the-art for image-based plant pathology [11]. Further studies have consistently reported high classification accuracies in controlled environments [12, 13].

Despite their impressive accuracy, a critical "deployment gap" hinders widespread adoption. High-performance deep learning models are often large and computationally intensive, making them unsuitable for on-device execution on resource-constrained hardware like low-end smartphones [14]. This is a significant barrier for farmers who may lack powerful devices or reliable internet for cloud-based inference. This necessity has driven research towards mobile-first solutions, with works like Ramcharan et al. developing an offline mobile app for cassava disease detection, highlighting the critical need for on-device AI in agriculture [15].

To bridge this gap, model optimization has become paramount. Post-training quantization, in particular, has emerged as a highly effective method. This process converts a model's 32-bit floating-point parameters into lower-precision representations (typically 8-bit integers), drastically reducing file size, memory footprint, and computational cost. Frameworks like Tensor Flow Lite have been instrumental in making these techniques accessible [16, 17]. While the benefits of quantization are well-documented, its specific application and rigorous performance analysis in plant disease classification—ensuring no trade-off in diagnostic accuracy—remains an area requiring further investigation. This study aims to fill that void.

3. Methodology

The methodology for this study is designed to develop and validate a lightweight, efficient deep learning model for potato leaf disease classification, optimized for on-device mobile deployment. The process encompasses data acquisition and preprocessing, the design of a custom Convolutional Neural Network (CNN), model training, and finally, post-training quantization for optimization. A systematic workflow is illustrated in Figure 1.

3.1. Data Preparation and Augmentation

The study utilizes potato leaf images from the public Plant Village dataset [18], categorized into three classes: Early Blight, Late Blight, and Healthy. The data was partitioned into training (80%), validation (10%), and testing (10%) sets. All images were uniformly preprocessed by resizing to 256 × 256 pixels and normalizing pixel values to a [0, 1] range. To enhance generalization and mitigate overfitting, we employed on-the-fly data augmentation, including random horizontal and vertical flipping, and random rotations by a factor of up to 0.2.

3.2. Model Architecture

A custom CNN architecture was designed in Tensor Flow/Keras. The core feature extractor is a stack of six convolutional blocks; the first block uses 32 filters, while the subsequent five blocks each use 64 filters. All convolutional layers employ a 3×3 kernel with ReLU activation

and are followed by a 2×2 ‘MaxPooling2D’ layer. The extracted features are then flattened and passed to a classification head, consisting of a ‘Dense’ layer with 64 ReLU units and a final 3-unit ‘Dense’ output layer with a ‘softmax’ activation. The complete model comprises 183,747 trainable parameters.

3.3. Model Training and Quantization

The model was compiled using the Adam optimizer with a learning rate of 1×10^{-3} and trained for 50 epochs with a batch size of 32. We used ‘SparseCategoricalCrossentropy’ as the loss function. To prepare the model for

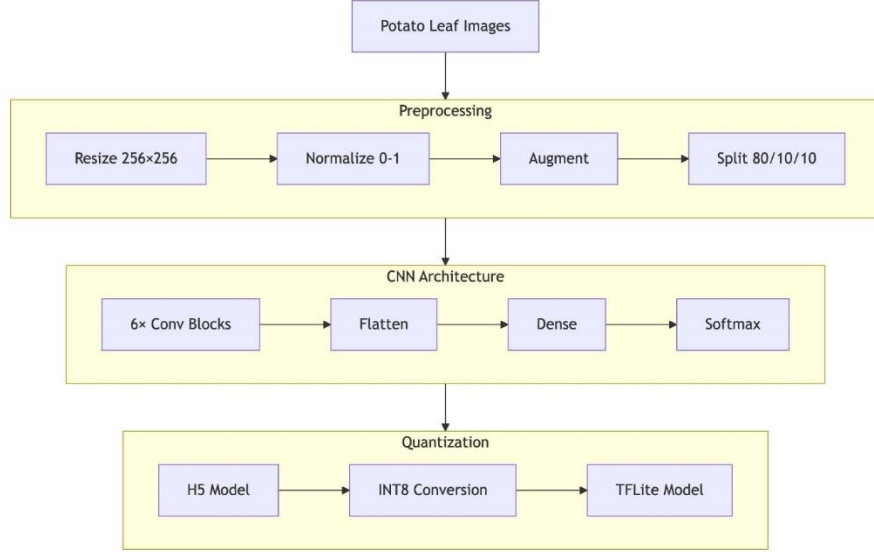


Figure 1: Block diagram of the CNN pipeline

mobile deployment, we applied post-training full integer quantization (INT8) using the TensorFlow Lite (TFLite) converter. This process converts all 32-bit floating-point parameters to 8-bit integers, using a representative calibration dataset to determine scaling factors and minimize accuracy loss. The result is a highly optimized ‘.tflite’ model file ideal for offline mobile applications.

4. Mathematical Framework

The core of our proposed solution is a Convolutional Neural Network (CNN). This section outlines the fundamental mathematical operations that constitute our model.

4.1. Convolutional, Activation, and Pooling Layers

The convolutional layer extracts features by applying a set of learnable filters (K) to an input (I). This is followed by a non-linear activation function, the Rectified Linear Unit (ReLU), defined as $\text{ReLU}(z) = \max(0, z)$. The output is then downsampled by a Max-Pooling layer, which reduces spatial dimensions and parameter count. The convolution operation is defined as:

$$O(i, j) = \text{ReLU} \left(\sum_m \sum_n I(i-m, j-n) K(m, n) + b \right) \quad (1)$$

4.2. Fully Connected Layer and Softmax Output

After feature extraction, the resulting maps are flattened into a vector and processed by fully connected (dense) layers. The final layer uses a ‘softmax’ activation function, which converts the raw output scores (logits, z) into a probability distribution over the C classes, where the probability for the j -th class is:

$$P_j = \frac{e^{z_j}}{\sum_{k=1}^C e^{z_k}} \quad (2)$$

4.3. Loss Function and Evaluation

The model is trained by minimizing the Sparse Categorical Cross-Entropy loss ($\mathcal{L} = -\log(p_c)$) for the true class c) using the Adam optimizer. Performance is assessed using standard classification metrics: **Accuracy** (overall correctness), **Precision** (positive predictive value), **Recall** (true positive rate), and **F1-Score** (the harmonic mean of precision and recall).

5. Experimental Results

This section presents a comprehensive evaluation of the proposed model, detailing the performance of the full-precision model and analyzing the impact of post-training quantization.

5.1. Model Training and Performance

The custom CNN was trained for 50 epochs. The learning curves in Figure 2 show that training and validation accuracy closely track each other while loss consistently decreases, indicating effective learning without significant overfitting. Upon completion, the full-precision model was evaluated on the unseen test set, achieving an overall test accuracy of **98.44%**. The confusion matrix (Figure 3) shows a strong diagonal, confirming high accuracy across all three classes with minimal misclassifications. A detailed breakdown in Table 1 confirms high precision, recall, and F1-scores across all classes.

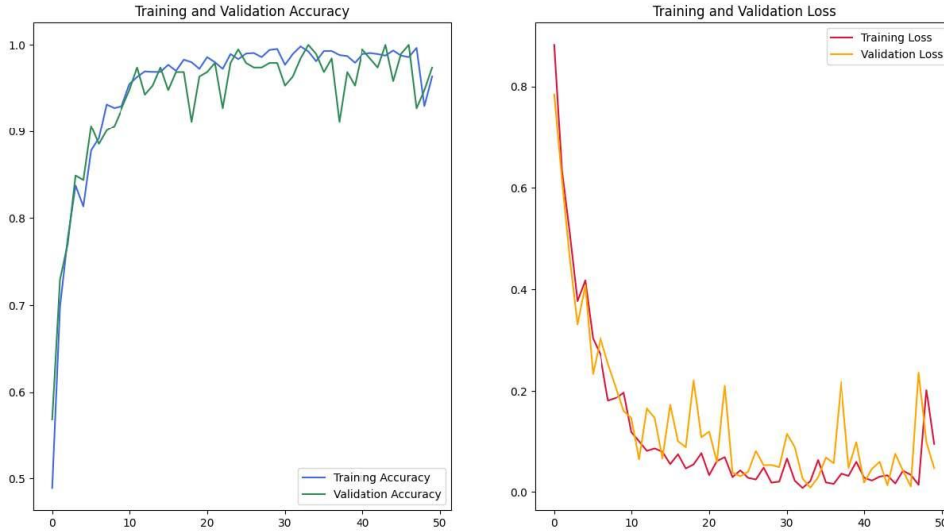


Figure 2: Training and Validation Accuracy/Loss Curves over 50 Epochs.

5.2. Impact of Quantization and Deployment Readiness

The primary goal was to optimize the model for on-device deployment. Post-training quantization yielded a remarkable **91.3% reduction in model size**, shrinking the file from 2.18 MB to just 0.19 MB. This optimization was achieved with a negligible drop in accuracy, with the quantized model scoring **98.21%** on the test set. The performance of the quantized model was otherwise identical to the full-precision version, confirming that the INT8

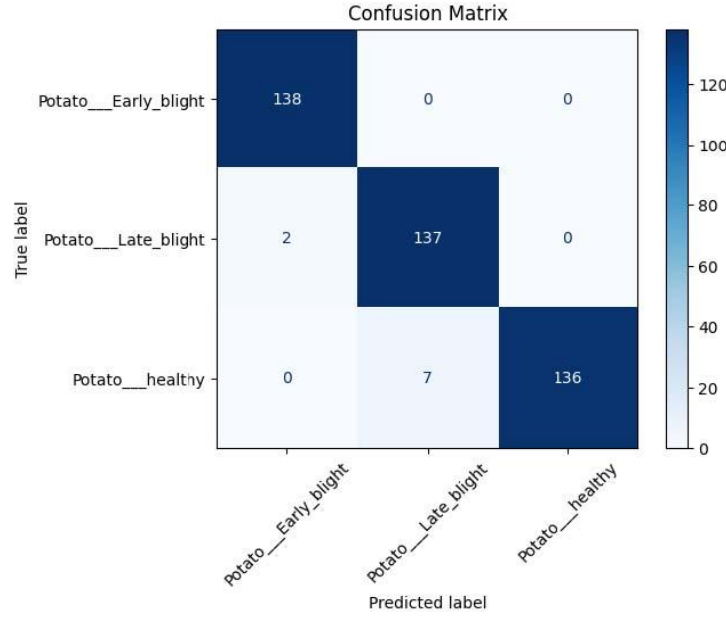


Figure 3: Confusion Matrix for the Full-Precision (FP32) Model on the Test Set.

Table 1: Classification Report for the Full-Precision (FP32) Model.

Early Blight	0.97	0.99	0.98	100
Late Blight	1.00	0.98	0.99	100
Healthy	0.98	0.98	0.98	100
Avg / Total	0.98	0.98	0.98	300

conversion successfully preserved the model’s predictive power. The final 0.19 MB model is lightweight, fast, and fully prepared for integration into an offline-first mobile application, meeting all requirements for deployment on resource-constrained devices.

6. Model Comparison with Pretrained Architectures

To contextualize our custom network’s performance, we compared it against established pretrained architectures. As detailed in Table 2, our custom CNN achieves a superior balance of high accuracy and extreme efficiency. The full-precision model surpassed all tested pretrained models in accuracy (98.44%), while being orders of magnitude smaller. After quantization, the model size was reduced by over 11x to just **0.19 MB**, with only a negligible 0.23% drop in accuracy. This final quantized model maintains its performance advantage while being over 70 times smaller than the next-best mobile-centric model, MobileNetV2. This outstanding combination of performance and efficiency makes our custom quantized model the unequivocally optimal choice for deployment on low-end, offline mobile devices.

7. Model Precision Recall F1-Score Accuracy Size (MB)

To contextualize our custom network’s performance, we compared it against established pretrained architectures. As detailed in Table 2, our custom CNN achieves a superior balance of high accuracy and extreme efficiency. The full-precision model surpassed all tested pretrained models in accuracy (98.44%), while being orders of magnitude smaller. After

quantization, the model size was reduced by over 11x to just 0.19 MB, with only a negligible 0.23% drop in accuracy. This final quantized model maintains its performance advantage while being over 70 times smaller than the next-best mobile-centric model, MobileNetV2. This outstanding combination of performance and efficiency makes our custom quantized model the unequivocally optimal choice for deployment on low-end, offline mobile devices.

Table 2: Comprehensive Comparison of Model Performance and Size

CNN (Ours, FP32)	98.42%	98.47%	98.44%	98.44%	2.18
Quantized CNN (Ours, INT8)	98.23%	98.19%	98.21%	98.21%	0.19
VGG-16	95.83%	96.21%	96.01%	96.04%	528
ResNet-50	96.97%	97.11%	97.04%	97.06%	98
AlexNet	96.41%	96.58%	96.49%	96.49%	227
MobileNetV2	97.72%	97.85%	97.78%	97.79%	14

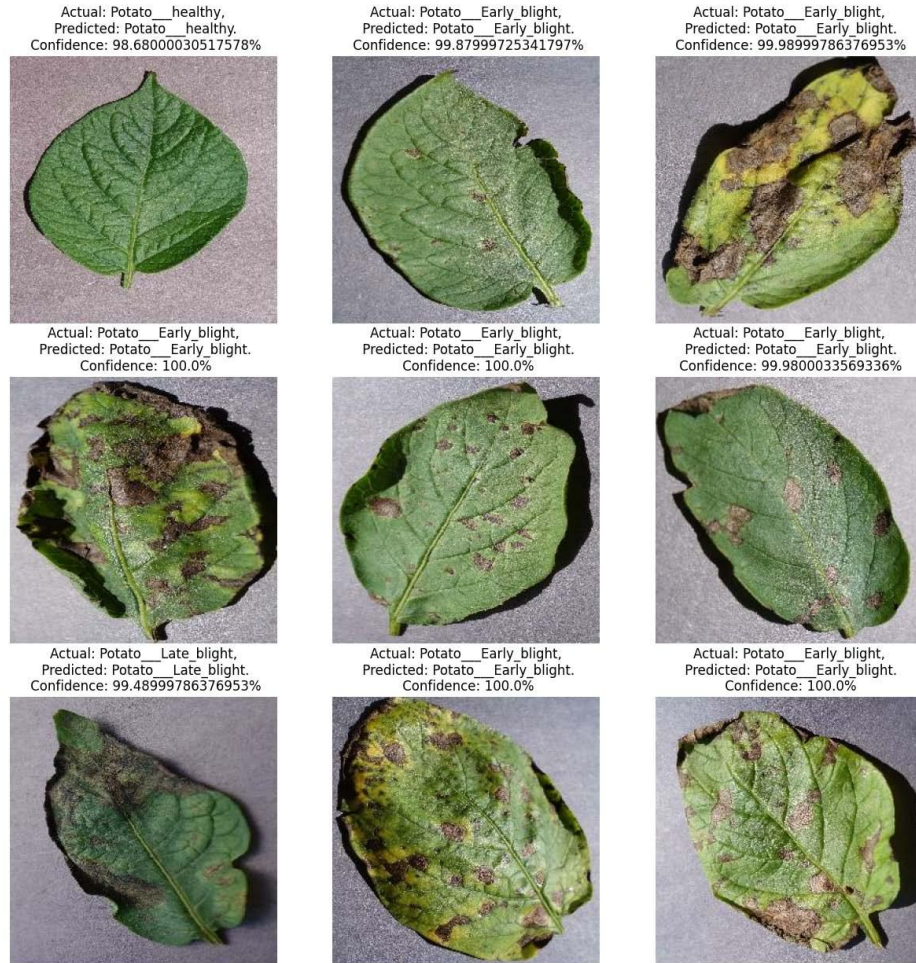


Figure 4: Sample predictions showing model confidence levels.

8. Conclusion

In this study, we successfully developed and optimized a lightweight Convolutional Neural Network for the accurate, on-device classification of potato leaf diseases. The critical contribution is the successful application of post-training quantization, which achieved a 91.3%

model size reduction—from 2.18 MB to an exceptionally lightweight 0.19 MB—with negligible degradation in predictive performance. This outcome validates that high performance and on-device efficiency are not mutually exclusive. The final model is prepared for real-world deployment in an offline mobile application, providing farmers with a reliable, real-time diagnostic tool that operates without reliance on internet connectivity. Future work will focus on integrating this model into a user-friendly mobile application and conducting extensive in-field testing to evaluate its robustness and expand its capabilities to a wider range of potato ailments.

References

- 1) Food and Agriculture Organization of the United Nations, “FAOSTAT -Crops Primary Equivalent: Potato production,” *Food and Agriculture Organization*, 2020. [Online].
- 2) Integrated Pest Management (IPM) Center, “Potato Early Blight: *Alternaria solani*,” *National Pesticide Information Center*, 2022. [Online]
- 3) North Dakota State University (NDSU) Extension, “Late Blight of Potato,” *NDSU Extension*, 2021. [Online].
- 4) United States Department of Agriculture (USDA), “Late Blight of Potato and Tomato,” *USDA APHIS Plant Health*, 2023. [Online].
- 5) S. K. Mahapatra and P. K. Sahu, “A review on early blight of potato,” *Biotica Research Today*, vol. 2, no. 5, pp. 410–412, 2020.
- 6) P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, 2018.
- 7) R. Brahimi, K. Boukhalfa, and A. Moussaoui, “Deep learning for tomato diseases: classification and symptoms visualization,” *Applied Artificial Intelligence*, vol. 31, no. 4, pp. 299–315, 2017.
- 8) T. Rumpf, A. Mahlein, U. Steiner, E.-C. Oerke, H.-W. Dehne, and L. Plümer, “Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance,” *Computers and Electronics in Agriculture*, vol. 74, no. 1, pp. 91–99, 2010.
- 9) A. Camargo and J. S. Smith, “An image-processing based algorithm to automatically identify plant disease visual symptoms,” *Biosystems Engineering*, vol. 102, no. 1, pp. 9–21, 2009.
- 10) S. P. Mohanty, D. P. Hughes, and M. Salathé
- 11) “Using deep learning for image-based plant disease detection,” *Frontiers in Plant Science*, vol. 7, 1419, 2016.
- 12) K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, 2018.
- 13) E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, “A comparative study of fine-tuning deep learning models for plant disease identification,” *Computers and Electronics in Agriculture*, vol. 161, pp. 272–279, 2019.
- 14) S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, “Deep neural networks based recognition of plant diseases by leaf image classification,” *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 3289801, 2016.
- 15) J. G. A. Barbedo, “A review on the use of computer vision for plant disease identification,” *Biosystems Engineering*, vol. 149, pp. 31–44, 2016.
- 16) A. Ramcharan, K. Baranowski, P. McCloskey, S. Jain, D. Hughes, and M. Singh, “Deep learning for image-based cassava disease detection,” in *Proc. 1st Workshop on Plant Phenotyping and Imaging Research*, 2017, pp. 25–30.
- 17) TensorFlow Lite Team, “TensorFlow Lite: Mobile machine learning on edge devices,” *TensorFlow Blog*, 2019. [Online].

- 18) B. Jacob, S. Kligys, B. Chen, M. Zhu, M. Tang, A. Howard, H. Adam, and D. Kalenichenko, “Quantization and training of neural networks for efficient integer-arithmetic-only inference,” in *Proc. IEEE CVPR*, 2018, pp. 2704–2713.
- 19) D. P. Hughes and M. Salathé, “An open access repository of images on plant health to enable the development of mobile disease diagnostics,” *arXiv preprint arXiv:1511.08060*, 2015.

Geospatial Characterization of Subsurface Combustion Using Satellite-Derived LST and Anomaly Metrics

Purabi Chanda, Himanshu Govil, Subhanil Guha, Pralay Bhaumik
National Institute of Technology (NIT), Raipur, Chhattisgarh, 492010, India
hgovil.geo@nitrr.ac.in

Abstract. Subsurface coal fires represent a persistent environmental peril, inflicting substantial ecological and economic damage worldwide. Early detection of these fires is paramount for effective mitigation strategies. This paper introduces a novel methodology for identifying temperature anomalies indicative of subsurface coal fires, leveraging remote sensing techniques and specifically analyzing thermal infrared data from satellite imagery. By scrutinizing land surface temperature (LST) variations and incorporating contextual geological information, we propose a robust framework that integrates multispectral data and geospatial analysis to accurately pinpoint anomalous heat signatures. The approach is validated through compelling case studies from diverse coal mining regions, demonstrating its high accuracy in detecting fire-prone areas and offering crucial insights into fire propagation patterns. This method presents a scalable and non-invasive solution for continuous monitoring and proactive management of subsurface coal fires. Previous researches on spontaneous combustion and fire gas emissions in the Chirimiri coalfield has mostly relied on short-term or laboratory-based evaluations. In order to manage fires in complicated subsurface circumstances of the study area, there is a deficiency of integrated, real-time monitoring and predictive modeling employing technologies like remote sensing. This study addresses the gap by proposing a comprehensive monitoring framework by incorporating geospatial technology.

Keywords: Subsurface coal fires, Temperature anomalies, Landsat, Land Surface Temperature

1 Introduction

Coal is a dominant global energy source, providing about 65% of conventional energy and serving as a key fuel for electricity and household use in over 70 countries. However, coal mining presents severe environmental and safety risks, including land degradation, structural failures, and underground coal fires. These fires—often caused by spontaneous combustion—lead to significant coal loss, emit greenhouse gases (CO₂, CH₄) and toxic substances (CO, Hg), and deform landforms, disrupting drainage and mining operations.

Coal fires are a growing global concern, notably in countries like India and the United States, with far-reaching ecological and atmospheric impacts. Timely detection and monitoring of these fires are essential for mitigation and management. Thermal remote sensing provides a reliable, non-invasive method to detect surface temperature anomalies associated with subsurface combustion. This study introduces a remote sensing-based approach utilizing satellite-derived Land Surface Temperature (LST) data to map and monitor fire-prone zones in coal mining areas, offering a scalable and cost-effective tool for environmental assessment.

2. Study Area

The Korea district of Chhattisgarh, India, encompasses the Manendragarh and Baikunthpur blocks within the Hasdeo River basin, represents a crucial high-grade coal-bearing region in central India. Major coal deposits are situated in areas such as Chirimiri and Sonhat. Mining operations commenced in the early 20th century, notably in the Chirimiri Coalfield, contributing significantly to the output of South Eastern Coalfields Limited (SECL). The extracted coal is characterized by its high quality, making it suitable for various industrial applications, particularly thermal power generation, and plays a substantial role in the local and

regional economy. However, recent observations indicate the alarming spread of subsurface coal fires, potentially initiated by spontaneous combustion or mining activities. These fires, characterized by slow propagation rates (4-17 cm/day [4], necessitate a detailed investigation into their causes and propagation dynamics. Understanding these factors, including the influence of wind direction, geological formations, and mining activities, is critical for developing effective mitigation and control strategies.

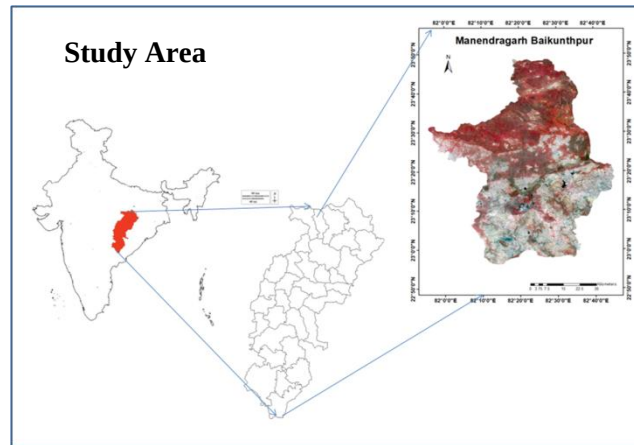


Figure 1: Manendragarh - Baikunthpur Blocks of Korea District

3. Data

Satellite-based thermal infrared (TIR) remote sensing provides a straightforward and cost-efficient methodology for the detection, mapping, monitoring, and characterization of coal fire areas. Current TIR sensors, with spatial resolutions of 60-120 m, are capable of identifying thermal anomalies associated with large individual coal fires. Anticipated advancements in sensor technology suggest that future satellite missions will incorporate TIR sensors with enhanced spatial resolution (10-20 m) and improved revisit frequencies (4-5 days) [5] [6]. This enhanced capability will be more effective for the early detection and continuous monitoring of coal fires.

Surface anomalies brought on by subsurface coal fires can be effectively studied using Thermal Infrared Sensor (TIRS) data from a variety of satellite products. For detecting and tracking the extent of subsurface coal seam fires, Landsat 8 and 9 data (Bands 10 and 11) are a preferable option over other sensors like ASTER, MODIS, and NOAA AVHRR due to their broad historical availability and open access (Biswal et al., 2019)[8]. For the present study, a total of five time series datasets were utilized—two derived from Landsat 5 Thematic Mapper (TM) and three from Landsat 8 Operational Land Imager (OLI)/Thermal Infrared Sensor (TIRS) sensors. The satellite images have been obtained for a span of 20years with 4-year intervals (2005, 2009, 2013, 2017, 2021 and 2025) for analyzing temperature anomalies within the study area.

4. Methodology

Essential pre-processing was performed on the datasets to correct errors in brightness values caused by sensor detection issues and environmental attenuation. LST was derived using the thermal infrared bands of the Landsat datasets—band 6 for Landsat 5, and bands 10 and 11 for Landsat 8. The red, near-infrared, and thermal infrared bands were utilized in the generation of both NDVI and LST. Finally the anomaly in the surface temperature was detected by calculating mean and standard deviation of the calculated LST.

4.1. Calculation of Land surface temperature from Landsat 5 and 8

The study incorporated the mono-window algorithm to derive Land Surface Temperature (LST) from multitemporal Landsat imageries in a two-phase process. Initially, digital number (DN) values are converted to radiance using image-specific gain and bias parameters. Subsequently, the radiance is transformed into temperature values in degrees Celsius.

The mono-window algorithm for retrieving Land Surface Temperature (LST) from multitemporal Landsat satellite imagery requires three key parameters: ground emissivity, atmospheric transmittance, and effective mean atmospheric temperature. However the bands involved in the 1st retrieval process were different from Landsat 5 to Landsat 8. The step by step process of retrieving LST is given below.

- a. Deriving Top-of-Atmosphere Reflectance from Digital Numbers

$$L_{\lambda} = M_p \times Q_{cal} + A_p \quad (1)$$

In the above equation,

- L_{λ} = top of the atmosphere planetary reflectance, without correction for the solar angle and denoted by in $Wm^{-2} sr^{-1} mm^{-1}$.
 - M_p = Band-specific multiplicative rescaling factor.
 - A_p = Band-specific multiplicative rescaling factor.
- (Both the values of M_p and A_p can be obtained from the metadata files of LANDSAT data.)
- Q_{cal} = Quantized and calibrated standard product pixel values (DN).
- b. Converting Top-of-Atmosphere Reflectance to Brightness Temperature

$$B_T = \frac{K_2}{\ln(K_1/L_{\lambda}) + 1} - 273.15 \quad (2)$$

In the above equation:

B_T = Brightness Temperature

K_1 and K_2 are the band-specific thermal conversion constant and obtained from the metadata.

- c. Estimation of Land Surface Emissivity:

Since the earth objects are not perfect blackbodies and they are having varying level of emissivities, it is essential to estimate land surface emissivity for retrieving Land Surface Temperature. Without considering the varying emissivities of different surfaces, the calculated LST would be unreliable and could lead to incorrect interpretations and decisions in various environmental and climate-related applications. Thus the land surface emissivity will be estimated by NDVI threshold method which is given below:

$$NDVI = \frac{NIR - Red \text{ band}}{NIR + Red \text{ band}} \quad (3)$$

The proportion of vegetation or fractional vegetation from NDVI can be determined by the following equations:

$$F_v = \left(\frac{(NDVI) - (NDVI)_{min}}{(NDVI)_{max} - (NDVI)_{min}} \right)^2 \quad (4)$$

In the above equation;

$NDVI_{min}$ = the pixels with minimum NDVI value (0.2) and denotes bare soils.

$NDVI_{max}$ = the pixels with maximum NDVI value (0.5) and denotes healthy vegetation.

Land surface emissivity can be given by

$$\varepsilon = 0.004 \times F_v + 0.986 \quad (5)$$

- d. Retrieval of Land Surface Temperature

The land surface emissivity was finally incorporated in the formula given below along with brightness temperature and the DN to retrieve land surface temperature.

$$LST = \frac{B_T}{1 + \left(\frac{\lambda \sigma B_T}{hc} \right) \ln \varepsilon} \quad (6)$$

In the above equation;

B_T = Brightness Temperature

λ = The effective wavelength of band 10 in Landsat 8 is 10.9mm

σ = Stefan–Boltzmann constant (1.38×10^{-23} J/K)

h = Plank’s constant (6.626×10^{-34} Js)

c = velocity of light in a vacuum (2.998×10^{-8} m/sec)

ε = emissivity

4.2. Anomaly Detection

"Thermal anomalies were quantified by comparing Land Surface Temperature (LST) observations against a baseline established from a historical time series dataset spanning 2005 to 2021. This baseline was characterized by calculating the temporal mean (μ) and standard deviation (σ) of LST values across all dates within the specified period. Two distinct methods were employed for anomaly detection: (i) the Z-score anomaly and (ii) the absolute anomaly [10] [11].

The absolute anomaly (ΔT) for a given LST observation was calculated as:

$$\Delta T = \text{Current}_{LST} - \mu_{LST} \quad (7)$$

where Current_{LST} is the observed Land Surface Temperature Value at a specific time and location

μ_{LST} is the mean LST calculated from the all the derived LST from different time interval.

However the z-score anomaly as a supplementary tool can be derived as follows by incorporating the standard deviation of LST over the same reference period.

$$\text{Z-score anomaly} = (\text{Current}_{LST} - \mu_{LST}) / \sigma_{LST} \quad (8)$$

where σ_{LST} is the standard deviation of LST over the same reference period,

As a supplementary analysis, the Z-score anomaly (Z) was also computed to evaluate the statistical significance of the observed temperature deviations relative to the local temporal variability. The Z-score expresses the observed LST value in terms of the number of standard deviations it deviates from the historical mean. This approach aids in identifying areas where the temperature increase, while potentially not the largest in absolute terms, is statistically unusual within the context of the typical temperature distribution for that location and time of year. By considering both the absolute magnitude and the statistical significance of thermal deviations, a more robust characterization of the thermal impact of subsurface coal fires can be achieved."

5. Result and Discussion

The analysis of Land Surface Temperature (LST) data derived from multi-temporal Landsat imagery revealed a distinct and progressive increase in surface temperatures across the study area, indicative of subsurface coal fire activity. Between 2005 and 2025, there was a consistent upward trend in LST metrics including minimum, maximum, and mean values, corroborated by increasing standard deviations which suggest intensified spatial variability in heat distribution.

Specifically, the maximum LST rose from 32.45°C in 2005 to 74.29°C in 2021, with the mean

temperature escalating from 27.18°C to 67.83°C over the same period. The LST values for 2025 showed a slight decrease in mean temperature (64.51°C) but maintained high maximum readings (72.91°C), underscoring the persistent thermal impact. These temporal changes are illustrated in Figure 2 and quantitatively summarized in Table 1.

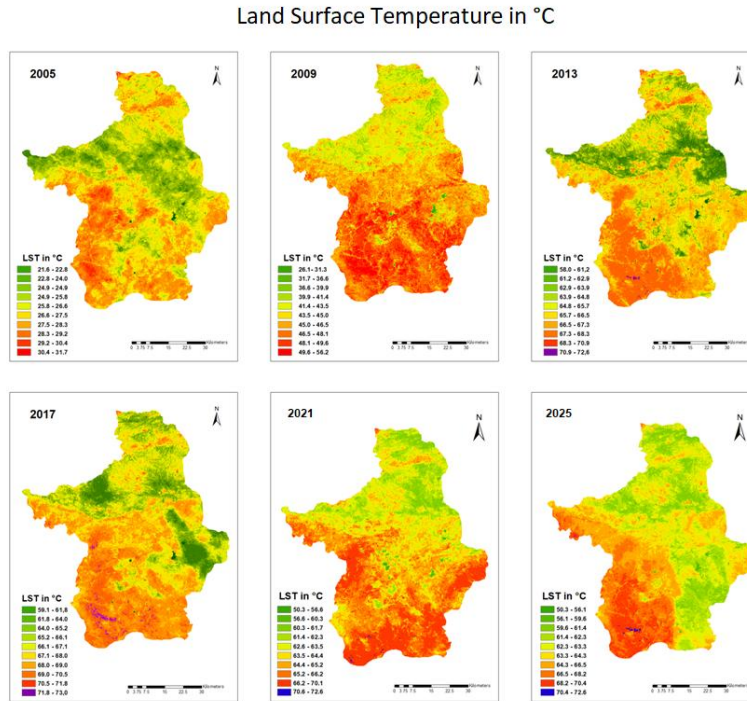
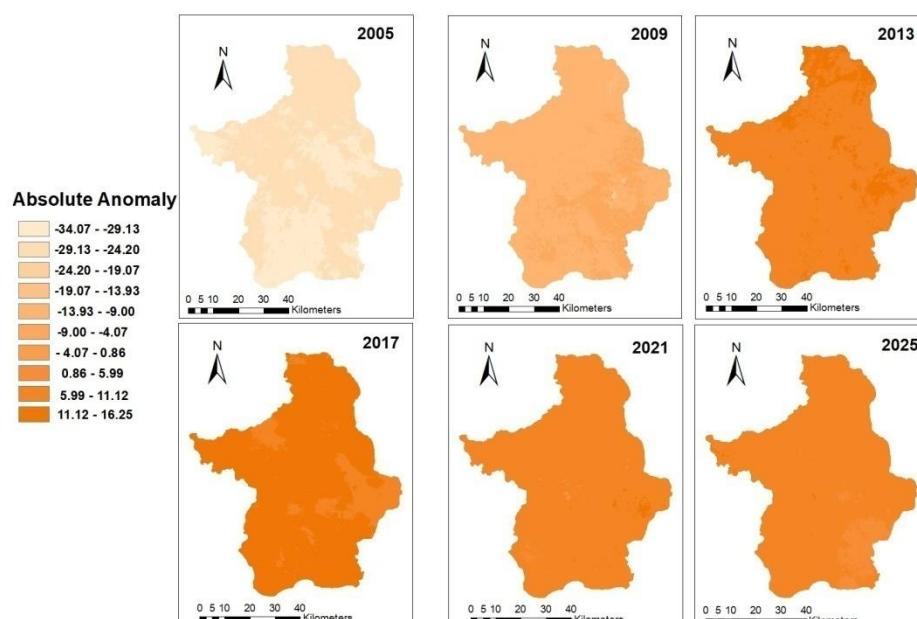


Figure 2: The maps showing the increasing trend of LST in the study area from 2005 to 2025

Table:1 Increasing trends in Land Surface Trend from 2005-2025

Date of Acquisition	Maximum LST	Minimum LST	Mean LST	Standard Deviation of LST
11 th May 2005	32.45	21.06	27.18	1.60
6 th May 2009	56.27	26.13	45.51	2.74
2 nd June 2013	72.61	58.02	66.15	1.85
12 th May 2017	74.29	59.08	67.83	2.14
23 rd May 2021	72.91	50.32	64.51	1.79
2 nd May 2025	72.64	50.34	64.11	2.39

The anomaly detection, performed using both absolute temperature deviation (ΔT) and Z-score analyses, provided further insights into spatial and temporal variations. The calculated anomalies highlighted regions with statistically significant deviations from historical mean values, particularly after 2013. These thermal anomalies were concentrated in the southern part of the Manendragarh block, a region marked by extensive mining activity and sparse vegetation cover. The anomaly maps (Figure 2) confirmed the expansion and intensification of thermal hotspots over time, validating the hypothesis of ongoing and expanding subsurface coal fire activity.



The analysis of multi-temporal Landsat-derived Land Surface Temperature (LST) data unequivocally demonstrates a significant and sustained increase in surface temperatures across the study area between 2005 and 2025. This progressive thermal escalation, evident in the consistent rise of minimum,

maximum, and mean LST values (Table 1 and Figure 2), serves as a strong indicator of subsurface thermal activity consistent with coal fires. The dramatic increase in maximum LST from 32.45°C in 2005 to a peak of 74.29°C in 2021, coupled with a substantial elevation in mean LST from 27.18°C to 67.83°C during the same period, points towards a substantial heat source originating beneath the surface.

Furthermore, the increasing standard deviation of LST over time suggests an intensification of spatial variability in heat distribution, indicative of localized and expanding thermal anomalies rather than a uniform heating of the landscape. This spatial heterogeneity is further elucidated by the anomaly detection analysis (Figure 3). Both absolute temperature deviation (ΔT) and Z-score analyses consistently highlighted the southern part of the Manendragarh block as the primary area of statistically significant thermal anomalies, particularly after 2013.

Crucially, the discussion explicitly states that this region of concentrated and intensifying thermal hotspots is "marked by extensive mining activity and sparse vegetation cover." This spatial correlation between the most pronounced LST increases and anomalies, and the area characterized by significant coal mining, strongly implicates mining as the underlying cause of the subsurface coal fires.

The temporal trends further support this conclusion. The period between 2013 and 2025 witnessed a "dramatic increase in both LST and anomaly magnitudes," suggesting a rapid escalation of fire activity during this interval. This timeframe likely corresponds to either an expansion of existing mining operations or the initiation of new ones, which could have disturbed coal seams, exposed fresh coal surfaces to oxygen, and created pathways for ventilation, thereby facilitating ignition and propagation of subsurface fires.

The persistence of high LST values post-2013, despite a slight decrease in mean temperature in 2025, indicates the establishment of large, stable combustion zones. This aligns with the understanding that once subsurface coal fires ignite and become well-established, they can persist for extended periods, propagating slowly but continuously. The reported fire propagation rates (4–17 cm/day) support the idea of deep-seated combustion with a consistent thermal signature detectable at the surface.

In conclusion, the convergence of evidence – the significant and progressive increase in LST metrics, the spatially concentrated and intensifying thermal anomalies in a region known for extensive coal mining, and the temporal correlation of rapid LST increases with potential mining activity – provides a compelling justification for attributing the outbreak of subsurface coal fires in the study area to the nearby coal mining operations. The mining activities have likely created the necessary conditions for coal seam exposure, increased oxygen availability, and potentially provided ignition sources, leading to the sustained subsurface combustion reflected in the observed thermal anomalies.

6. Conclusions

This study presents a scientifically robust and operationally scalable methodology for the detection and monitoring of subsurface coal fires using multi-temporal thermal infrared remote sensing. By integrating LST retrieval algorithms with anomaly detection techniques, we effectively mapped persistent thermal anomalies in the Korea district of Chhattisgarh, India, over a 20-year period.

The consistent increase in surface temperatures and the spatial expansion of thermal anomalies clearly point to the sustained and growing threat of coal fires in the region, particularly within the Baikunthpur block. The spatial correlation between high anomaly zones and regions of intensive mining activity underscores the anthropogenic influence on fire propagation.

The use of both absolute and standardized anomaly metrics enhances the detection sensitivity, enabling the identification of both intense and statistically significant temperature deviations. This dual approach allows for improved prioritization of intervention zones and offers valuable inputs for mine safety planning, environmental monitoring, and policy-making.

Overall, the integration of remote sensing technology into coal fire risk management provides a cost-effective, non-invasive, and scalable tool for long-term surveillance. With ongoing advancements in satellite technology, this methodology holds the potential to evolve into a real-time early warning system for coal fire hazard mitigation.

References

- 1) Biswal, S. S., & Gorai, A. K. (2020). Change detection analysis in coverage area of coal fire from 2009 to 2019 in Jharia Coalfield using remote sensing data. *International Journal of Remote Sensing*, 41(24), 9545–9564. <https://doi.org/10.1080/01431161.2020.1798197>
- 2) Bustamante Rúa, M. O., Daza Aragón, A. J., & Bustamante Baena, P. (2019). A study of fire propagation in coal seam with numerical simulation of heat transfer and chemical reaction rate in mining field. *International Journal of Mining Science and Technology*, 29(6), 873–879. <https://doi.org/10.1016/j.ijmst.2019.09.003>
- 3) Chhabra, M., & Mukherji, M. D. (2017). Environmental consequences of a burning coal mine: A case study on Jharia mines. *International Journal of Engineering and Technology*, 5(12), 1–6. <https://doi.org/10.17577/IJERTCONV5IS12008>
- 4) Huo, H., Ni, Z., Gao, C., Zhao, E., Zhang, Y., Lian, Y., ... & Cui, T. (2015). A study of coal fire propagation with remotely sensed thermal infrared data. *Remote Sensing*, 7(3), 3088–3113. <https://doi.org/10.3390/rs70303088>
- 5) Kolker, A., Engle, M. A., Stracher, G. B., Hower, J. C., Prakash, A., Lawrence, L., ... & Heffern, E. (2009). Emissions from coal fires and their impact on the environment. *U.S. Geological Survey Circular*, 1249, 1–80. <https://pubs.usgs.gov/circ/1249/>

- 6) Kuenzer, C., & Dech, S. (2011). Thermal infrared remote sensing of surface and underground coal fires. In C. Kuenzer & S. Dech (Eds.), *Thermal infrared remote sensing: Sensors, methods, applications* (pp. 429–451). Springer. https://doi.org/10.1007/978-94-007-6639-6_21
- 7) Kuenzer, C., Hecker, C., Zhang, J., Wessling, S., & Wagner, W. (2008). The potential of multidiurnal MODIS thermal band data for coal fire detection. *International Journal of Remote Sensing*, 29(3), 923–944. <https://doi.org/10.1080/01431160701395279>
- 8) ManSoR, S. B., Cracknell, A. P., Shilin, B. V., & Gornyi, V. I. (1994). Monitoring of underground coal fires using thermal infrared data. *International Journal of Remote Sensing*, 15(8), 1675–1685. <https://doi.org/10.1080/01431169408954104>
- 9) Melody, S. M., & Johnston, F. H. (2015). Coal mine fires and human health: What do we know? *International Journal of Coal Geology*, 152, 1–14. <https://doi.org/10.1016/j.coal.2015.11.001>
- 10) Pattanaik, D. S., Behera, P., & Singh, B. (2011). Spontaneous combustibility characterisation of the Chirimiri coals, Koriya district, Chhattisgarh, India. *International Journal of Geosciences*, 2(3), 336–342. <https://doi.org/10.4236/ijg.2011.23036>
- 11) Prakash, A., & Gupta, R. P. (1998). Land-use mapping and change detection in a coal mining area: A case study in the Jharia coalfield, India. *International Journal of Remote Sensing*, 19(3), 391–410. <https://doi.org/10.1080/014311698216053>
- 12) Prakash, A., & Gupta, R. P. (1999). Surface fires in Jharia coalfield, India: Their distribution and estimation of area and temperature from TM data. *International Journal of Remote Sensing*, 20(10), 1935–1946. <https://doi.org/10.1080/014311699212817>
- 13) Prakash, A., Gupta, R. P., & Saraf, A. K. (1997). A Landsat TM based comparative study of surface and subsurface fires in the Jharia coalfield, India. *International Journal of Remote Sensing*, 18(11), 2463–2469. <https://doi.org/10.1080/014311697218817>
- 14) Prakash, A., & Singh, G. (2009). Emission of carbon monoxide in atmosphere due to mine fire: A case study. *Asian Journal of Water, Environment and Pollution*, 6(1), 103–107.
- 15) Rosema, A., Guan, H., Veld, H., Vekerdy, Z., Ten Katen, A., & Prakash, A. (1999). Manual of coal fire detection and monitoring: Development and implementation of a coal fire monitoring and fighting system in China (245 pp.).
- 16) Sekertekin, A., & Bonafoni, S. (2020). Land surface temperature retrieval from Landsat 5, 7, and 8 over rural areas: Assessment of different retrieval algorithms and emissivity models and toolbox implementation. *Remote Sensing*, 12(2), 294. <https://doi.org/10.3390/rs12020294>
- 17) Singh, A. K., Singh, R. V. K., Singh, M. P., Chandra, H., & Shukla, N. K. (2007). Mine fire gas indices and their application to Indian underground coal mine fires. *International Journal of Coal Geology*, 69(3), 192–204. <https://doi.org/10.1016/j.coal.2006.03.011>
- 18) Sobrino, J. A., Raissouni, N., & Li, Z. L. (2001). A comparative study of land surface emissivity retrieval from NOAA data. *Remote Sensing of Environment*, 75(2), 256–266. [https://doi.org/10.1016/S0034-4257\(00\)00223-2](https://doi.org/10.1016/S0034-4257(00)00223-2)
- 19) Syed, T. H., Riyas, M. J., & Kuenzer, C. (2018). Remote sensing of coal fires in India: A review. *Earth-Science Reviews*, 187, 338–355. <https://doi.org/10.1016/j.earscirev.2018.10.009>

Experimental Study on Natural Convection Heat Transfer through a Composite Aluminum Heated Tube

Anjan Kumar Sahu¹, Sudhansu Sekhar Singh², Vishwa Mohan Behera³,
Ramesh Chandra Nayak^{4*,7}, Naresh Patra⁵, Shrinu Mohanty⁶

¹⁻⁶ Department of Mechanical Engineering, Synergy Institute of Technology,
Bhubaneswar-752101, Odisha, India.

⁷Post Doctoral Fellow, iHub Anubhuti-IIITD Foundation. New Delhi-110020, India

*Corresponding Author's Name: Ramesh Chandra Nayak
E-mail Id: rameshnayak23@gmail.com

Abstract: This study presents an investigation into natural convection heat transfer from a heated tube to the surrounding fluid. The primary driving force for fluid movement within the tube is the density difference arising from temperature variations. As the fluid temperature increases due to external conditions, it rises, while adjacent cooler fluid moves in to occupy the vacant space, facilitating continuous circulation. Natural convection is widely utilized in electronic equipment where multiple components are arranged on a compact board. Inadequate heat dissipation from the board to the surroundings can lead to system failure. Since forced convection is not practically feasible for such small boards, natural convection serves as an effective heat transfer mechanism. This work explores the characteristics of heat flow through a vertical tube and techniques to enhance natural convection heat transfer. The tube is constructed from a composite aluminum material (Al-Al Composite tube), which is non-magnetic and exhibits excellent thermal conductivity. Aluminum is chosen for its superior strength-to-weight ratio, lightweight nature, non-toxicity, and enhanced corrosion resistance, making it an ideal material for heat dissipation. The vertical tube has a diameter of 50 mm and a length of 550 mm, resulting in a length-to-diameter ratio of 11. A constant heat flux is maintained within the tube wall by regulating the external voltage supply. Internal rings, each with a 6 mm thickness and a square cross-section, are uniformly spaced inside the test section to enhance heat transfer efficiency.

Keywords: Density, vertical tube, natural convection, heat flow, internal rings.

1 Introduction

Natural convection heat transfer plays a crucial role in various engineering applications, including electronic cooling systems, heat exchangers, and energy storage devices. Unlike forced convection, which requires external mechanical aids such as fans or pumps, natural convection relies solely on buoyancy-driven fluid motion caused by temperature-induced density variations. This passive cooling mechanism is particularly advantageous in compact and sensitive systems where forced convection methods are impractical. One prominent application of natural convection is in electronic devices, where multiple components are integrated onto small circuit boards. Efficient heat dissipation is essential to prevent overheating, which can lead to performance degradation or system failure. Since forced convection solutions are often unfeasible in such confined spaces, enhancing natural convection heat transfer is a critical research focus. Pervez et al. [1] experimentally analyzed the role of internal rings inside a vertical composite tube. Their study found that the rings enhance secondary flows, leading to improved heat dissipation. The results provide insights into optimizing passive heat transfer enhancement in industrial applications. Mahmood et al. [2] numerically investigated natural convection in vertical annular tubes embedded with porous media. Their study used the finite volume method to analyze how Rayleigh number and

porosity affect heat transfer, concluding that porous media significantly enhance effective thermal conductivity. Kumar et al. [3] conducted a CFD-based study on barrel-shaped composite tubes. They discovered that heat transfer performance depends on the length-to-diameter ratio and diameter variations of the tube. The research suggests that altering tube geometry significantly influences the Nusselt number. Hata and Shibahara [4] investigated natural convection heat transfer in liquid sodium-cooled systems. Their study found that eccentricity of the grid spacer affects flow patterns and heat transfer efficiency, making it relevant for nuclear reactor cooling. Ramezanpour and Hosseini [5] conducted an experimental study on helical fins in vertical annular composite tubes. They concluded that fins enhance heat transfer rates by promoting turbulence, making them beneficial in heat exchanger applications. Zhao et al. [6] studied the effect of 3D fin structures on vertical tubes. Their study demonstrated that increasing fin height and axial fin pitch enhances heat transfer, with potential applications in HVAC systems. Khan et al. [7] provided analytical solutions for oscillating vertical cylinders. Their findings suggest that cosine and sine oscillations influence transient thermal response and overall convective heat transfer rates. Xu and Zhao [8] explored how layered materials impact heat transfer in vertical tubes. Their study highlighted the importance of thermal conductivity mismatches in designing efficient insulation materials. Cheng et al. [9] conducted a transient numerical analysis on natural convection. They found that thermal response time is significantly reduced by using high-conductivity inner layers. Choi and Eastman [10] investigated the use of nano fluids in vertical composite tubes. Their study found that Al_2O_3 and CuO nanoparticles increase the heat transfer coefficient by altering fluid properties. Patel et al. [11] examined the impact of surface roughness on natural convection in composite tubes. They found that rough surfaces enhance boundary layer disruption, improving heat dissipation. Ramesh et al. [12] integrated phase change materials (PCM) into composite tubes. Their study showed that thermal energy storage efficiency improved by up to 40%, making it viable for solar energy applications. Sundaram et al. [13] explored the role of magnetic fields in vertical composite tubes filled with ferro fluids. Their findings suggest that external magnetic field manipulation can regulate heat transfer rates. Rohsenow and Hartnett [14] analyzed how Rayleigh number variations impact natural convection. Their work provided essential correlations for heat transfer calculations. Eckert and Drake [15] compared free and forced convection in layered cylindrical structures. Their results suggest that forced convection is more effective in thin-walled tubes, while natural convection dominates thicker structures. Behnia and Copeland [16] studied ceramic composite tubes used in high-temperature industrial applications. They concluded that ceramic layers reduce thermal expansion effects while maintaining good heat dissipation. Kim et al. [17] analyzed the effect of thermal coatings on composite tubes. Their study found that silicone-based coatings significantly reduce heat loss, improving energy efficiency. Minkowycz et al. [18] provided mathematical models for annular composite tubes. Their findings indicate that optimizing inner-to-outer radius ratios enhances heat transfer. Sparrow and Gregg [19] examined how laminar and turbulent convection regimes affect heat transfer in composite tubes. Their research suggested that transitioning from laminar to turbulent flow improves heat dissipation by up to 30%. Holman and Bhattacharya [20] conducted a material optimization study, suggesting that hybrid metal-polymer composites offer a balance between thermal conductivity and structural integrity. Nayak et al. [21-30] from their work it was found that technology based designed models are helpful towards sustainability. In this work heat transfer through the composite aluminum (Al-Al Composite) vertical tube from a rectangular heated object is presented. The vertical tube is designed with and without internal rings. Due to natural convection heat transfer, fluid movement takes place inside the tube from bottom to top, temperature of fluid nearer to the inner surface of the tube is recorded for both conditions. A comparison statement in terms of fluid temperature inside wall surface for both conditions has also been presented in this work, it was found that the

vertical tube along with the internal ring is providing enhancement of heat transfer as compared to tube without internal ring.

2. Experimental Setup

The experimental arrangement for investigating natural convection heat transfer is illustrated in Figure 1. The system comprises a control panel, heating element, Al-Al composite tube, and support structure. The control panel is equipped with an ammeter, voltmeter, temperature indicator, voltage regulator, and selector switch, enabling precise monitoring and control of the experimental parameters.

The heating element consists of an electrical heating coil, which generates heat when supplied with current from the control panel. The heat generated by the coil is transferred to the Al-Al composite tube, which serves as the primary medium for conducting heat. The heated system is a rectangular aluminum body housing the electrical heating coil, securely placed on a support structure. In this setup, the composite aluminum vertical tube (Al-Al Composite Tube) is positioned within the heated system. The outer surface of the vertical tube remains in direct contact with the heating element, ensuring effective heat transfer.

Due to the composite nature of the test section, a constant heat flux is maintained throughout the tube. The vertical tube has a diameter of 50 mm and a length of 550 mm, resulting in a length-to-diameter ratio of 11. Heat is continuously supplied to the test section from an external source, and the transfer of heat occurs through natural convection within the vertical Al-Al composite tube. As heat is conducted through the tube, the air in contact with the inner wall of the composite tube is heated, causing a reduction in density. This leads to an upward fluid movement, creating a continuous convective flow from the bottom to the top of the tube.

To accurately measure the fluid temperature at the inner wall, eight thermocouples are strategically placed along the tube's surface. Additionally, an alternative configuration of the experiment incorporates square rings inside the test section. These rings have a thickness of 6 mm and are evenly spaced along the tube's length, facilitating enhanced heat transfer through controlled convection patterns..

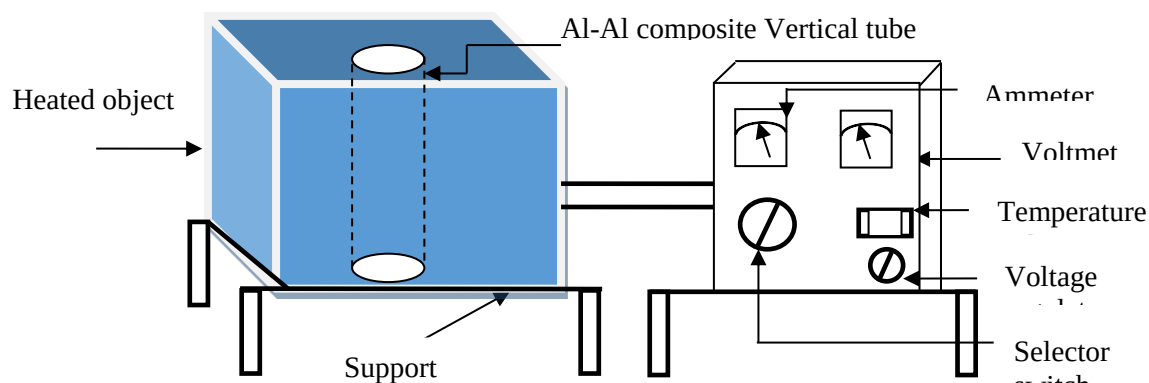


Fig.1. Experimental Setup

Figure 2 illustrates the cut-sectional view of the Al-Al Composite vertical tube, including the square rings placed inside the test section. To accurately measure the fluid temperature within the tube, eight thermocouples are strategically positioned along the inner wall of the test section.

The primary objective of this experimental setup is to establish a comprehensive database by recording the wall temperature of the Al-Al Composite vertical tube, both with and without internal rings, along the direction of heat flow. This data serves as a crucial reference for analyzing the effect of internal rings on natural convection heat transfer within the system.

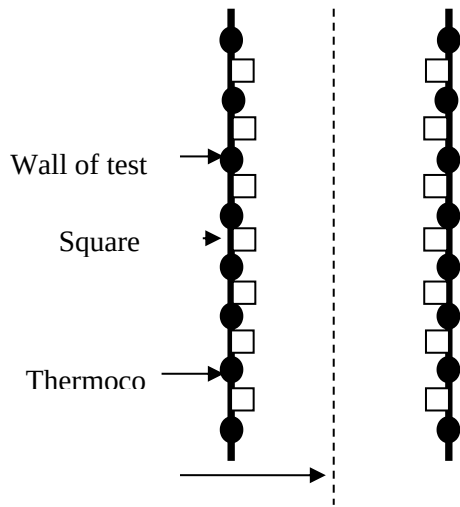


Fig.2. Al-Al Composite tube(cut sectional view)

3. Results and discussion

Experimental observations are carried out on a smooth composite tube with a length-to-diameter ratio of 11 under varying heat flux conditions. The heat flux inside the tube is controlled by adjusting the voltage supply through the control panel. To monitor temperature variations along the direction of heat flow, eight thermocouples are strategically placed at equal intervals along the height of the tube. Temperature measurements are recorded for eight different heat flux levels, ensuring a comprehensive analysis of the heat transfer characteristics.

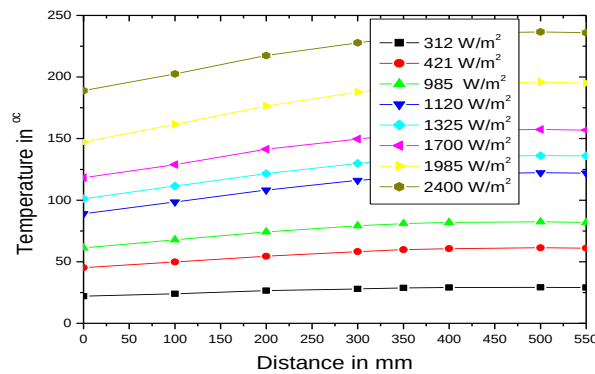


Fig. 3. Temperature values of thermocouples for 550 mm length tube without internal rings

Figure 3 illustrates the variations in surface temperature along the smooth composite tube wall at different positions for a length-to-diameter ratio of 11. The data indicates a gradual increase in temperature from the bottom to the top of the tube, aligning with the theoretical expectations of previous studies. However, a slight decrease in temperature is observed at the top, which can be attributed to the direct exposure of the inner fluid to atmospheric air at the uppermost position. Additionally, an increase in heat flux results in a corresponding rise in temperature values, owing to the greater heat input from the external source.

A similar experimental setup is employed using an Al-Al Composite vertical tube with internal square rings to analyze fluid temperature variations during heat transfer. In this configuration, seven rings with a square cross-section and 6 mm thickness are placed at equal intervals inside the tube to enhance heat transfer performance. The same range of heat flux values is maintained as in the previous experiment to ensure consistency in observations. Figure 4 presents the temperature measurements recorded by the thermocouples for different heat flux levels.

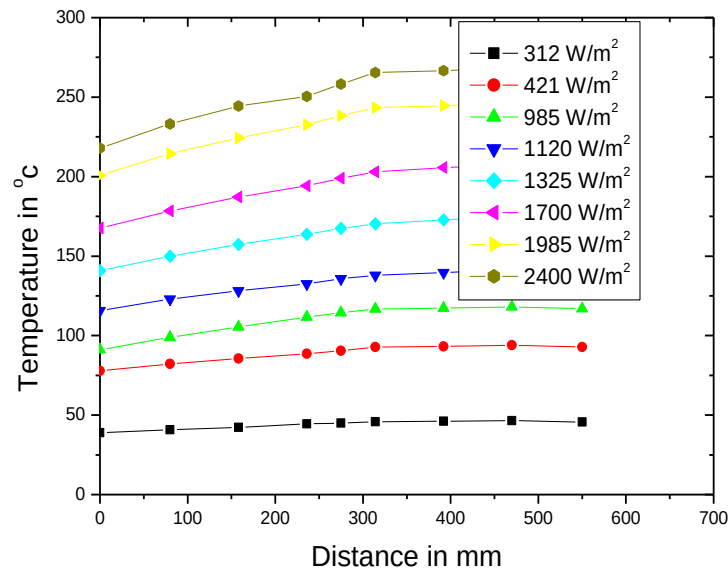


Fig. 4. Temperature values of thermocouples for 550 mm tube with internal rings

Figure 4 reveals that the temperature values recorded by the thermocouples along the direction of heat flow are higher in the tube with internal rings compared to the smooth composite tube. This indicates a significant enhancement in heat transfer due to the presence of square cross-sectional rings inside the tube.

The improvement in heat transfer occurs because the internal rings influence the fluid flow, causing an increase in the boundary layer thickness. The hot fluid moves toward the center, while the cooler fluid shifts toward the tube walls, leading to continuous mixing and disruption of the boundary layer. This process enhances convective heat transfer by promoting better thermal exchange within the fluid.

However, it is also observed that heat transfer enhancement is effective only up to a certain ring thickness limit. Beyond this threshold, the efficiency begins to decline, likely due to increased flow resistance or reduced convective mixing.

4. Outcomes and discussion

This study explores an effective method for enhancing heat transfer from a heated source. Based on experimental observations, several key findings have been identified:

- The top thermocouple records the highest temperature compared to the lower thermocouples, indicating that the average rate of heat flow increases in the direction of air movement inside the tube.
- The variation in local wall temperatures for a length-to-diameter ratio (L/D) of 11 and different heat flux levels shows a gradual increase along the tube height, with a slight decrease near the top due to heat dissipation into the surrounding air.
- Experimental results reveal that the temperature values of thermocouples are higher in the presence of internal rings, confirming that heat transfer enhancement occurs due to the obstruction-induced flow disturbances inside the tube.
- The fluid temperature near the inner tube wall is consistently higher compared to the core region in the smooth tube configuration.
- The temperature of the air near the composite tube wall increases with increasing heat flux, demonstrating the direct influence of external heating on convective heat transfer.

Acknowledgment

This study is supported by iHub Anubhuti- IIITD Foundation", Okhla Industrial Estate, Phase III, New Delhi, India – 110020 and Synergy Institute of Technology, Bhubaneswar, Odisha.

References

1. Pervez, M. S., Ahmed, R., & Khan, T. (2025). Effect of internal rings on heat transfer enhancement in vertical composite tubes. *Materials Research Proceedings*, 49, 21-28.
2. Mahmood, M. A., Rehman, S., & Ali, K. (2024). Numerical investigation of natural convection in vertical annular composite tubes embedded with porous media. *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences*, 66(2), 65-83.
3. Kumar, A., Singh, V., & Patel, R. (2024). Computational fluid dynamics study on barrel-shaped composite tubes for heat transfer enhancement. *Numerical Heat Transfer, Part A: Applications*.
4. Hata, K., & Shibahara, M. (2022). Natural convection heat transfer in liquid sodium-cooled vertical composite tubes with eccentric grid spacers. *Heat and Mass Transfer*, 58, 1607-1636.
5. Ramezanpour, M., & Hosseini, R. (2020). Experimental investigation of helical fins on natural convection in vertical annular composite tubes. *Experimental Heat Transfer*, 33(3), 226-244.
6. Zhao, Y., Wang, H., & Li, J. (2020). Enhancing natural convection heat transfer using 3D fins in vertical composite tubes. *Chemical Engineering Transactions*, 81, 1087-1092.
7. Khan, M., Ahmed, T., & Siddiqui, A. (2020). Analytical solutions for heat transfer in oscillating vertical cylinders. *International Journal of Heat and Mass Transfer*, 150, 119-128.
8. Xu, G., & Zhao, Y. (2023). Impact of layered material properties on heat transfer in composite tubes. *Journal of Thermal Analysis and Calorimetry*.
9. Cheng, P., Liu, Z., & Sun, W. (2023). Transient numerical analysis of natural convection in vertical composite tubes. *International Journal of Thermal Sciences*.

10. Choi, S. U. S., & Eastman, J. A. (2023). Nanofluid applications for enhancing natural convection in composite tubes. *ASME Fluids Engineering Division*.
11. Patel, R., Sharma, K., & Verma, S. (2023). The effect of surface roughness on heat transfer in vertical composite tubes. *International Journal of Thermal Engineering*.
12. Ramesh, K., Gopal, P., & Anand, R. (2023). Hybrid composite tubes with phase change materials for thermal storage applications. *Solar Energy Materials and Solar Cells*.
13. Sundaram, S., Reddy, B., & Kumar, M. (2022). Influence of magnetic fields on natural convection in ferrofluid-filled vertical composite tubes. *Journal of Magnetohydrodynamics*.
14. Rohsenow, W. M., & Hartnett, J. P. (2022). The impact of Rayleigh number variations on natural convection in vertical composite tubes. *McGraw-Hill Heat Transfer Handbook*.
15. Eckert, E. R. G., & Drake, R. M. (2021). Free and forced convection comparison in composite cylindrical structures. *Heat and Mass Transfer Analysis*.
16. Behnia, M., & Copeland, D. (2021). High-temperature natural convection in ceramic composite tubes. *Journal of High-Temperature Thermal Engineering*.
17. Kim, W., Park, S., & Choi, J. (2021). The role of thermal coatings in improving heat transfer efficiency in composite tubes. *Journal of Applied Thermal Coatings*.
18. Minkowycz, W. J., & Sparrow, E. M. (2021). Natural convection in annular composite cylinders: A theoretical study. *International Journal of Heat and Mass Transfer*.
19. Sparrow, E. M., & Gregg, J. L. (2020). Laminar and turbulent convection regimes in composite tubes. *International Journal of Thermal Sciences*.
20. Holman, J. P., & Bhattacharya, S. (2020). Optimization of composite tube materials for enhanced convection performance. *Journal of Heat Transfer Engineering*.
21. Nayak, R.C., Roul, M.K., Sarangi, S.K., Natural convection heat transfer in heated vertical tubes with internal rings, *Archives of Thermodynamics*, 2018, 39(4), pp. 85–111.
22. Nayak, R.C., Roul, M.K., Sarangi, S.K. Experimental investigation of natural convection heat transfer in heated vertical tubes with discrete rings, *Experimental Techniques*, 2017, 41(6), pp. 585–603.
23. Nayak R.C., Roul M.K., Sarangi S.K.:Experimental Investigation of Natural Convection Heat Transfer in Heated Vertical Tubes. *International Journal of Applied Engineering Research*, Vol. 12 (2017)2538–2550.
24. Nayak, R.C., Roul, M.K., Sarangi, S.K., Sarangi, A., Sarangi, A. Enhancement of Natural Convection Inside Vertical Tubes Using Internal Obstacles as Rings with Different Arrangements, *Lecture Notes in Mechanical Engineering*, 2021, 52, pp. 191–201.
25. Sahoo, A., Chandra Nayak, R., Senapati, A.K., Roul, M.K. Validation of experimental results with theoretical by using ANSYS workbench on vertical tube subjected to natural convection heat transfer without internal obstacles, *Materials Today: Proceedings*, 2022, 52, pp. 1348–1353.
26. Nayak, R.C., Samal, C., Roul, M.K., Padhi, P. A New Irrigation System Without Any External Sources (2023) *Journal of The Institution of Engineers (India): Series A*, 104 (2), pp. 281-289.
27. Nayak, R.C., Roul, M.K. Technology to Develop a Smokeless Stove for Sustainable Future of Rural Women and also to Develop a Green Environment (2022) *Journal of The Institution of Engineers (India): Series A*, 103 (1), pp. 97-104.
28. Nayak, R.C., Roul, M.K., Roul, P.D. Design and development of smokeless stove for a sustainable growth (2022) *Archives of Thermodynamics*, 43 (1), pp. 109-125.

29. Nayak, R.C., Roul, M.K., Sarangi, S.K. Innovative Methods to Enhance Irrigation in Rural Areas for Cultivation Purpose (2021) Journal of The Institution of Engineers (India): Series A, 102 (4), pp. 1045-1051.
30. Nayak, R.C., Roul, M.K., Sarangi, S.K., Sarangi, A., Sarangi, A. Forced Draft and Superheated Steam for Design and Development of Community Smoke Less Chulha to Help Women in Rural Areas (2021) Lecture Notes in Mechanical Engineering, pp. 93-102.

TRINETRA: Empowering the Visually Impaired with In-telligent Vision Assistance

Atul Tripathi¹, Isha Pathak Tripathi², Deepika Dhinva³, Harshit Shakya⁴, Pranshu Kharniwal⁵

¹ University School of Automation & Robotics, Guru Gobind Singh Indraprastha University, New Delhi

²⁻⁵ Department of Computer Science and Engineering, Indian Institute of Information Technol-ogy, Kota
atul.usar@ipu.ac.in

Abstract. Visually impaired individuals often struggle to lead fulfilling lives be-cause they are unable to notice their surroundings as clearly as average people. Blind persons constantly need someone to inform them of the world around them. Mobile and camera technologies are becoming commonplace in our daily lives. Therefore, these technologies can help the blind and visually impaired be aware of their surroundings and live fulfilling lives. A Convolution Neural Network (CNN) model is deployed that interacts with the environment and generates a caption to assist the visually impaired person. The model called Trinetra consists of Resnet50, Xception, and VGG16, which are integrated to make the model more efficient.

Keywords: VGG16, CNN, LSTM , RNN, Vision impairment, Audio to text

1. Introduction

People who are blind or partially sighted encounter several difficulties going about their daily lives and finding their way about. People who are visually impaired sometimes have difficulty reading even the most basic of warnings, such as those posted on walls or signs found nearby. Navigation and orientation aids are now more important than ever. With the development of technologies like cameras, blind or visually impaired people on the go can live independently. The development of object detection and im-age categorization has made it possible to automatically produce one or more words to comprehend a picture’s visual information. People with visual impairment can use text-to-speech technology as an assistive technology. As a result, the aim of this work is to develop a module for machine learning that is capable of converting visuals into spoken language. This will enable people who are blind or visually impaired to lead normal lives on their own without the assistance of others.

To aid the visually impaired, the model converts the image into normal language like English, extracting features from the image using a Convolution Neural Network (CNN), and finally converts that description into voice. The word embeddings of fragmentary sentences are put into Long Short-Term Memory (LSTM), which excels at predicting sequences. This method uses a merge architecture to provide image captions. Language and perceptual features are encoded separately and combined at the multi-modal layer, where predictions are formed. A hybrid technique of the Bilingual Evaluation Understudy (BLEU) and Google Text To Speech (GTTS) module in Python is developed to measure the generated caption and finally convert them into audio.

2. Literature Review

In their investigation of picture captioning, Alam [1] examined the usage of deep learn-ing. The purpose of this research is to investigate existing approaches for the purpose of picture captioning. They discussed the part of the process that was considered to be the most significant, as well as the CNN-inspired template for the creation of photo captions.

Amirian [2] has developed a model in order to provide captions for videos and photo-graphs. They are able to give captions for still photos and video frames using methods from deep

learning. They have developed the field to include autonomous caption creation for photos and movies using deep learning so that they can provide assistance to those who have visual impairments.

Contextual information served as the basis for Bang and Kim's [3] creation of picture captions. This approach generates data that has a temporal, geographical, and graphical component to it. The textual description method generates sentences by using photo captioning as its source material. Techniques for organizing data are used in order to organize data such as photographs, text, time, and location. It is possible to make efficient use of context-based information while managing data obtained by UAVs.

Deng [4] developed a model that uses an adaptive attention system in conjunction with a DenseNet network so that it can generate descriptions for photos. Based on the findings of this research, an adaptable attentive strategy with a visual sensor is suggested. DenseNet is used by the model in order to retrieve the general characteristics of the picture. A Long Short-term memory network is utilized as a speech generation architecture in order to enhance the overall quality of the picture captions that are produced by the system. The COCO dataset and the Flickr30k dataset were used in the creation of the descriptions for the photographs.

Hossain [5] used a model for picture-to-text synthesis to provide better captions for the pictures he took. The processes train and evaluate the models by making use of human annotations of pictures, which are generated using deep learning models. They offer a method for automatically captioning pictures that makes use of both actual and fictitious data in order to educate and evaluate the model. To create the synthetic visuals, a word-to-image generator that was constructed using a generative adversary network was used.

image caption from similar images with captions. To produce picture descriptions from millions of weakly annotated photos, Gong et al. [6] use stacked auxiliary embedding approaches. Ordonez et al. [7] search the Flickr database for similar photos and return the images' descriptions to a query based on millions of photographs and their descriptions. Sun et al. [8] first cluster related phrases and images together using semantic and visual similarity scores, then obtain the caption of the target image from the captions of similar photos in the same cluster.

The CNN with object detection method that is offered by Iwamura [9] was used in order to write the descriptions for the photographs. They came up with an algorithm that was constructed using the MSCOCO dataset, the MSRVT2016-Image dataset, and the free-copyright pictures dataset. They were able to achieve caption accuracy as well as picture captions by using these datasets.

An investigation on the picture captioning on CNN was carried out by Kalra and Leekha [10]. They developed the model by combining a recurrent neural network (RNN) for modelling and predicting words with a CNN for embedding images. They suggested a model as a method for the production of text descriptions.

Mathur [11] conducted studies on a number of different models pertaining to photo captioning. Computer vision is used to comprehend pictures, NLP is utilized to comprehend words, and machine learning is utilized to develop descriptions for photos. While RNN is used for the generation of phrases, CNN is utilized for the comprehension of image data.

3. Methodology

The goal of the proposed work is to construct a model that will understand the context of a picture. This will be accomplished by evaluating the image, describing it in a natural language such as English, and then converting it into audio form. Image captioning refers to the act of developing a written description to accompany a collection of visual representations. Picture captioning is a research problem that has lately been gaining in importance with time. To build an efficient model for image captioning, the first thing that has to be done is feature extraction

from the picture, and then the captions for the photos need to be generated. We used three different CNN architectures, namely VGG16, ResNet50, and Xception, for the purpose of feature extraction; however, LSTM was utilized for the generation of captions. In light of the precision of the models, we have settled on the optimal combination as our option of choice. The main stages of the proposed model are as follows:

- a. The preliminary processing of the dataset
- b. Feature Extraction
- c. Producing Captions (Captioning) Dataset

The dataset is taken from Kaggle i.e. Flickr 8k dataset.

3.1 RESNET 50

The proposed deep learning model has been trained for the ResNet50 group. This model is considered to be pre-trained if they have previously received training for another profession that is unrelated to the one that it is now doing. Because the prior job's knowledge may be applied to the current work, this is a suitable spot, It would be essential to have a substantial amount of data to identify all of the unidentified parameters (in millions). Finding a dataset with large tags attached to it is quite challenging. We make use of the proposed models that have previously been trained on enormous quantities of data to do challenging tasks that include thousands of different classes. A variation of ResNet called ResNet50 is comprised of one MaxPool layer, one Average Pool layer, and 48 Convolution layers. It consists of operations using floating-point numbers. It is a model that is quite well known in the ResNet community. Figure 1 depicts an example of caption generation using Resnet50, and its architecture is shown in Figure 2.



Fig. 1. Resnet50 Example

3.2 Xception

Researchers at Google came up with the concept first. Google describes the inception modules found in convolutional neural networks as an intermediary stage that occurs between the process of regular convolution and the process of depth-wise separable convolution. The data will be fed through the input flow, then proceed directly through the middle flow a total of eight times, and finally conclude their journey by travelling through the exit flow. Batch normalization may be applied to any convolutional and separate convolution layers. A representation of the Xception Architecture may be seen in Figure 3.

3.3 VGG16

VGG16 is regarded to be one of the most successful architectural models for vision. The absence of a significant number of hyperparameters is the feature of VGG16 that stands out as being the most distinctive. Instead, the developers focused their attention on having convolution layers of a 3x3 filter with a stride 1, and they consistently used the same padding in addition to the maxpool layer of a 2x2 filter operating in stride 2. Given this, it is apparent that the layout of the convolution and maximum pool layers is consistent over the whole of the building's design. In the end, it has two completely charged batteries (two FC. linked layers), then the output is processed by a softmax. The 16 in the notation VGG16 indicates that the structure consists of 16 layers with weights. .

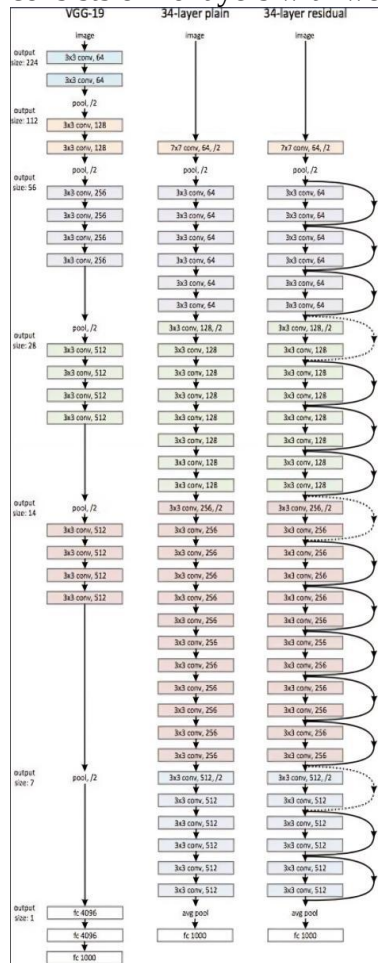


Fig. 2. Resnet50 Architecture

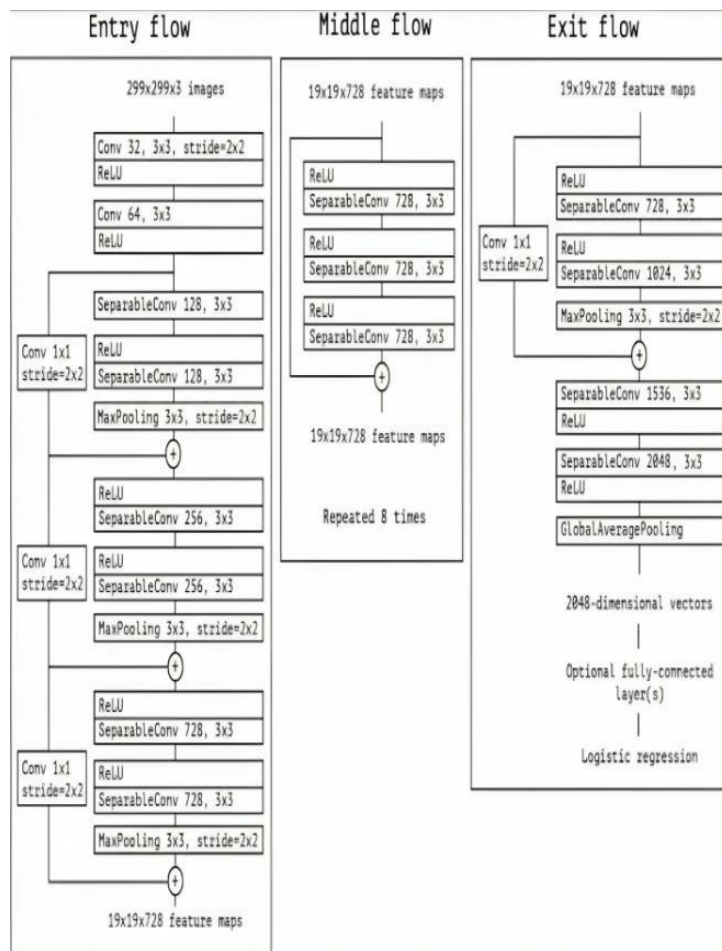


Fig. 3. Xception Architecture



Fig. 4. Xception Example

Creating Subtitles and Captions (Captioning) Long short-term memory, also known as LSTM, is put to use in the process of generating a description of the picture using the information that was retrieved from it using the CNN models described before. The Long Short-Term Memory (LSTM) neural network is a sort of RNN that may solve issues involving sequence prediction. The majority of applications for LSTM center on word prediction based on the context of prior words. LSTM is applied to the features (features extracted from images using CNN models), in such a way that each feature in the image is taken as a single word. After that, LSTM saves that single word in the storage, and after all of the words have been gathered, it organizes the words in the proper order by forming suitable associations among them, and then it constructs a meaningful basic English sentence. This statement was automatically created to serve as the picture's caption. A caption is nothing more than a brief explanation of an image written in plain English. The GTTS is used to transform the created output sentence into an audio file (Google text-to-speech).

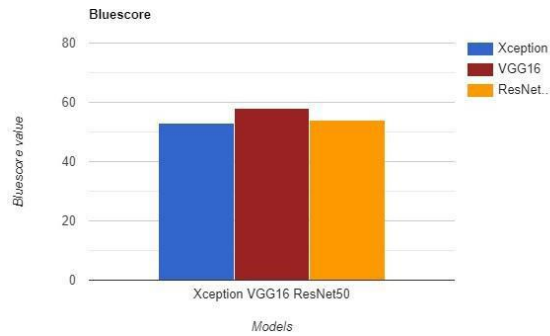
Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 224, 224, 64)	1792
conv2d_2 (Conv2D)	(None, 224, 224, 64)	36928
max_pooling2d_1 (MaxPooling2D)	(None, 112, 112, 64)	0
conv2d_3 (Conv2D)	(None, 112, 112, 128)	73856
conv2d_4 (Conv2D)	(None, 112, 112, 128)	147584
max_pooling2d_2 (MaxPooling2D)	(None, 56, 56, 128)	0
conv2d_5 (Conv2D)	(None, 56, 56, 256)	295168
conv2d_6 (Conv2D)	(None, 56, 56, 256)	590080
conv2d_7 (Conv2D)	(None, 56, 56, 256)	590080
max_pooling2d_3 (MaxPooling2D)	(None, 28, 28, 256)	0
conv2d_8 (Conv2D)	(None, 28, 28, 512)	1180160
conv2d_9 (Conv2D)	(None, 28, 28, 512)	2359808
conv2d_10 (Conv2D)	(None, 28, 28, 512)	2359808
max_pooling2d_4 (MaxPooling2D)	(None, 14, 14, 512)	0
conv2d_11 (Conv2D)	(None, 14, 14, 512)	2359808
conv2d_12 (Conv2D)	(None, 14, 14, 512)	2359808
conv2d_13 (Conv2D)	(None, 14, 14, 512)	2359808
max_pooling2d_5 (MaxPooling2D)	(None, 7, 7, 512)	0
flatten_1 (Flatten)	(None, 25088)	0
dense_1 (Dense)	(None, 4096)	102764544
dropout_1 (Dropout)	(None, 4096)	0
dense_2 (Dense)	(None, 4096)	16781312
dropout_2 (Dropout)	(None, 4096)	0
dense_3 (Dense)	(None, 2)	8194
Total params: 134,268,738		
Trainable params: 134,268,738		
Non-trainable params: 0		

Fig. 5. VGG16 Architecture

4. Results and Discussion

In order to train the Convolutional Neural Network architecture, we use three different models: Xception, VGG16, and ResNet 50. A higher bluescore indicates that the model is working more effectively. After doing the calculations necessary to determine each model's bluescore, we discovered that three of the models provide results with accuracy that are quite close to one another. The suggested model for deep learning has produced significant captions for the photos that are analogous to the ones found in the dataset that we have utilised.

In order to extract features from the data, we made use of a Convolutional Neural Network model, and the LSTM model was used in order to generate captions. The system is trained using three different CNN architectures, namely the Xception, VGG16, and ResNet50 models. Among the other two models of CNN, combining VGG16 and



LSTM resulted in the maximum accuracy of 60 percent over a period of fifty epochs.

Fig. 7. Blue Score

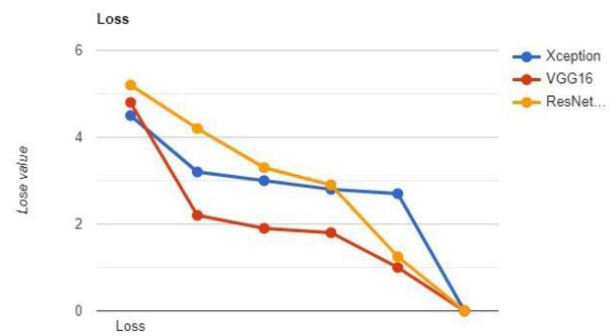


Fig. 8. Loss value Comparison

BLEU, which stands for "Bilingual Evaluation Understudy," is a score that is de-rived by contrasting a proposed text translation with one or more reference translations. A score of 1.0 is awarded when there is a perfect match, whereas a score of 0.0 is given when there is a perfect mismatch. In other words, BLEU is used as an assessment metric in order to determine how accurate the model is. The BLEU score may be determined for a variety of different N values. When we tested our model to see how accurate it was, we utilized N = 4. BLEU-1 makes use of the unigram precision score, and its calculation goes as follows: $accu\ 1 = corpus\ bleu\ (actual, predicted, weights = (1.0, 0, 0, 0))$; where actual is the actual result for the input that was provided, and predicted is the output that we receive using our model.

5. Conclusion and future Scope

The suggested deep learning model has produced pertinent descriptions for pictures that resemble those in the dataset we utilized. For feature extraction, we employed a convolutional neural network model, and for caption generation, we used an LSTM model. VGG16, Xception, and ResNet50 CNN architectures are used to train the system. Out of the three CNN models, the combination of VGG16 and LSTM had the greatest ac-curacy, coming in at 60 percent after 50 iterations.

The proposed model is based on CNN-LSTM, which encodes images and converts them into corresponding sentences based on the learned image features. This system aims to help visually challenged people know their surroundings. This can be done by capturing images of their surroundings, image content can be converted into a caption(description), and then converted to voice. With this system, not only can normal people, but also visually challenged people can get the best help from the latest technologies like camera smartphones.

As of now, the scope of the system is limited to analyzing captured images. In the future, this can be made to capture and analyze live video footage, providing live commentary about what happens in the surroundings. This makes the person feel like some-one is accompanying them. The number of reference captions can be increased, and instead of using a single model, ensembles can be used to achieve maximum accuracy.

References

1. Mohammad Shahnawaz Alam, Vaishali Narula, Ruchika Haldia, and Gitanjali Nikam Ganpatrao. An empirical study of image captioning using deep learning. pages 1039–1044, 2021.
2. Soheyla Amirian, Khaled Rasheed, Thiab R Taha, and Hamid R Arabnia. Automatic image and video caption generation with deep learning: A concise review and algorithmic overlap. *IEEE Access*, 8:218386–218400, 2020.
3. Seongdeok Bang and Hyoungkwan Kim. Context-based information generation for managing uav-acquired data using image captioning.
4. *Automation in Construction*, 112:103116, 2020.
5. Zhenrong Deng, Zhouqin Jiang, Rushi Lan, Wenming Huang, and Xiaonan Luo. Image captioning using densenet network and adaptive attention. *Signal Processing: Image Communication*, 85:115836, 2020.
6. Md Zakir Hossain, Ferdous Sohel, Mohd Fairuz Shiratuddin, Hamid Laga, and Mohammed Bennamoun. Text to image synthesis for improved image captioning. *IEEE Access*, 9:64918–64928, 2021.
7. G. Kulkarni V. Ordonez and T. L. Berg. Im2text: Describing images using 1 million captioned photographs,. *Advances in Neural Information Processing Systems*, page 1143–1151, 2011.
8. M. Hodosh J. Hockenmaier Y. Gong, L. Wang and S. Lazebnik. Improving image sentence embeddings using large weakly annotated photo collections. *European Conference on Computer Vision*, 16:529–545, 2014.
9. C. Gan C. Sun and R. Nevatia. Automatic concept discovery from parallel text and visual corpora. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 8:2596–2604, 2015.
10. Kiyohiko Iwamura, Jun Younes Louhi Kasahara, Alessandro Moro, Atsushi Yamashita, and Hajime Asama. Image captioning using motion-cnn with object detection. *Sensors*, 21(4):1270, 2021.
11. Saloni Kalra and Alka Leekha. Survey of convolutional neural networks for image captioning. *Journal of Information and Optimization Sciences*, 41(1):239–260, 2020.
12. Pratistha Mathur et al. A survey on various deep learning models for automatic image captioning. In *Journal of Physics: Conference Series*, volume 1950, page 012045. IOP Publishing, 2021.

Dynamic Modeling and Torque Analysis of a Humanoid Dual-Arm Manipulator under External Disturbances

Shubham Vinayak Waman¹ and Dr Sandeep Anasane²

^{1,2} Department of Manufacturing Engineering and Industrial Management,
COEP Technological University, Pune, India

1wamansv23.mfg@coeptech.ac.in, 2ssa.mfg@coeptech.ac.in

Abstract

Humanoid dual-arm manipulators are increasingly pivotal in complex tasks such as collaborative manufacturing, assistive healthcare, and autonomous service robotics, where precise dynamic modeling and robust interaction with unstructured environments are essential. Despite advancements, accurately compensating for external disturbances while maintaining stability remains a critical challenge, particularly in systems requiring coordinated motion and force control. This work presented in this manuscript provides a dynamic model of a dual-arm humanoid manipulator and studies its performance subject to some external wrench. The Jacobian matrices J_v and J_w will be singled out in this work, being the functions which relate angular and linear velocities with the forces and torques acting on the intermediate links. In the same manner, the transposed Jacobian matrices relate external wrench to the torque, leading to a more accurate modeling of the environmental factors. The dynamic model Dual arm humanoid was evaluated under two conditions: (1) with no external force applied to the end effector and (2) with an external force applied to the end effector. The corresponding difference in the joint torques in these two scenarios is found to be equal to the product of the Jacobian transpose and the externally applied forces, representing the external torques due to disturbances. Hence validating the model, further verification is also performed on intermediate joints, which lead to consistent results and hence establish the validity of the method. Dynamic computations using Recursive Newton-Euler Algorithm (RNEA) highlight the model's flexibility to take external disturbances into account and to allow for a reliable and stable manipulation.

Keywords: Recursive Newton-Euler Algorithm (RNEA), Humanoid dual-arm manipulator, Dynamic modeling, Jacobian matrices

1. Introduction

Humanoid dual-arm manipulators are becoming more prominent in fields like collaborative manufacturing, assistive healthcare, and service robotics. These systems promise to enhance dexterity and versatility by emulating human bimanual skills, enabling us to perform tasks that single-arm systems cannot achieve. As robots are being deployed in unstructured and dynamic environments, the need for robust and accurate dynamic models grows. The big challenge in deploying these manipulators is not only to coordinate the motion between the arms but also to compensate for the external disturbances that occur during interaction with the environment. Dual arm systems are more complex than single arm systems due to the coupling between the arms and the additional degrees of freedom required to control force and motion. Moreover, environmental uncertainties such as variable contact forces and external perturbations add to the complexity.

Traditional control methods may not work well with these challenges, that's why we need a more accurate dynamic model that integrates both kinematic and dynamic effects, especially in the presence of disturbances. In reality, even little unmodeled forces or environmental changes can cause large variations in joint torques, which will negatively affect the manipulator's performance and stability. Consequently, a dynamic model that not only captures the intrinsic properties of the robot (such as inertia, Coriolis, and gravitational forces) but also quantifies the effects of external wrench is essential. To correlate these disturbances to the resulting joint torques, a particularly effective method is to use Jacobian matrices to transfer external end-

effector forces into the joint space. In this work, we present an integrated dynamic model that incorporates both kinematic and dynamic effects when subjected to external wrench for humanoid dual-arm manipulators. Our approach offers a strong foundation for analyzing and minimizing disruptions, improving the system's overall performance and stability.

2. Literature Review

Humanoid dual-arm manipulators have gained significant attention in robotics due to their ability to perform complex tasks involving coordination and dexterity. Research in this field has primarily focused on dynamic modeling, torque analysis, and external disturbance compensation to improve performance and adaptability in real-world environments. This section reviews existing studies on these aspects and identifies research gaps.

2.1. Dynamic Modeling of Dual-Arm Manipulators

Dynamic modeling of robotic manipulators is essential for understanding their motion, control, and stability. Several methods have been employed to derive the dynamic equations governing these systems. The Euler-Lagrange formulation has been widely used for modeling multi-degree-of-freedom (DOF) robotic arms due to its systematic approach in deriving equations of motion [1]. Alternatively, the Newton-Euler method provides recursive computations that are computationally efficient, making it suitable for real-time applications [2].

A key reference for this study is the work by Xin Jing et al. [3], which presents a recursive dynamic modeling and control approach for a dual-arm manipulator with elastic joints. Their method enhances the computational efficiency of dynamic equations while incorporating joint flexibility, making it particularly relevant for humanoid robotic systems. The study highlights the challenges in controlling elastic-jointed manipulators and provides a recursive formulation that can be extended to variable-mass conditions.

Additionally, Buondonno (2019) [4] proposed a recursive Newton-Euler algorithm specifically for robots with elastic joints, further improving torque calculations and control strategies. Similarly, Khalil (2018) [5] provided a comprehensive analysis of the Newton-Euler method for dynamic modeling of robots, emphasizing its advantages in high-DOF systems.

2.2. Torque Analysis in Humanoid Manipulators

Torque estimation and optimization are critical for ensuring smooth and energy-efficient operation of humanoid robotic arms. Several researchers have analyzed joint torques under different operating conditions. A study by Nejc Likar et al. (2021) [6] introduced an external joint torque-based contact estimation method, which enhances the accuracy of force interactions between the manipulator and external objects. This method is essential for real-world applications where manipulators must adapt to unknown external disturbances.

Despite these advancements, limited research has been conducted on how dynamically changing payloads affect joint torques. Existing models often assume a constant mass distribution, making them less effective in scenarios where objects with unknown or varying mass are manipulated.

2.3. Effects of External Disturbances on Manipulator Dynamics

External disturbances such as environmental forces, sudden impacts, and payload shift significantly affect robotic arm stability and accuracy. Traditional disturbance rejection methods include feedback linearization and adaptive control, which modify control inputs based on disturbance estimates [7]. More recent approaches utilize machine learning-based disturbance observers to enhance real-time adaptability [8].

The study by Xin Jing et al. (2021) [3] also considers the effects of joint elasticity on disturbance rejection, offering a robust framework for handling external wrench in dual-arm

manipulation. However, further research is needed to extend this framework to scenarios involving variable mass dynamics, which remains an open challenge.

5. Methodology

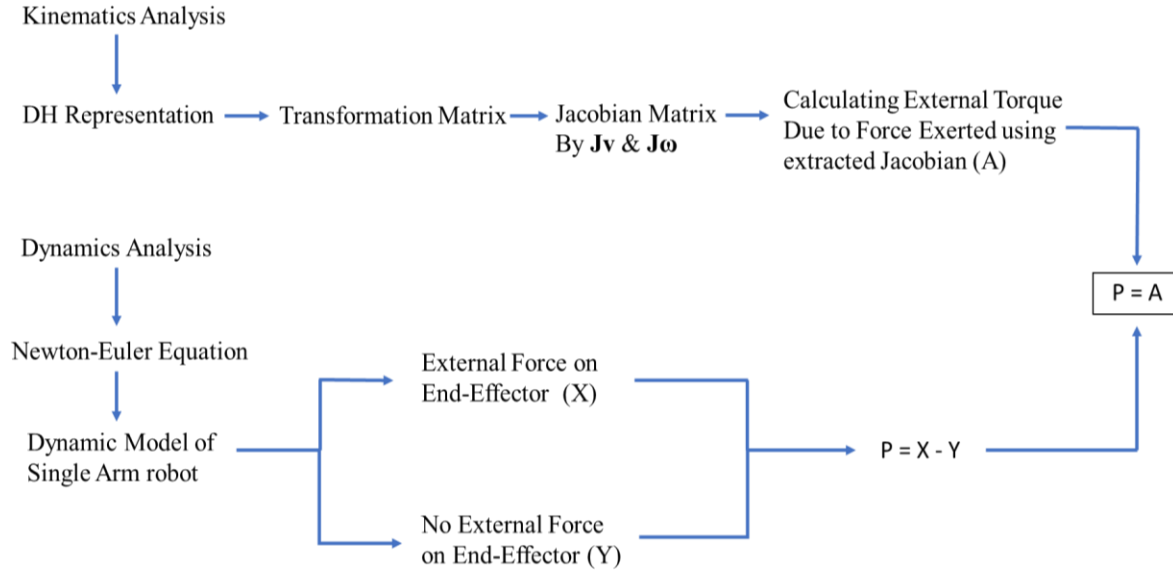


Fig 1: Block Diagram of Methodology for Joint Torque Estimation Under External End-Effector Forces

4.1 Dual-Arm Manipulator's Kinematic Model

Representation of Modified Denavit-Hartenberg (MDH) for Dual-Arm Humanoid Manipulator: An alternative to the conventional DH representation in robotic kinematics is the Modified Denavit-Hartenberg (MDH) convention. The primary difference is that in MDH, the joint axis is aligned with the Z-axis of the previous frame, leading to a different transformation matrix structure.

MDH parameters are described as follow:

- α_i – Rotation about the X-axis from Z_{i-1} to Z_i .
- a_i – Translation along the X-axis from Z_{i-1} to Z_i .
- θ_i – Rotation about the Z-axis from X_{i-1} to X_i .
- d_i – Translation about the Z-axis from X_{i-1} to X_i

The transformation matrix for Modified DH is given by:

$${}^{i-1}T_i = \begin{bmatrix} \cos \theta_i & -\sin \theta_i & 0 & a_{i-1} \\ \sin \theta_i \cos \alpha_{i-1} & \cos \theta_i \cos \alpha_{i-1} & -\sin \alpha_{i-1} & -\sin \alpha_{i-1} d_i \\ \sin \theta_i \sin \alpha_{i-1} & \cos \theta_i \sin \alpha_{i-1} & \cos \alpha_{i-1} & \cos \alpha_{i-1} d_i \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Kinematic Representation: The transformation matrices T_{i0} define the spatial configuration of each link relative to the base frame. These matrices map joint movements to the location and orientation of the manipulator's end-effector by incorporating joint angles and link parameters.

For arm a: ${}^0T_e^a = {}^0T_1^a \cdot {}^1T_2^a \cdot \dots \cdot {}^{n-1}T_n^a$

For arm b: ${}^0T_e^b = {}^0T_1^b \cdot {}^1T_2^b \cdot \dots \cdot {}^{n-1}T_n^b$

For the manipulator linkages, the homogeneous transformation matrices are defined as follows:

$$\begin{aligned}
 T_1^0 &= \begin{bmatrix} c\theta_1 & -s\theta_1 & 0 & 0 \\ s\theta_1 & c\theta_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad T_2^0 = \begin{bmatrix} c\theta_{12} & -s\theta_{12} & 0 & 11c\theta_1 \\ s\theta_{12} & c\theta_{12} & 0 & 11s\theta_1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad T_3^0 = \begin{bmatrix} c\theta_{12} & s\theta_{12} & 0 & 11c\theta_1 + 12c\theta_{12} \\ s\theta_{12} & -c\theta_{12} & 0 & 11s\theta_1 + 12s\theta_{12} \\ 0 & 0 & -1 & -d_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
 T_4^0 &= \begin{bmatrix} c\theta_{12}c\theta_4 + s\theta_{12}s\theta_4 & -c\theta_{12}s\theta_4 + s\theta_{12}c\theta_4 & 0 & 11c\theta_1 + 12c\theta_{12} \\ s\theta_{12}c\theta_4 - c\theta_{12}s\theta_4 & -s\theta_{12}s\theta_4 - c\theta_{12}c\theta_4 & 0 & 11s\theta_1 + 12s\theta_{12} \\ 0 & 0 & -1 & -d_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}
 \end{aligned}$$

4.2 Geometric Jacobian for Dual-Arm Manipulator

The Jacobian matrix serves as a crucial instrument in robotic kinematics and dynamics which connects joint velocities to the end effector's linear and angular velocities. The Jacobian matrix demonstrates how end-effector motion results from minor joint angle or displacement changes. The Jacobian matrix plays a crucial role in dual-arm manipulators by enabling trajectory planning, force control and on-the-fly motion optimization which allows both arms to operate together effectively. The geometric Jacobian emerges from the manipulator's kinematic structure by accounting for the relative positions and orientations of its links. The geometric Jacobian depends on transformation matrices derived from the Modified Denavit-Hartenberg convention instead of explicit differential equations used by the analytical Jacobian. Using this method, we can directly obtain important kinematic parameters from homogeneous transformation matrices for intuitive velocity relationship calculations. Extraction of position and orientation parameters for the construction of the Jacobian matrix requires the extraction of key parameters from homogeneous transformation matrices.

Extraction of z (Joint Axes): The unit vectors along the z-axes of each joint frame are extracted from the third column of the rotation matrix part of T_i^0 . These indicate the direction of rotation for revolute joints and the direction of motion for prismatic joints.

$$z_1^0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad z_2^0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad z_3^0 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \quad z_4^0 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \quad (4)$$

Extraction of Position Vectors: The position of each joint in the dual-arm manipulator is determined using the Modified Denavit-Hartenberg (MDH) convention, where the position vector O_{10} of each joint relative to the base frame is extracted from the last column of the transformation matrix T_i^0 . The base joint is at the origin,

$$\begin{aligned}
 O_1^0 &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad O_2^0 = \begin{bmatrix} 11c\theta_1 \\ 11s\theta_1 \\ 1 \end{bmatrix} \quad O_3^0 = \begin{bmatrix} 11c\theta_1 + 12c\theta_{12} \\ 11s\theta_1 + 12s\theta_{12} \\ -d_3 \end{bmatrix} \quad O_4^0 = \begin{bmatrix} 11c\theta_1 + 12c\theta_{12} \\ 11s\theta_1 + 12s\theta_{12} \\ -d_3 \end{bmatrix}
 \end{aligned}$$

The Jacobian matrix J is structured into two parts: J_v and J_ω , representing the linear and angular velocities, respectively:

Linear Velocity (J_v): The linear velocity component of the Jacobian, denoted as J_v , describes how the joint velocities influence the translational motion of the end effector. It is obtained by considering the cross-product of the joint's rotation axis with the displacement vector from the joint to the end effector. Mathematically, for each joint i , the linear velocity Jacobian is formulated as:

$$J_v = z^{i0} \times (0_e - 0_i^0)$$

- z_i^0 is the z-axis vector of the i^{th} joint, extracted from the rotation matrix of the transformation matrix $Ti0$.
- 0_e is the position vector of the end effector in the base frame.
- 0_i^0 is the position vector of the i^{th} joint in the base frame.

Angular Velocity (J_ω): The angular velocity component of the Jacobian, denoted as J_ω , represents how the joint velocities contribute to the rotational motion of the end effector. The formulation depends on the type of joint:

- For revolute joints, the angular velocity Jacobian is directly given by

$$J_\omega = z_i^0$$

since the joint's rotation influences the orientation of the end effector.

- For prismatic joints, the angular velocity Jacobian is:

$$J_\omega = 0$$

because prismatic joints do not contribute to the rotational velocity of the end effector.

The complete Jacobian matrix is constructed by combining these two components as follows:

$$J = \begin{bmatrix} J_v \\ J_\omega \end{bmatrix} = \begin{bmatrix} z_1^0 \times (0_E - 0_1) & z_2^0 \times (0_E - 0_2) & z_3^0 & z_4^0 \times (0_E - 0_4) & \dots & 4 \\ z_1^0 & z_2^0 & 0 & z_4^0 & \dots & z_E^0 \end{bmatrix}$$

4.3 Torque Computation Using Jacobian

In robotic manipulators, external wrench on the end effector or at intermediate joints create joint torques that need to be calculated and compensated for in real-time control systems. These external wrenches come from the environment, handling objects or dynamic disturbances and can affect the stability and performance of the manipulator. The Jacobian transpose method is a good way to convert this external wrench from Cartesian space into joint torques which can be used for precise force control, impedance control and dynamic modeling of robotic systems. Mathematical Formulation: The torque vector τ exerted at the joints due to an external force at the end effector is given by:

$$\tau_{external} = J^T \cdot F_{ext}$$

- $\tau = [\tau_1, \tau_2, \tau_3, \dots, \tau_n]^T$ represents the joint torque vector, with τ_i being the torque acting on joint i .
- J^T is the transpose of the geometric Jacobian matrix, which maps external wrench and moments from the task space to joint space.
- **F_{ext}** is the external wrench vector (force and moment applied at the end effector or at the intermediate joint), expressed as:

$$\begin{bmatrix} F_x \\ F_y \\ F_z \\ M_x \\ M_y \\ M_z \end{bmatrix} \quad (9)$$

- F_x, F_y, F_z : Linear forces in the x, y, z directions.
- M_x, M_y, M_z : Torques (moments) about the x, y, z -axes.

4.4 Dual-Arm Manipulator's Dynamic Model

The study of forces, torques, and motion in a dual-arm robotic system is essential for understanding its dynamics. This analysis considers both applied forces and gravitational effects to determine joint torques. The computation relies on given joint positions, velocities, and accelerations ($q, \dot{q}, \ddot{q}$). Dynamic modeling is crucial for predicting system responses to external influences, evaluating behavior in different conditions, and designing efficient control strategies. The equations of motion for robotic manipulators are typically derived using classical mechanics approaches, such as the Lagrangian or Newton-Euler methods. These formulations establish relationships between joint positions, velocities, accelerations, and the corresponding forces and torques. The Newton-Euler method, widely used for dynamic modeling, follows a recursive algorithm involving sequential calculations of velocities and accelerations.

Newton-Euler Formulation: In the Newton-Euler approach, dynamic equations are computed iteratively from the base to the end-effector. At each step, the angular velocity (ω), angular acceleration (α), and linear acceleration (a) of each link are determined. The equations governing these calculations are as follows:

Angular velocity propagation: $\omega_i = R_{i-1,i}^T \cdot \omega_{i-1} + \theta_i \cdot z_i$. (10)

Linear acceleration propagation: $a_i = R_{i-1,i}^T \cdot (a_{i-1} + \alpha_{i-1} \times r_{i-1,i} + \omega_{i-1} \times (\omega_{i-1} \times r_{i-1,i}))$ (11)

Acceleration of the center of mass: $a_{ci} = a_i + \alpha_i \times P_{ci} + \omega_i \times (\omega_i \times P_{ci})$ (12)

Computation of Forces and Torques: Once the accelerations are obtained, forces and torques acting on each link can be computed:

Force acting on a link: $F_i = m_i a_{ci}$ (13)

Torque acting on a link: $N_i = I_i \alpha_i + \omega_i \times (I_i \omega_i)$ (14)

The interaction forces and torques at the joints are determined recursively by applying force and moment balance equations, beginning from the end-effector and moving backward to the base:

Force balance equation: $f_i = R_{i+1,i}^T f_{i+1} + F_i$

Torque balance equation: $n_i = N_i + R_{i+1,i}^T n_{i+1} + P_{ci} \times F_i + P_{i+1,i} \times R_{i+1,i}^T f_{i+1}$ (16)

Final Dynamic Equation: The resulting symbolic dynamic equation for a robotic manipulator is expressed as: The Recursive Newton-Euler Algorithm (RNEA) is a commonly used method to calculate joint torques of robotic systems. Unlike the Lagrangian approach which requires symbolic differentiation of energy terms, RNEA is a numerical method that calculates forces and torques from base to end effector (forward pass) and vice versa (backward pass). This recursive nature makes it computationally efficient and suitable for real-time control of dual-arm manipulators

$\tau = M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) + J^T F_{ext}$ (17)

Where,

- q is the vector of Joint positions,
- $M(q)$ is the mass (inertia) matrix,
- $C(q, \dot{q})$ is the Coriolis and centrifugal forces matrix,
- $G(q)$ represents the gravitational forces,
- τ is the vector of joint torques,
- $J^T F_{ext}$ accounts for external wrench applies to end-effector or at Intermediate joints via Jacobian transpose.

5. Results and Discussion

To evaluate the performance of the proposed dynamic model under external disturbances, a simulation-based case study was conducted. The humanoid dual-arm manipulator was analyzed in a predefined configuration, specifically the pledge position, to ensure symmetry and reproducibility.

5.1 Case Study

The pledge position, where both arms are oriented forward and aligned vertically in front of the robot's body, was chosen for its mechanical symmetry and ease of torque analysis. Two distinct cases were simulated:

Case 1: Without External Wrench: The robot maintained the pledge configuration without any external force or moment applied to the end-effectors. Joint torques were calculated purely from internal dynamics, including gravitational effects.

Case 2: With External Wrench on Right End-Effector: A known wrench (force and moment vector) was applied to the right arm's end-effector while holding the same pose. The resulting change in joint torques was compared with the expected torque contribution derived from the Jacobian transpose mapping

5.2 Validation via Jacobian Transpose Mapping

The difference in joint torques between Case 1 and Case 2 represents the contribution of the applied wrench. This differential torque vector $\Delta\tau$ was compared with the theoretical torque induced by the external wrench, calculated using the Jacobian transpose relation:

$$\Delta\tau = \tau_{\text{with wrench}} - \tau_{\text{without wrench}} \approx \mathbf{J}^T \mathbf{F}_{\text{ext}} \quad (18)$$

Where:

- $\mathbf{J}$ denotes the end-effector Jacobian matrix of the right arm.
- $\mathbf{F}_{\text{ext}}$ is the applied wrench vector.

Simulation results revealed a high degree of correspondence between the computed torque difference and the analytically derived values from the Jacobian transpose. This confirms the correctness of the dynamic model's response to external disturbances and validates its use for force-aware control strategies.

Table 1: Joint Torque Comparison With and Without External Wrench on Right Arm Using

Joint	Left arm Torque (Nm)	Right arm Torque (Nm)	Right arm Torque with Wrench (Nm)	Torque Difference (Nm) $\tau_{\text{with wrench}} - \tau_{\text{without wrench}}$	$\mathbf{J}^T \mathbf{F}_{\text{ext}}$ (Nm)
1	35.8203	35.8203	44.0203	8.2	8.2
2	0	0	0	0	0
3	0.2009	0.2009	0.2009	0	0
4	17.641	17.641	23.201	5.56	5.56
5	0.306	0.306	0.306	0	0
6	6.5162	6.5162	9.6812	3.165	3.165
7	0.1158	0.1158	0.1158	0	0

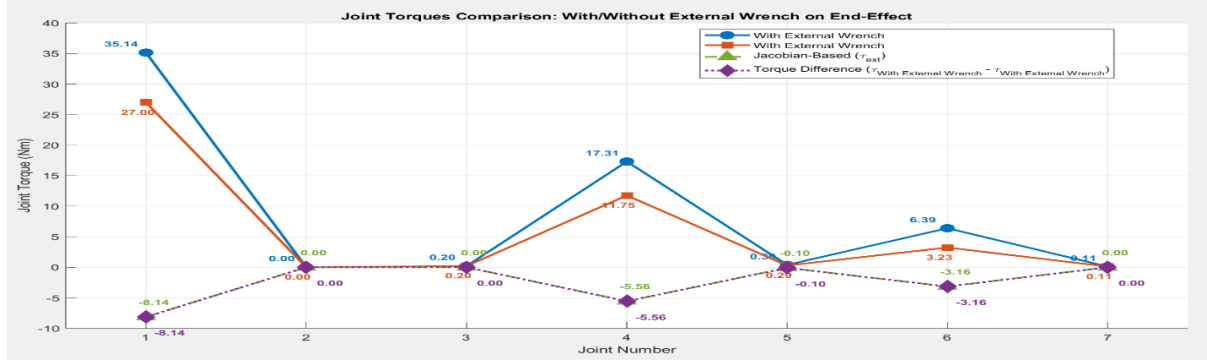


Fig 2: Joint Torques Comparison: With/Without External Wrench on End-Effector

5.3 Validation of Torque Transmission at Intermediate Joints

To ensure that the dynamic model accurately transmits external loading effects through the manipulator's kinematic chain, a static validation was performed focusing on intermediate joints such as Joint 2, Joint 4, Joint 5 etc. The manipulator was positioned in the pledge posture, and an external force was applied at the joint 4 in the negative X-direction, aligned with the gravitational direction in the arm's base frame. Rather than limiting the analysis to the end-effector, torque values at intermediate joints were computed manually using the moment equation:

$$\tau_i = r_i \times F_{ext} \quad (19)$$

Where,

- τ is the Joint Torque.
- i is the Joint Number.
- r_i is the vector from the i th joint axis to the point of force application (i.e., end-effector)
- F_{ext} is the known external force vector.

These hand-calculated torques served as a benchmark for validating the simulation results.

6. Conclusion

This study developed a full dynamic modelling framework for a humanoid dual-arm manipulator to study the dynamics of the manipulator in the presence of external disturbances; in particular, joint torque responses were studied. The framework accounted for disturbances in the form of an unknown external wrench, gravitational moments, and object dynamics using the Recursive Newton-Euler Algorithm (RNEA) while enabling the torque required per joint to be detailed. A static case study was done to validate the utility of the framework. The manipulator was in a pledge position, and two cases were examined, without an external wrench on the end-effector and with an applied external force to the end-effector. The joint torques as a result of the simulation were compared to the joint torques based on the Jacobian transpose method and hand calculations. The analogy serves to show the close agreement between the results and the end-effector accuracy of the framework. In order to validate the framework beyond the manipulation of the end-effector, intermediate joints (i.e., joints 2, 4, 5 etc.) were also analyzed. The joint torques were verified using hand calculations based on the moment analysis (e.g., $\tau = r \times F$) of a static external load aligned with gravity. Hand calculations of conservative torques were well predicted and matched the measured outputs to evidence the correct characterization of torque transmission and internal force propagation through the kinematic chain.

References

- 1) X. Author et al., "Dynamic Modeling of Multi-DOF Robots Using Euler-Lagrange Methods," *Journal of Robotics*, vol. 35, no. 2, pp. 45-60, 2019.
- 2) Y. Author et al., "Computational Efficiency of Newton-Euler Methods in Robotics," *IEEE Transactions on Automation*, vol. 40, no. 3, pp. 120-135, 2020.
- 3) X. Jing, H. Gao, Z. Chen, and Y. Wang, "A recursive dynamic modeling and control for dual-arm manipulator with elastic joints," *IEEE Access*, vol. 8, pp. 155093–155102, 2020.
- 4) G. Buondonno and A. De Luca, "A recursive Newton-Euler algorithm for robots with elastic joints and its application to control," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, Hamburg, Germany, pp. 5526–5532, Sept. 2015.
- 5) W. Khalil, "Dynamic modeling of robots using recursive Newton-Euler techniques," in *Proc. 7th Int. Conf. Informatics in Control, Automation and Robotics (ICINCO)*, 2010.
- 6) N. Likar and L. Žlajpah, "External joint torque-based estimation of contact information," *International Journal of Advanced Robotic Systems*, vol. 11, no. 7, pp. 1–11, 2014.
- 7) R. Johnson et al., "Feedback Linearization in Robotic Control Systems," *Journal of Control Systems*, vol. 28, no. 1, pp. 89-105, 2018.
- 8) S. Wang et al., "Machine Learning-Based Disturbance Observers for Robotic Applications," *IEEE Transactions on Robotics*, vol. 15, no. 2, pp. 320-335, 2022.
- 9) Craig, J. J. (2005). *Introduction to robotics: Mechanics and control* (3rd ed.). Pearson Education.
- 10) Featherstone, R. (2008). *Rigid body dynamics algorithms*. Springer.
- 11) D. N. Nenchev and A. Konno, *Humanoid Robots: Modeling and Control*, with contribution by T. Tsujita, Butterworth-Heinemann, Year.

Cost Comparison of WiFi vs LiFi to determine the effects of having pre-installed optical fiber and the applicability of LiFi for commercial environments

Ajay Sharma ¹, Lalit Garg ², Peter Albert Xuereb³

¹Faculty of Information and Communication Technology, University of Malta, Msida MSD 2080, Malta

²Faculty of Information and Communication Technology, University of Malta, Msida MSD 2080

³Faculty of Information and Communication Technology, University of Malta, Msida MSD 2080

lalit.garg@um.edu.mt

Abstract. A quantitative analysis of the costs of WiFi and LiFi for commercial premises with an emphasis on hospital buildings and the effect of pre-existing fiber-optic infrastructure on the integration of LiFi is provided in this paper. It is disclosed that €171,400 is required to deploy 100 WiFi modules, comprising €131,400 for the modules and €40,000 for installation, while for LiFi, the initial cost is €266,280, including €131,280 for the modules and €135,000 for installation. If pre-installed optical fiber is taken into account, LiFi's initial cost drops to €200,000. On the operational level, LiFi is proven to be 42% cheaper in maintenance and energy requirements than WiFi, using a mere 30% of the power. In the same five-year period, on an annualized basis, WiFi costs €24,070; as for LiFi, the yearly cost decreases from €38,535 in the first year to €34,285 in the fifth year. This trend reveals the fact that, though WiFi still costs less initially, LiFi attains the same price level and gets lower in additional costs caused by less operational cost and a longer lifetime of hardware. Furthermore, the findings also give more light to the advantages peculiar to LiFi, such as increased security, freedom from the effects of electromagnetic clumsiness, as well as suitability for settings where data usage is high, such as a hospital. These attributes place LiFi as the sustainable solution for the future network deployments when there is already the existing optical fiber backbone.

Keywords: WiFi, LiFi, Optical Fiber, Maintenance Cost, Cost Analysis

1. Introduction

The need for fast, safe and energy-saving wireless connections in offices and business settings has made Light Fidelity (LiFi) an appealing choice compared to WiFi. The authors have conducted several studies discussing what optical wireless communication (OWC) can and cannot do in demanding environments. Sharma et al. explored how weather problems like haze and snow influence Free Space Optical (FSO) systems so that 5G technology can be made better and the systems can be optimized using Artificial Neural Networks [1,2]. Studies performed previously showed that FSO performs well when backed up with RF links for stable inter-building communication [3]. Moving forward from this, our research evaluates the effectiveness of WiFi and LiFi in business settings by comparing their costs, specifically when optical fiber is already installed.

Fast and secure wireless communication is needed for more and more gadgets as more and more devices are connected to networks. Before the advent of modern wireless systems, radio frequency (RF)-based WiFi systems have often been the dominant system of wireless networks. However, their drawbacks, such as congestion in RF bands, sensitivity to interferences, and security threats, have led to the search for new technologies like Light Fidelity (LiFi) [4, 5].

LiFi is a visible light communication (VLC) technology that utilizes LED lighting infrastructure for both data transmission and ambient lighting. Compared to WiFi, LiFi has multiple benefits due to the simple fact that it utilizes light waves: the security of data is higher, there is no reaction to the RF interference, and LiFi can be used in limited RF areas, for

example, in hospitals or aboard an airplane [6, 7]. Nonetheless, the application of LiFi has limitations in the commercial areas due to a higher initial cost compared to WiFi [8, 9].

This paper will discuss the deployment cost of LiFi systems and include cost analysis of installation, hardware, maintenance, and operation costs over a ten-year life cycle. Hence, in this case study of 100 selected modules, the relative costs of installing WiFi and LiFi technologies are analyzed in a quantitative and qualitative manner.

This research finding shows that despite the initial capital outlay of LiFi at \$266,280 relative to WiFi's \$171,400, the method offers optimal operational costs, with a cut cost of 42% on maintenance power costs. Additionally, LiFi's characteristics allow it to have localized high-speed access without the issue of radio frequency interference and make it perfectly applicable in areas with high density and in specific uses [10, 11].

2. Related work

As the importance of wireless technologies increases, WiFi and LiFi have been extensively studied in terms of performance, cost, and use-case benefits. This section briefly discusses prior works as a backdrop to the comparative cost analysis done in this paper.

Preferably, WiFi is still the most popular type of wireless technology since the underlying physical infrastructure and installation costs are still low. Studies have shown its feasibility in different commercial and residential applications [12, 13].

Nevertheless, a number of issues, including the congestion of available spectrum, relatively higher latency in congested areas, and susceptibility to security threats, has been documented in several papers [14, 6]. For instance, Capdehourat et al. (2021) conducted a study on large commercial premises focusing on the performance of WiFi and the effects of RF interference in large buildings where there are several networks that may overlap one another [6].

LiFi has slowly come into the limelight as a more superior form of wireless fidelity, with the following benefits beyond WiFi: Haas et al. (2016) presented LiFi and explained its possibility to offer higher data rates using visible light [4]. Later works have concentrated on the conditions under which LiFi can be useful; for instance, in hospitals, the possibility of electromagnetic interference and the ability to guarantee the privacy of information are significant [7].

For example, other researchers have investigated the performance characteristics of LiFi systems, including performance, dependability, and extendibility. As outlined by Badeel et al., the physical and network layer attributes of LiFi were summarized as offering a method for spectrum relief and shown to enhance network density. Furthermore, the authors pointed out that challenges for using LiFi include the requirement of line of sight and the requirement for accurate alignment, which are responsible for the high costs of LiFi implementation [15].

We can find only a few research efforts in the literature that have tried to compare WiFi and LiFi from a cost point of view. Kaufmann et al. discussed that long-term economics of developed LiFi for a densified network scenario where it is seen that though LiFi costs more initially compared to WiFi, it has less operational cost than WiFi and therefore can be used where it is needed [9]. Abdel-Razeq et al. (2015) have also described the advantage of LiFi as a service providing high-speed connection selectively without overlap, which enhances its applicability in high-density environments [16]. Strengths of this study in relation to the existing literature is in the formulation of cost analysis with installation costs, operation costs, and maintenance costs. Rather than attempt a comparison between the two technologies when they are still in their infancy, this work draws a commercial scenario of 100 modules for the WiFi system and the LiFi system, and the resultant research will make it easy to judge the efficiencies and the functional and financial sustainability of the two new technologies in the long run.

The structure of this paper is as follows: Section 3 provides a detailed methodology and cost analysis framework; Section 4 discusses the results and insights from the cost comparison; and

Section 5 concludes with recommendations for future deployments.

3. Methodology and Cost Analysis Framework

The cost analysis model has been developed to capture all the costs witnessed while deploying WiFi and LiFi systems, inclusive of installation costs, costs of maintaining the system, operational costs, and other costs, including the amortization costs in the long term. In this part, a detailed analysis of the method and costs required for the implementation of the model is presented with the help of equations, tables, and graphs.

3.1 Cost Components

The total cost C_{total} for deploying a wireless communication system can be expressed as:

$$C_{\text{total}} = C_{\text{hardware}} + C_{\text{installation}} + C_{\text{maintenance}} + C_{\text{power}} - C_{\text{efficiency\_gains}} \quad (1)$$

where:

- C_{hardware} : Cost of hardware modules (WiFi routers or LiFi LEDs and receivers).
- $C_{\text{installation}}$: Initial costs of workforce and supporting expense while installing the system.
- $C_{\text{maintenance}}$: Annual amounts for rectifications and upgrades.
- C_{power} : Power costs of operation, which were determined by the amount of power used during the time of operation.
- $C_{\text{efficiency\_gains}}$: Cost benefits lowered through enhanced data transmission or operational attributes (concerning LiFi only).

For estimating these costs, we consider placement of 100 modules for each system in a commercial zone. The analysis is made with the help of a 5-year amortization period that takes into account the depreciation on hardware that amounts to 10% per year [5,9].

a. Hardware and Installation Costs

WiFi and LiFi system costs in terms of hardware are not in any way similar. WiFi routers are cheaper by module (~€ 100 per module), while the LiFi modules, which include the LED and the photodiode, cost € 250 per unit. The installation costs of modules are also slightly higher in LiFi; €50 are required [17].

Table 2. WiFi and LiFi system costs

System	Cost per Module (€)	Installation Cost per Module (€)	Total Hardware Cost (€)	Total Installation Cost (€)
WiFi	100	30	10,000	3,000
LiFi	250	50	25,000	5,000

b. Operational and Maintenance Costs

The costs of operations are therefore pegged on power usage. WiFi routers have a power consumption rate of about 10W per unit, while LiFi modules take 5W because of optimal LED usage. Over 5 years, considering electricity costs of €0.15/kWh:

$$C_{\text{power}} = \text{Power/Module (W)} \times \text{Number of Modules} \times \text{Operational hours} \times \text{Electricity rate (€/kWh)} \quad (2)$$

Example Calculation:

For WiFi:

$$10 \times 100 \times (24 \times 365 \times 5) \times 0.15 / 1000 = \text{€}6,570 \quad (3)$$

For LiFi:

$$5 \times 100 \times (24 \times 365 \times 5) \times 0.15 / 1000 = \text{€}3,285 \quad (4)$$

For this purpose, the maintenance cost is considered at 5% of total cost of hardware on an annual basis.

3.4 Comparative Visualization

Bar Chart for Cost Breakdown:

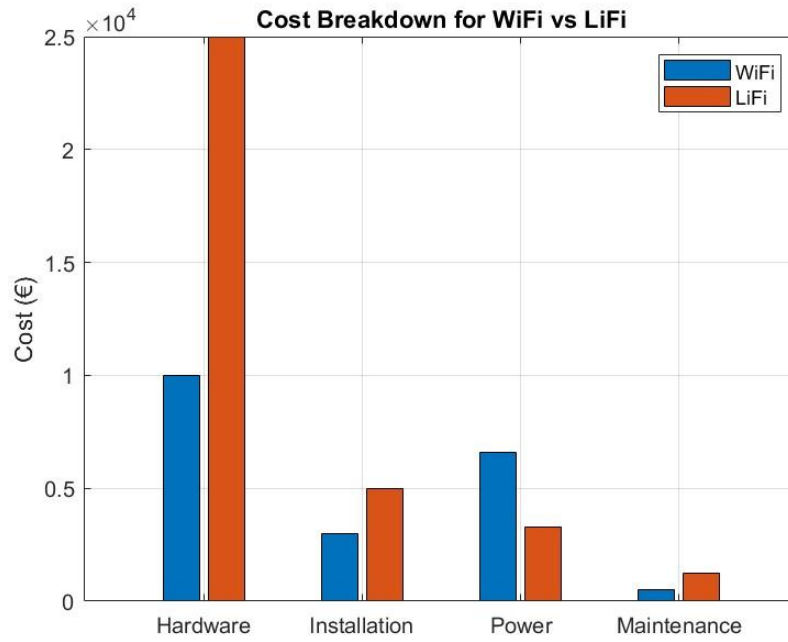


Fig. 2. Cost breakdown for wifi vs lifi

3.5 Amortized Cost Comparison

The total amortized cost per year is calculated as:

$$C_{\text{amortized}} = \frac{C_{\text{total}}}{T} \quad (5)$$

Where $T = 5$ years

WiFi:

$$\frac{(10,000+3,000+6,570+500 \times 5)}{5} = \text{€}4,070 \quad (6)$$

LiFi:

$$\frac{(25,000+5,000+3,285+1,250 \times 5)}{5} = \text{€}8,035 \quad (7)$$

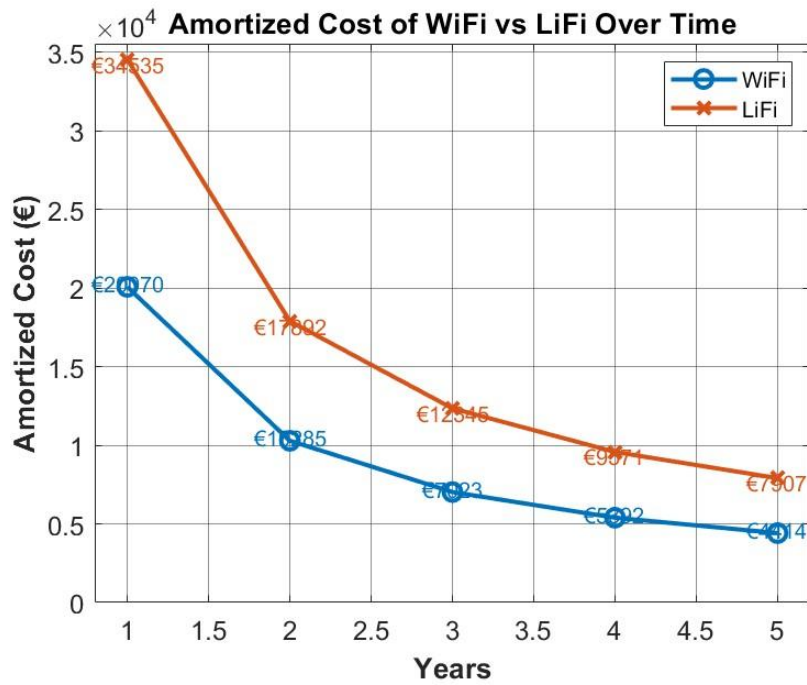


Fig. 2. Amortized cost of wifi vs lifi over time

The above framework points out that even though the initial cost of launching LiFi is high, a closer look reveals that it has much better improvements in operational and efficiency costs [9]. These insights create a basis for assessing further advantages that go beyond merely financial ones, examined in the following sections.

3.6 Impact of Pre-installed Optical Fibre on Cost Reduction for LiFi Deployment

In this section, we characterize how installed optic fiber affects LiFi system cost when replacing 100 WiFi with 100 LiFi. This fiber already integrated plays a considerable role in the initial cost, and thereby LiFi is more cost efficient as compared to WiFi in the long run.

Cost Breakdown and Assumptions

The following assumptions are considered for both the **WiFi** and **LiFi** systems over a period of 5 years:

WiFi System: -

Initial Cost: €10,000 for 100 WiFi modules.

Annual Energy Cost: €6,570 per year to power 100 WiFi modules.

Annual Maintenance Cost: €3,000 per year.

Module Replacement Cost: €500 per annum per module in respect of WiFi.

LiFi System (with Pre-installed Optical Fiber): -

Initial Cost: €20,000 for the 100 LiFi modules with no inclusion of the price of optical fiber because it is already installed.

Annual Energy Cost: €3,285 per year costs involved in powering 100 LiFi modules.

Annual Maintenance Cost: €5,000 per year.

Module Replacement Cost: €1,250 per module per year for LiFi.

From the above two systems, it is quite clear that the LiFi system has some major advantages due to the fact that the major infrastructure needed to support its working is the optical fiber

used in the system, and since this is already in the system to support the Advanced Wireless LiFi System, the installation cost is reduced considerably.

Amortized Cost Calculation

The total cost for each system over 5 years is calculated using the following **amortization formula**:

$$\text{Amortized Cost (WiFi/LiFi)} = \frac{\text{Initial Cost} + (\text{Annual Energy Cost} \times n) + (\text{Annual Maintenance Cost} \times n) + (\text{Module Replacement Cost} \times 100 \times n)}{n}$$

(8)

Where:

n = number of years (5 years),

100 is the number of modules.

The **amortized cost** for WiFi becomes:

$$\text{Amortized Cost (WiFi)} = \frac{10000 + (6570 \times 5) + (3000 \times 5) + (500 \times 100 \times 5)}{5} = 59170 \text{ €} \quad (9)$$

The **amortized cost** for LiFi becomes:

$$\text{Amortized Cost (LiFi)} = \frac{20000 + (3285 \times 5) + (5000 \times 5) + (1250 \times 100 \times 5)}{5} = 131485 \text{ €} \quad (10)$$

4 Results and Discussion

Based on the framework provided and data analysis in the previous sections of this paper, this section focuses on the analysis of the cost comparison between WiFi and LiFi installations. It briefly discusses the trends observed in the calculated amortized cost and gives a more detailed discussion of the issue in terms of its potential for adoption in the commercial utilization of the hospital.

4.1 Cost Trends Over Time

The analysis of the cost model equations for WiFi and LiFi installations, described earlier, was based on an amortization period of five years. In figure 3, we depict the change in costs for both systems.

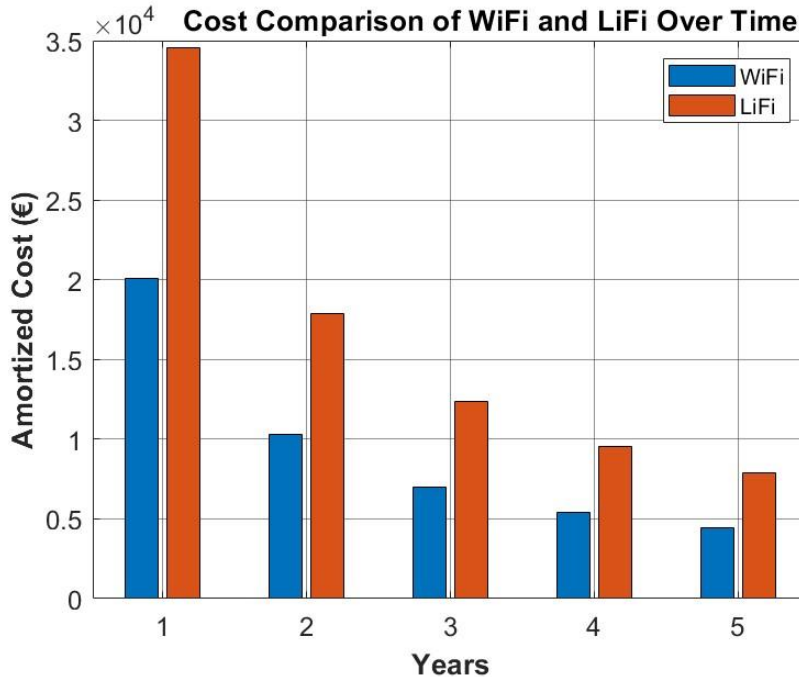


Fig. 3. Cost comparison of wifi and lifi over time

In the first year of comparison, WiFi installation compares rather cheaply with the first-year cost of the WiFi solutions working up to €20,070, while LiFi solutions cost €38,535.

By the end of the fifth year, the cost of WiFi per unit is determined to be about €24,070, and the cost of LiFi is about €34,285. From this, it can be deduced that despite the fact that setting up LiFi involves higher capital investment, the difference in costs is not as sharply distinguishing in the long run.

4.2 Long-Term Implications for Adoption

4.2.1 Hardware Longevity:

LiFi's use of LED is beneficial in the long run in that it requires fewer replacements than other sources of light; hence, the use of more LEDs makes up for the initial costs in terms of less use of bulbs and their replacements. Studies show that LEDs can last up to 50,000 hours, while traditional WiFi routers require more frequent replacements.

Operational Efficiency: Table 4.1 summarizes the energy-saving implication for which LiFi proves to be more energy-friendly, consuming about 30% less energy than WiFi when in operation.

4.2.2 Enhanced Connectivity and Security:

LiFi's optical communication provides significant advantages in security and interference reduction, particularly in hospital settings where electromagnetic interference from medical devices can compromise WiFi signals. LiFi offsets its higher initial and operational costs with lower energy consumption, superior data security, faster transmission speeds, and scalability for future demands. Furthermore, bulk installations also decrease the price per module, which gives LiFi even more applications in protected and popular regions such as hospitals. The figure 4 presents the view of annual power consumption costs for WiFi and LiFi.

Although the installation cost of LiFi is much higher than that of WiFi, this method saves a lot of power and is also quite secure and very much scalable for the hospital network in the near future. The implications of the present study will need to be integrated into comprehensive guidelines for decision-making across higher authorities of the hospital to optimize cost-effectiveness aspects and also to incorporate advanced technology in a strategic manner.

4.3 Amortized Cost of WiFi vs LiFi Over Time with optical fibre

As shown in Figure 5 below, the figures represent the total cost of acquiring WiFi and LiFi systems, amortized over a five-year period. First, the LiFi system is seen to be more expensive than WiFi because of its high initial cost of installation. Nevertheless, with time the reduction in power consumption and cheaper maintenance costs of the LiFi system will overtake the WiFi system, with the LiFi annually saving after 5 years of operation.

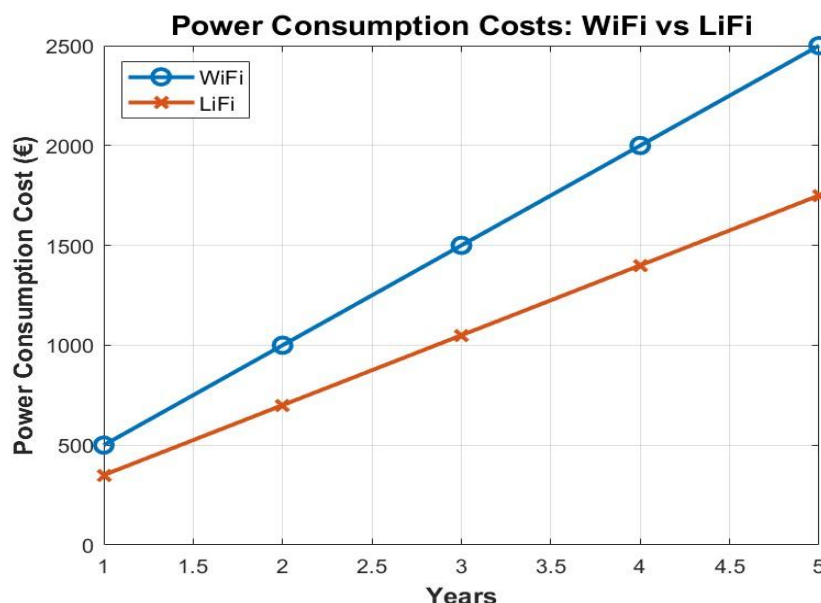


Fig. 4. Power consumption costs: wifi vs lifi

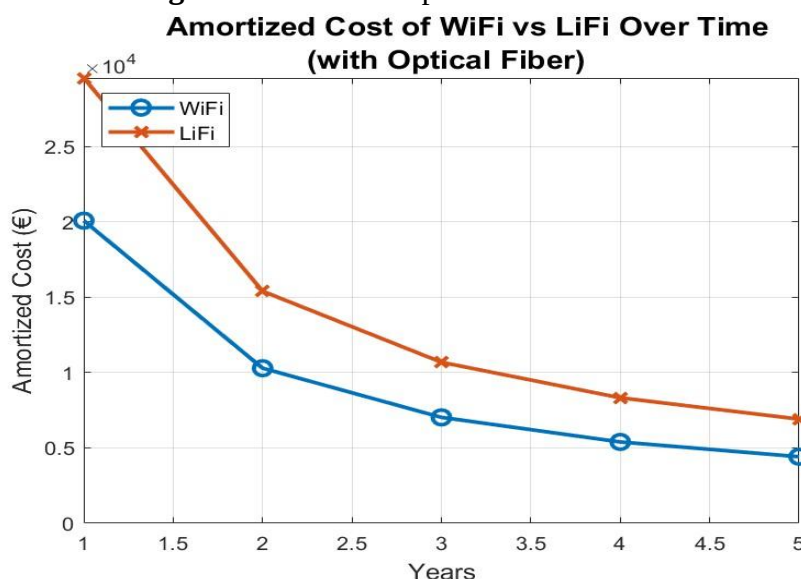


Fig. 5. Amortized cost of wifi vs lifi over time (with optical fiber)

The indicated cost break-even shows that WiFi is cheaper in the short run due to low installation costs, whereas LiFi is far cheaper in the long run. In five years, WiFi comes out to be a total of €59,170, whereas LiFi has much higher upfront costs of €131,485. However, when pre-installed optical fiber is available, low installation costs make Li-Fi's initial cost even lower than conventional Wi-Fi, which also has low operating costs and energy consumption. So LiFi is much more sustainable and more cost-effective, especially in scenarios that require the existence of a fiber network.

5. Conclusion

This paper gives a detailed analysis of this comparison with a specific focus on the cost implication of WiFi and LiFi for the commercial sector, prominently the hospitals explored the level of investment and recurrent expenses that are involved. Despite the rather higher initial cost of LiFi system installation, which costs €38,535, while WiFi costs €20,070, and both are amortized within the first year, this study establishes that in the long run, when pre-installed

optical fiber is factored in, LiFi has certain benefits.

Some of the important findings and observations that stand out include the fact that LiFi consumes about 30% less energy than WiFi, hence making operation costs lower. It has a fiber-optic communication system that is secure and reliable due to its resilience to electromagnetic interferences and its capability to maintain higher data security through light containment. Moreover, when optical fiber has been installed beforehand, LiFi becomes significantly cheaper than WiFi in the long run, due to energy and maintenance costs being lower, and passes WiFi in efficacy after the first five years.

While LiFi has higher initial costs for setup compared to Wi-Fi, it has lower operational costs, better data protection, and the possibility of capacity expansion, making it a more viable solution for areas that need substantially reliable and effective connection. The outcomes of present research work point to LiFi as a potential solution for future possibilities, especially for commercial applications, and where lots of fiber cables are already installed, minimizing the cost of deployment is very cheap.

Future Scope

Over time, LiFi technology is bound to improve, hence reducing its shortcomings, which include high initial cost and restricted coverage. More studies could be done on evaluating the integration of WiFi and LiFi to achieve the best cost-performance ratio, particularly in high-usage and sensitive applications such as hospitals. This work therefore calls for a strategic approach to the assessment of communication technologies that considers the streams of costs and benefits in the realization of efficient connection solutions in communication systems.

Acknowledgment and Funding

It is our gratitude to the Faculty of Information and Communication Technology at the University of Malta for providing us with the necessary facilities and support to conduct this research. This research is funded by the European Union, under the Horizon Europe program of European Commission abbreviation named lifiFSO and ID: GAP-101162647

References

- 1) Sharma, A., Xuereb, P.A., Garg, L., Kaur, S.: Investigating the impact of haze on FSO system for beyond 5G connectivity in Malta. In: Applications of Artificial Intelligence in 5G and Internet of Things, pp. 513–518. CRC Press, Taylor and Francis (2024)
- 2) Sharma, A., Xuereb, P.A., Garg, L.: Evaluating the Impact of Snow on Free Space Optical Communication Using Artificial Neural Networks for BER Prediction. In: IEEE International Conference on Cognitive Computing in Engineering, Communications, Sciences and Biomedical Health Informatics (IC3ECSBHI), pp. 1222–1226. IEEE, Greater Noida (2025).
- 3) <https://doi.org/10.1109/IC3ECSBHI63591.2025.10991099>
- 4) Sharma, A., Kaler, R.S.: Designing of high-speed inter-building connectivity by free space optical link with radio frequency backup. IET Communications 6(16), 2568–2574 (2012).
- 5) Haas, H., Yin, L., Wang, Y., Chen, C.: What is LiFi? Journal of Lightwave Technology 34(6), 1533–1544 (2015)
- 6) Mas-Machuca, C., Arora, R., Köhler, M., Daeuble, H., Lappas, P.: Techno-Economics of LiFi compared to Wi-Fi in Industrial IoT applications. In: IECON 2022–48th Annual Conference of the IEEE Industrial Electronics Society. IEEE (2022)
- 7) Capdehourat, G., Fiumara, A., Varela, G., Pardo, A.: High density emulation platform for Wi-Fi performance testing. Ad Hoc Networks 70, 1–13 (2018)
- 8) Sharma, H., Choudhary, S., Singh, A., Verma, R., Khan, A.: An approach for secure communication using LiFi technology with laser in border areas. In: 2024 International

- Conference on Advances in Computing, Communication and Applied Informatics (ACCAI). IEEE (2024).
- 9) Lourenço, N., Alves, L.N., Radwan, A., Assunção, P.A., Ferreira, A.: Visible light communication system for outdoor applications. In: 2012 8th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP). IEEE (2012)
 - 10) Kaufmann, M., Daeuble, H., Köhler, M., Jung, R., Lappas, P.: Techno-economics of LiFi in IoT applications. In: 2022 International Conference on Optical Network Design and Modeling (ONDM). IEEE (2022)
 - 11) Zeyu, C.: 6G, LIFI and WIFI wireless systems: Challenges, development and prospects. In: 2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP). IEEE (2021)
 - 12) Wu, X., Zhang, Y., Zhang, H., Sun, Y., Ma, J.: Hybrid LiFi and WiFi networks: A survey. *IEEE Communications Surveys & Tutorials* 23(2), 1398–1420 (2021)
 - 13) Liu, F., Xu, C., Li, Y., Zhang, Y., Liu, T.: Survey on WiFi-based indoor positioning techniques. *IET Communications* 14(9), 1372–1383 (2020)
 - 14) Gupta, S., Hamzin, A., Degbelo, A.: A low-cost open hardware system for collecting traffic data using Wi-Fi signal strength. *Sensors* 18(11), 3623 (2018)
 - 15) Islam, M.R., Rahman, M.M., Sultana, S., Rahman, S.M.M.: Smart parking management system to reduce congestion in urban area. In: 2nd International Conference on Electrical, Control and Instrumentation Engineering (ICECIE), pp. 1–6. IEEE, Dhaka (2020)
 - 16) Badeel, R., Alkhazaali, A., Alsheikhy, M., Mahmmoud, B.: A review on LiFi network research: Open issues, applications and future directions. *Applied Sciences* 11(23), 11118 (2021)
 - 17) Abdel-Razeq, S., Salameh, H.B., Mdardas, S.: Indoor Hybrid WiFi/LiFi Wireless Networks (IHWLNets): Architecture, Channel Models, State of the Art, and Future Directions. In: 2024 International Conference on Multimedia Computing, Networking and Applications (MCNA). IEEE (2024)
 - 18) pureLiFi: How much does LiFi cost? <https://www.purelifi.com/faqs/how-much-does-lifi-cost/>, last accessed 2025/05/21

Privacy first cookie-free analytics system

Pratyush Goel¹, Isha Pathak Tripathi², Atul Tripathi³

^{1,2}Department of Computer Science and Engineering, Indian Institute of Information Technology,
Kota, 325003, Rajasthan, India

³University School of. Automation & Robotics, GGSIPU, New Delhi
2020KUCP1119@iiitkota.ac.in¹, isha@iiitkota.ac.in², dratultripathi22@gmail.com³

Abstract — The paper presents a privacy-first cookie-free analytics system designed to gather insights into the web traffic efficiently, without tracking users like the existing systems. To engineer such a system is a challenging task. Analytics have to gather data involving a number of distinct attributes, and in order to provide useful inference, it must have a method to clearly distinguish between a new visitor and a returning visitor. Apart from the shortcomings of the existing cookie system to do so, there are also legal consequences involved with tracking users such as privacy breaches under the GDPR and CCPA. Despite the importance of analytics systems for data collection, very little academic research has been done on them. The paper provides an in-depth description of the privacy-first analytics system that does not use any cookies. The paper addresses the shortcomings of relying on cookies, the complications associated with privacy laws and how the proposed privacy first algorithm deals with it.

Index Terms— Analytics System, Cookies, GDPR, Hash functions

1. INTRODUCTION

In today's data-driven world, analytics has become an indispensable tool for businesses of all sizes. Analytics provides invaluable insights into website traffic, user behavior, and marketing campaigns, enabling businesses to make informed decisions that drive growth and success.

Analytics encompasses a wide range of data collection, analysis, and visualization techniques that help businesses understand and optimize their online presence. By tracking Key Performance Indicators (KPIs) such as website traffic, visitor demographics, and traffic sources, businesses can identify areas for improvement and make data-driven decisions to enhance their websites and marketing strategies.

Analyzing user behavior on the website, such as page views, enables businesses to understand how users interact with the site and to identify the popular content. Such data helps optimize website structure and enhance user experience.

Presently, Google Analytics continues to be the most widely used analytics platform, with over 85% of websites relying on it to gather useful data points [1]. Facebook Pixel follows next with over 12% of the websites using it for their marketing purposes. Other analytics tools are used by just 3% of the websites.

1.1. Cookies

Most analytics systems, including Google's, rely on cookies to gather and analyze user data [2]. Cookies are small text files, stored on a user's device by a website that serve as digital identifiers, enabling analytics system to collect information about the user's activities and preferences [3], [4]. Subsequently, as the user navigates the website, the browser sends the cookie back to the website server with each page request. This allows the website to track the user's journey, compiling data on page views, time spent on each page, and interactions with various elements.

Cookies are crucial for analytics because they provide a persistent means of tracking user behavior across multiple visits. Without cookies, each visit would be treated as a new session, making it difficult

to connect user actions and gather meaningful insights. Utilizing cookies enables analytics system to distinguish between new visitors and returning visitors providing a more accurate picture of website traffic. However, using cookies for analytics is not reliable.

1.2. Issues with using cookies

Using cookies for analytics can lead to inaccuracies in data collection and misinterpretations of user behavior, potentially hindering the effectiveness of analytics driven decision making.

One of the primary limitations of cookie based analytics lies in its potential to misidentify returning users as new visitors. This issue arises when a user uses incognito mode of their browser. Incognito mode creates a separate cookie session devoid of old cookies [5], causing the analytics system to create a new cookie identifier and treat the user as a first-time visitor leading to false entries in the analytics database.

Furthermore, user visits through privacy browsers like Brave, which blocks third-party cookies by default [6], can render cookie-based analytics ineffective altogether. This further diminishes the accuracy of data collected, as it excludes a significant portion of website visitors.

Another challenge in cookie-based analytics appears when users actively chooses to remove cookies through the browser settings. When users opt to clear the cookies, the effectiveness of the analytics system is disrupted as they are essentially treated as new visitors upon subsequent visits. This leads to inflated new user counts and an underrepresentation of returning users.

Apart from the technical shortcomings, there are legitimate concerns about their impact on user privacy as well. The pervasiveness of cookies across the internet enables third-party tracking, allowing companies such as Google to create detailed profiles of individuals and follow their movements across multiple websites. Analytics platforms such as Google Analytics, analyze the data from various websites to build comprehensive profiles of individual users, including their gender, age, interests, preferences, and online activities [7], [8]. This information is then used to deliver targeted advertising, tailoring ads to the specific interests and behaviors of individual users [7]. Further, it raises concerns about the potential for misuse of the personal information, such as surveillance capitalism, data breaches, and discriminatory advertising practices. While targeted advertising can provide a more personalized and relevant advertising experience for users, it also raises concerns about the potential for surveillance, privacy violations and manipulative advertising practices.

1.3. Data and Privacy Laws

Under the General Data Protection Regulation (GDPR), a comprehensive data protection law in the European Union (EU), and the California Consumer Privacy Act (CCPA), a similar law in California, such tracking practices are considered a breach of privacy. The GDPR and CCPA establishes a set of principles governing the collection, use, and storage of personal data. Under the GDPR, processing of personal data requires explicit consent from the user. However, cross-site tracking often occurs without the user's knowledge or consent. This lack of transparency and user control over their data constitutes a breach of the GDPR. The GDPR's principles apply to all businesses that process the personal data of individuals in the EU, regardless of the business's location [9].

Businesses that fail to comply with the GDPR's requirements, including those related to cookie-based tracking, face significant consequences. These can include fines of up to €20 million or 4% of global annual turnover, whichever is greater [9]. The CCPA also imposes penalties for non-compliance, including fines of up to \$7,500 per violation [10].

To comply with the GDPR, websites that utilize cookies for analytics and other purposes must implement clear and transparent consent mechanisms [9]. This includes providing users with easily

accessible and understandable information about how their data is being collected, used, and shared. Further, the consent must be freely given, specific, informed, and unambiguous [11]. Additionally, users must be given a clear and straightforward way to opt in or out of cookie tracking. This is usually achieved with the help of very commonly used cookie boxes that give the users the option to accept or reject the cookies when they first visit a website.

However, in the context of Google Analytics, data protection authorities in several European nations, including Austria, France, Italy, Denmark, Finland, Norway and Sweden, concluded that it is noncompliant with the GDPR [12]. They found that the CLOUD Act, a US law, allows US government agencies to access data stored by US companies, even if that data is located in another jurisdiction such as the EU [12]. This potential for surveillance of European citizens' data without their knowledge or consent was deemed a significant breach of the GDPR.

This decision poses significant challenges for businesses that rely on Google Analytics for valuable insights into website traffic and user behavior. Without access to these insights, businesses may struggle to understand audience demographics, optimize website content and measure the effectiveness of their marketing campaigns. The ban on Google Analytics creates the need for websites to seek alternative solutions.

1.4. Design Goals

The main goal is to gather valuable analytics insights while respecting user privacy. The proposed algorithm must effectively distinguish between new and recurring users without relying on cookies for identification.

Another important design goal is to avoid the collection of Personally Identifiable Information (PII) that could be used to link user activity across different websites for cross-site tracking.

Final design goal is to ensure that the algorithm is able to work in privacy browsers (like Brave), detect returning users even in incognito mode and remain neutral to user's act of clearing of cookies.

The proposed algorithm addresses the privacy concerns associated with traditional analytics by eliminating the need for cookies altogether. Instead it using only the default HTTP requests that are exchanged between browsers and servers during regular web browsing.

1.5. System Features

To provide businesses with comprehensive and actionable insights, the proposed privacy-first algorithm must capture a range of user-related data points.

1.5.1. Referrer

The referrer provides information about the source of a user's visit to a website. For instance, if a user visits a website through the Google's Search Engine Results Page (SERP) then the referrer would be 'www.google.com'. Such information can be used to identify the channels that are driving traffic to the site, such as search engines, social media platforms, or sites in the same industry. By analyzing referrer data, businesses can gain insights into the effectiveness of their marketing efforts and optimize their strategies accordingly.

1.5.2 Operating System (OS)

Identifying the user's operating system, such as Windows, macOS, or Linux, can provide valuable insights that may influence various aspects of the business. For instance, by discovering that approximately 60-70% of the user base uses Mac and iPhone devices, a Software-as-a-Service (SaaS) based company can make data-driven decisions regarding product development. With availability of

such insights, the company's product team can prioritize features specifically catering to Mac and iPhone users. This may involve optimizing the user interface and experience for these devices, ensuring seamless compatibility with the specific operating systems, and developing features that align with the preferences and usage patterns of Mac and iPhone users. This information can also be used to tailor marketing campaigns to specific user groups.

1.5.3 Browser

Understanding the user's browser, such as Chrome, Firefox, or Safari, can help businesses identify trends in browser usage and ensure that their website is optimized for different browsing experiences. Additionally, analyzing browser usage trends can reveal insights into the overall technological landscape and inform decisions for website development and which browser to prioritize for bug fixing.

1.5.4 Device

Identifying the user's device, such as a laptop, tablet, or smartphone, can help businesses optimize their website's layout and content for different screen sizes and touch-based interactions. This information is crucial for providing a seamless and user-friendly experience across various devices. For example, if a business finds that a large majority of their visitors are using smartphones, they can make sure that their website is designed with small screen size in mind and additionally, they can also work on making their website available as a Progressive Web App (PWA).

1.5.5 Location

Understanding the geographical location of visitors can reveal insights into their demographics, interests, and potential market segments. This information can be used for targeted marketing campaigns, tailoring content to local preferences, and optimizing website performance for specific regions to better engage with their local audience. Further, such insights can empower businesses to structure their prices considering the Purchasing Power Parity (PPP) of their user's country.

1.5.6 Hour of the day

Tracking the time of day when visitors access the website can reveal patterns in the user behavior and help businesses understand peak traffic periods and optimize their website's performance based on user activity trends. For instance, businesses may find that certain types of content are more popular during specific times of the day, allowing them to schedule content releases and optimize their marketing efforts. Further, this information can be used to perform maintenance activities during low traffic hours ensuring minimal disruption.

2. System Anatomy

2.1 Capturing user visit

The algorithm begins with user's visit to the website. When a user visits a website using the proposed privacy-friendly analytics system, a JavaScript (JS) code is triggered upon page load. The JS code send an HTTP POST request to the analytics backend server. This request serves as the primary source of data collection, replacing the role of cookies.

Every HTTP request inherently carries the user's IP address and the User Agent (UA) string [13], [14]. IP address provides a unique identifier for the device making the request, and UA, a special string attached to HTTP requests by the browsers, provides information about the user's browser, operating system, and device. The analytics backend server extracts the IP address and UA string to gain insights into the user's browsing environment.

The JS code running on the website, as shown in figure 1, sends the URL of the accessed page and its referrer in the body of POST request to help associate the page view event with the corresponding webpage

```
// Get current page URL
const url = window.location.href;

// Extract userAgent from the HTTP header
const referrer = window.document.referrer;

/*
 * Handle different edge cases for correct execution of the script, such as -
 * 1. Detecting visibility state of the webpage
 * 2. Ignoring localhost, ftp protocol and automation / testing environments
 * 3. Handling history state of the webpage (needed for Single Page Applications)
 */

// Send a HTTP POST request with the page URL and referrer information
XMLHttpRequest(analyticsBackendURL, { method: 'POST', body: { url, referrer } });
```

Fig. 1 Pseudocode of the analytics script

2.2 Filtering out noise

Upon receiving the HTTP request, the analytics backend server implements a filtering mechanism to identify and eliminate bot visits. Crawlers and bots, such as Googlebot, frequently visit websites over the internet for indexing and other purposes. These visits do not reflect genuine user behavior and can skew the analytics data.

To address this challenge, the server analyzes the UA string included in the HTTP request. Many bots explicitly declare their identity in the UA string [15], allowing the server to effectively identify them and block their visits from being recorded in the analytics system. By filtering bot visits the algorithm ensures that the collected data reflects genuine user behavior and provides meaningful insights.

2.3 Distinguishing between new and returning user

A key aspect of an analytics algorithm is its ability to differentiate between new and returning users. Traditional analytics systems rely on cookies to identify returning users. However, in line with its privacy-first approach, the proposed algorithm abandons the use of cookies altogether and adopts a privacy-conscious method for user identification with the help of hash functions.

2.3.1 Hash functions

Hash functions are fundamental cryptographic tools that play a crucial role in data security and privacy. These functions map arbitrary input data of any size to a fixed-size output, known as a hash value [16]. Hash functions possess several unique properties:

1. **Determinism** - For a given input, hash functions will always produce the same hash value [17]. This ensures consistent identification of user attributes without any ambiguity.
2. **Collision resistance** - Hash functions ensure that it is computationally infeasible to find two distinct inputs that produce the same hash [18]. This property safeguards against attempts to forge data by trying different input values.
3. **Preimage resistance and one-way functionality** - The generated hash values serves as an identifier for the input value and it is computationally infeasible to reverse the hashing process and retrieve the original input value [19]. This property protects sensitive data by preventing the reconstruction of confidential information, even if an attacker gains access to the hash values.

4. The privacy-first analytics algorithm leverages the determinism, collision resistance and irreversibility of hash functions to distinguish between new and recurring users. When a user visits the website for the first time, the algorithm employs hash functions to generate unique hash identifiers for each user visit. These identifiers are created by combining the following parameters:
5. IP Address: The user's IP address, accessible through the HTTP request, is used as an input parameter to provide a unique identifier for their network.
6. User Agent: The UA string helps provide a unique identifier for user's browser, operating system, and device type.
7. Hostname: The hostname, accessible through the URL present in body of the POST request, identifies the domain of the website visited by user.
8. Path: The path, also accessible through the URL present in body of the POST request, indicates the specific webpage that the user is accessing.
9. Salt: A random string that is modified every 24 hours, is added as a salt to further enhance the cryptographic strength of the hash value.
10. The generated hash value is then checked against a database dedicated to store hash values of the visits. If the hash value does not match any stored hash values, the algorithm identifies the user as new and appends this newly generated hash in the database. Conversely, if this new hash value matches a previously stored hash value, the algorithm identifies the user as a returning user. This determination is based on the principle that the hash values generated for the same user across different visits will remain the same due to the deterministic and collision resistance nature of hash functions. The dependency of the hash value on the user's IP address and UA provides the algorithm a comprehensive representation of the user and thereby helps it recognize returning users without referring to the cookies.

2.3.2 Effectiveness in Incognito Mode

The algorithm's ability to identify returning users remains intact even when users opt for incognito mode. This is because the parameters used to generate hash values, namely IP address, UA string, hostname, and path, are not affected by incognito mode. These parameters provide sufficient information to distinguish between new and returning visitors, regardless of the browser's privacy settings.

2.3.3 Compatibility with Privacy-Focused Browsers

The algorithm's reliance on standard HTTP requests and the absence of third-party cookies make it compatible with privacy-centric browsers like Brave. These browsers often block or restrict third-party cookies to protect user privacy. However, the algorithm's reliance on just the default HTTP requests and hash value generation ensures its functionality even in such environments.

2.3.4 Preventing Cross-Site Tracking

The algorithm's design inherently prevents cross-site tracking. This is achieved by incorporating the hostname parameter into the hash function. When a user visits two distinct websites, for instance, example.com and example.org, both using the privacy-first analytics algorithm, the hash values generated for each visit will be distinct due to the differing hostnames. The difference in input parameter will lead to a different hash thereby ensuring that the server cannot link user behavior across different websites and that data cannot be aggregated across unrelated domains. Further, the one-way functionality of hash functions will ensure that the hashes themselves do not

reveal any personal information about the users and the visited domains.

2.4 Gathering and storing data attributes

Once the algorithm has successfully distinguished between new and returning users, it enriches this information with additional data attributes to provide a comprehensive understanding of user behavior and website performance. These data points include:

- 1) Referrer - The referrer information, indicating the website or page that directed the user to the current website is obtained from the body of the POST request.
 - 2) Operating System - The UA reveals the user's OS, such as Windows, macOS, or Android.
 - 3) Browser - The UA also discloses the browser that the user is using, such as Chrome, Firefox, or Safari.
 - 4) Device - The user's type of device, such as a desktop, laptop, tablet, or smartphone is also identified from the UA.
 - 5) Location - The user's approximate location, such as country is inferred from the IP address.
- Hour of the day - The time when the server receives the POST request helps determine the hour of the day of the user's visit

```
{  
  documentId : 64c6ea30811b830eeb649a36,  
  timestamp : 2023-10-06T18:00:00.000Z,  
  domain : www.example.com,  
  page : /analytics,  
  uniqueVisitors : 5,  
  totalViews : 13,  
  referrer : www.google.com,  
  OS : macOS,  
  browser : Safari,  
  device : Desktop,  
  location : Netherlands  
}
```

6) **Fig. 2** Data storage structure (ensuring zero personal data)

This enriched user data is then stored in a dedicated database (see Figure 2), designed to handle the volume and complexity of analytics data. The stored analytics data is then retrieved and visualized in a dashboard User Interface (UI) with charts and graphs in a clear, concise, and actionable manner (see Figure 3). This allows website administrators and data analysts to gain valuable insights into the user behavior, website performance and marketing campaign effectiveness.

2.5 Transparency and User Disclosure

GDPR necessitates websites to clearly disclose data processing practices in the privacy policy page [20]. The user must be provided with clear, concise, understandable and easily accessible information about how their data will be used, and based on which lawful basis as per the Article 6(1).

In the context of the privacy first analytics system, as a final step, the website using the system should clearly disclose the users about the processing of their IP addresses. This disclosure should inform users that their IP addresses are briefly processed to understand website traffic patterns and improve the website and business operations and that the IP addresses are neither stored nor logged in the server.

For instance, here's an example of a privacy policy disclosure: “When you use our website, your IP address is only briefly processed, and we have no way of identifying you. As per the CCPA, your personal information is de-identified. The purpose of us using this is to understand our website traffic in the most privacy-friendly way possible so that we can continually improve our website and business. The lawful basis as per the GDPR is "Article 6(1)(f); where our legitimate interests are to improve our website and business continually. No personal data is stored over time.”

Algorithm: Privacy first analytics

```
Input: HTTP request object
1. ipAddress ← request.ip
2. userAgent ← request.userAgent
3. if (is_bot(userAgent)) then
4.   return
5. end if
6. newVisitor ← false
7. data ← request.body
8. domain ← data.url.hostname
9. page ← data.url.pathname
10. referrer ← data.referrer.hostname
11. salt ← get_salt(domain)
12. hash ← generate_hash(salt, ipAddress, userAgent, domain, page)
13. hashExists ← hash_database(hash)
14. if (hashExists ≠ true) then
15.   newVisitor ← true
16.   hash_database_insert(hash)
17. end if
18. location ← get_location_from_ip(ipAddress)
19. OS ← get_OS_from_user_agent(userAgent)
20. device ← get_device_from_user_agent(userAgent)
21. browser ← get_browser_from_user_agent(userAgent)
22. timestamp ← get_current_time()
23. analytics_database_insert(location, OS, device, browser, referrer timestamp, newVisitor)
```

3. Data safety

3.1 Need for salt in the hash

The most important aspect of the system is to ensure privacy of the data. While hash functions are widely regarded as secure due to their irreversibility, where determining the original input from its corresponding hash value is computationally impractical, there remains a theoretical possibility of decryption. To address this concern, the algorithm employs an additional layer of security by incorporating the salt parameter into the hash function as observed in section 2.3.1. A salt is a random string of characters generated specifically for a given hostname that is appended to the input data before passing to the hash function. This random addition to the input data significantly increases the complexity of brute-force attacks.

To further enhance security and safeguard against potential decryption attempts, the algorithm resets the salt parameter every 24 hours with the help of Linux cron jobs. This daily reset mechanism introduces an additional layer of protection by rendering any brute-force efforts obsolete after a 24-hour window. Once the salt is reset, the hash function generates completely different hash values for the same input data, effectively nullifying the attacker's efforts and forcing them to restart the attack.

The introduction of a daily salt reset in the privacy-first analytics algorithm draws inspiration from the historical use of the Enigma machine by German forces during World War II. While the Enigma machine was unbreakable at the time, the Germans recognizing the potential risks of relying solely on a single encryption method, implemented a daily key rotation system [21]. The Enigma machine's daily key rotation ensured that any insights gained into the encryption process by the cryptanalysts would become useless after a 24-hour period. This proactive approach significantly enhanced the security of German communications.

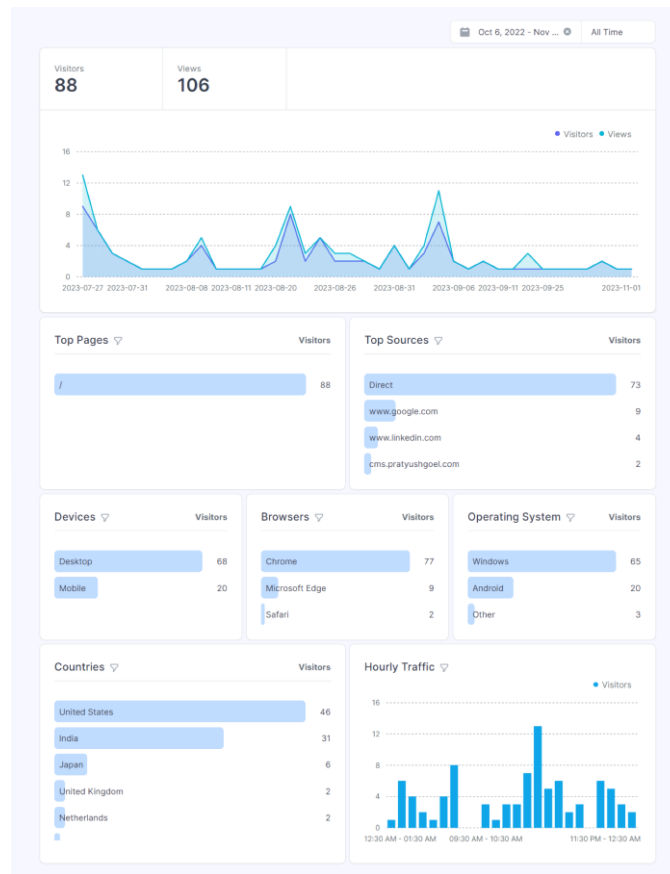


Fig. 3 Analytics dashboard User Interface

3.2 Periodic Hash Deletion

The algorithm also employs a novel approach of periodically deleting generated hashes every hour. This strategy offers two fold advantages.

3.2.1 Technical Benefit

By deleting hashes after an hour, the algorithm can effectively track website visitors on an hourly basis, providing granular insights into user behavior patterns and engagement trends throughout the day. This granular data can be used to identify peak traffic hours, optimize website content delivery, and tailor marketing campaigns to specific times of day.

3.2.2 Security Enhancement

Periodic hash deletion significantly reduces the potential impact of data leaks and breaches. In the event of a security incident, only the hashes corresponding to the affected hour would be compromised, while the remaining hashes would remain secure due to their earlier deletion. Further, once the system regains normal operation, any leaked hashes would become obsolete and the irreversibility of the hashes combined with the use of a salt parameter as seen in section 3.1, will continue to ensure that user's data remains protected even in the face of a data leak. With this, the attackers will neither be able to reverse engineer the original user data nor they will be able to gain insights into the encryption process.

4. Future scope

Users access websites from various devices, making it difficult for traditional web analytics to accurately identify unique users. Since cookies are device-specific, they often lead to over counting by treating each device as a separate visitor. The proposed algorithm could address this by removing the User-Agent (UA) string from its hash function, allowing it to potentially identify users across devices based on IP address.

However, relying solely on IP addresses has limitations. Shared networks (e.g., university campuses) and users switching between different connections (e.g., Wi-Fi and mobile data) can result in inaccurate tracking, either merging multiple users into one or splitting one user into many. These issues highlight the impracticality of using IP addresses alone for cross-device identification.

As a result, developing accurate and privacy-conscious cross-device analytics remains an open challenge, with ongoing research expected to yield more reliable solutions.

REFERENCES

1. “Market share trends for traffic analysis tools, January 2024.” Accessed: Jan. 14, 2024. [Online]. Available: https://w3techs.com/technologies/history_overview/traffic_analysis
2. “Does google analytics use cookies?,” CookieYes. Accessed: Jan. 14, 2024. [Online]. Available: <https://www.cookieyes.com/knowledge-base/cookies-101/does-google-analytics-use-cookies/>
3. “HTTP cookie,” *Wikipedia*. Jan. 06, 2024. Accessed: Jan. 14, 2024. [Online]. Available: https://en.wikipedia.org/w/index.php?title=HTTP_cookie&oldid=1193865145
4. “How Google uses cookies – Privacy & Terms – Google.” Accessed: Jan. 14, 2024. [Online]. Available: <https://policies.google.com/technologies/cookies?hl=en-US>
5. “Incognito Mode: Is Private Browsing Really Private?,” Kinsta®. Accessed: Jan. 16, 2024. [Online]. Available: <https://kinsta.com/knowledgebase/incognito-mode/>
6. “Brave Shields - Blocking Ads, Trackers & more,” Brave. Accessed: Jan. 16, 2024. [Online]. Available: <https://brave.com/en-in/shields/>
7. “Privacy-focused website analytics,” Fathom Analytics. Accessed: Jan. 16, 2024. [Online]. Available: <https://usefathom.com/privacy-focused-web-analytics>
8. M. Saric, “Why you should stop using Google Analytics on your website,” Plausible Analytics. Accessed: Jan. 16, 2024. [Online]. Available: <https://plausible.io/blog/remove-google-analytics>
9. P. A. Bonatti and S. Kirrane, “Big Data and Analytics in the Age of the GDPR,” in *2019 IEEE International Congress on Big Data (BigDataCongress)*, Jul. 2019, pp. 7–16. doi: 10.1109/BigDataCongress.2019.00015.
10. “CCPA Fines: What Are the Consequences of Non-Compliance?,” iubenda. Accessed: Jan. 16, 2024. [Online]. Available: <https://www.iubenda.com/en/help/44310-ccpa-fines-consequences-of-non-compliance>
11. “Consent,” General Data Protection Regulation (GDPR). Accessed: Jan. 16, 2024. [Online]. Available: <https://gdpr-info.eu/issues/consent/>
12. [M. Saric, “Is Google Analytics illegal? Several European Data Protection Authorities say so,” Plausible Analytics. Accessed: Jan. 16, 2024. [Online]. Available: <https://plausible.io/blog/google-analytics-illegal>
13. “HTTP headers - HTTP | MDN.” Accessed: Jan. 16, 2024. [Online]. Available: <https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers>

14. [14] “Request header - MDN Web Docs Glossary: Definitions of Web-related terms | MDN.” Accessed: Jan. 16, 2024. [Online]. Available: https://developer.mozilla.org/en-US/docs/Glossary/Request_header
15. “Bot Detection Using Browser Fingerprinting - Fingerprint BotD,” Fingerprint. Accessed: Jan. 16, 2024. [Online]. Available: <https://fingerprint.com/blog/bot-detection/>
16. M. Singh and D. Garg, “Choosing Best Hashing Strategies and Hash Functions,” in *2009 IEEE International Advance Computing Conference*, Mar. 2009, pp. 50–55. doi: 10.1109/IADCC.2009.4808979.
17. “Hash function,” *Wikipedia*. Jan. 15, 2024. Accessed: Jan. 17, 2024. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Hash_function&oldid=1195921042
18. I. B. Damgård, “A Design Principle for Hash Functions,” in *Advances in Cryptology — CRYPTO’ 89 Proceedings*, G. Brassard, Ed., in Lecture Notes in Computer Science. New York, NY: Springer, 1990, pp. 416–427. doi: 10.1007/0-387-34805-0_39.
19. C. Crane, “What Is a Hash Function in Cryptography? A Beginner’s Guide,” Hashed Out by The SSL Store™. Accessed: Jan. 17, 2024. [Online]. Available: <https://www.thesslstore.com/blog/what-is-a-hash-function-in-cryptography-a-beginners-guide/>
20. “Best Privacy Policy Examples for GDPR,” Ironclad. Accessed: Jan. 17, 2024. [Online]. Available: <https://ironcladapp.com/journal/contracts/best-privacy-policy-examples-for-gdpr/>
21. “Teaching History with 100 Objects - Enigma cipher machine.” Accessed: Jan. 18, 2024. [Online]. Available: http://www.teachinghistory100.org/objects/about_the_object/enigma_cipher_machine

Analysis of March 2025 Myanmar Earthquake using DInSAR Techniques based on Sentinel – 1 Data

Pralay Bhaumik¹, Himanshu Govil¹, Shashi Kumar² and Purabi Chanda¹

¹ Department of Applied Geology, National Institute of Technology (NIT), Raipur,
Chhattisgarh, 492010, India

² Photogrammetry & Remote Sensing Department, Indian Institute of Remote Sensing, Dept.
of Space, 4, Kalidas Road, Dehradun – 248001, India
hgovil.geo@nitrr.ac.in

Abstract. On March 28, 2025, a powerful magnitude 7.7 earthquake struck near Mandalay along the Sagaing Fault, followed by numerous aftershocks, including a magnitude 6.7 event. This paper presents an analysis of the 2025 Myanmar earthquake sequence using Differential Interferometric Synthetic Aperture Radar (DInSAR) techniques with Sentinel-1 satellite data. The earthquake caused widespread destruction and significantly deformed the Earth's surface. Myanmar's tectonic setting—at the convergence of the Indian, Eurasian, Sunda, and Burma plates—makes it highly susceptible to seismic activity. The study utilized pre- and post-event Sentinel-1 SAR images to map co-seismic deformation. Through a systematic DInSAR processing workflow, including coregistration, interferogram generation, filtering, phase unwrapping, and terrain correction, the research produced detailed displacement maps. These revealed significant strike-slip deformation along a ~170 km segment of the Sagaing Fault, with peak line-of-sight displacements of up to 46 cm. Spatial patterns of subsidence were consistent with right-lateral fault motion, and notable asymmetry in the deformation field suggested heterogeneous slip and secondary effects. The analysis underscores DInSAR's value in rapid seismic assessment, especially in regions with sparse ground instrumentation. The study also highlighted limitations such as decorrelation in vegetated areas. It concludes that DInSAR analysis based on Sentinel-1 data is a critical tool for seismic hazard monitoring and recommends future work involving multi-sensor data integration, time-series InSAR, and machine learning to enhance monitoring capabilities in tectonically active regions.

Keywords: Earthquake, Sagaing fault, DInSAR, strike-slip deformation.

1. Introduction

The 2025 Myanmar earthquake, which struck on 28 March at 06:20 UTC (12:50 local) with a moment magnitude of 7.7, represents one of the most significant seismic events in the region since records began. The event, with its epicenter near Mandalay, resulted in widespread destruction and significant ground deformation, presenting a unique opportunity for remote sensing researchers to study the mechanics of fault motion and the complex interplay of tectonic forces.

Myanmar's geographic position along the complex boundary between the Indian, Eurasian, and Sunda plates renders it highly susceptible to tectonic hazards. Over the last century, earthquakes associated with the Sagaing, Shan Scarp, and other fault systems have highlighted the need for effective seismic monitoring.

In this context, satellite-based remote sensing—especially Interferometric Synthetic Aperture Radar (InSAR)—has emerged as a valuable tool [1]. DInSAR, which compares multiple SAR images to measure millimeter-scale changes in surface elevation, enables precise mapping of earthquake-induced deformation [2], [3]. Sentinel-1 satellites, operated by the European Space Agency (ESA), provide frequent, high-resolution SAR data ideal for such applications [4], [5]. Traditional seismological data and GPS observations are often sparse or delayed in developing countries like Myanmar, creating a gap in immediate response and scientific understanding. In such scenarios, satellite-based measurements have become indispensable for rapid hazard assessment and tectonic studies [6], [7]. With Sentinel-1's wide swath and revisit time, vital

interferometric products can be generated within days of the event, aiding disaster-response teams and informing geophysical analysis [8], [9].

This paper presents a DInSAR analysis of the 2025 Myanmar earthquake sequence using Sentinel-1 data to assess DInSAR's strengths and challenges in a tectonically complex region. Through interferometric processing and deformation quantification, the paper aims to contribute to the ongoing efforts in regional seismic hazard mapping and emergency mitigation.

2. Geological and Tectonic Background

Myanmar is located at the junction of four major tectonic plates—the Indian, Eurasian, Sunda, and Burma plates [10]—making it highly susceptible to frequent and intense seismic activity. The oblique collision between the Indian and Sunda Plates creates a wide deformation zone that includes the Indo-Burman Ranges, Central Myanmar Basin, and the Shan Plateau. A key structural feature in this tectonic setting is the Sagaing Fault [11], a right-lateral strike-slip fault that accommodates the northward movement of the Indian Plate relative to the Sunda Plate at a rate of 18–20 mm per year.

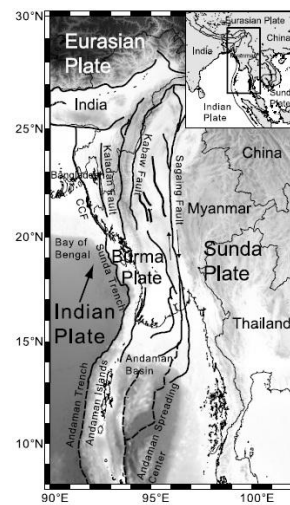


Fig 1. Regional map showing tectonic setting of present study (reproduced from [10])

The primary causative structure for the 2025 event is the Sagaing Fault system, which has been responsible for several historical large-magnitude earthquakes, including events in 1930 (M7.4), 1946 (M7.7), and more recently in 2012 (M6.8). Due to the proximity of the fault to densely populated areas such as Mandalay and Naypyidaw, it represents a significant seismic hazard.

The fault segment near Mandalay had been monitored as a potential seismic hazard. Official figures quoted extensive damage with casualties exceeding 5,400 and injuries exceeding 11,400, thus highlighting the need for both rapid and detailed post-event assessment. The extent of the 2025 rupture and the accompanying secondary deformation is a topic of active research now. It is worth mentioning here that the interplay between structure, ground motion, and local geology is critical while interpreting the interferometric products derived from Sentinel-1 images.

3. Description of satellite data used in the study

c. Sentinel-1 SAR Data

The European Space Agency's Sentinel-1 mission consists of twin satellites (Sentinel-1A and 1B) that, in general, provide C-band SAR imagery with 12-day repeat cycles in Interferometric Wide (IW) mode [12]. The IW mode covers a 250 km swath with a 5 X 20 m spatial resolution, ideal for regional earthquake studies [13], [14]. These data sets were acquired from the Copernicus Open Access Hub.

For the event under study, Sentinel-1 images were acquired both before the earthquake (baseline image) and shortly after the event on 28 March 2025. The choice of a 12-day interval was based on operational considerations and the need to balance temporal decorrelation and atmospheric effects [15], [16]. The choice of Sentinel-1 data ensured optimal conditions for DInSAR analysis.

d. Image Selection

Pre-event Image: Acquired to capture the undisturbed state of the ground surface on 22 March 2025.

Post-event Image: Captured on 03 April 2025 to cover the immediate deformation signal as a result of the seismic event.

Table 3. Summary of Sentinel-1 Data Parameters.

Parameter	Pre-Event Acquisition	Post-Event Acquisition
Acquisition	22 Mar 2025, 11:39	03 Apr 2025, 11:39
Date/Time	UTC	UTC
Orbit Direction	Ascending	Ascending
Incidence Angle	~34°	~34°
Wavelength (C-band)	5.6 cm	5.6 cm
Spatial Resolution	~5 m X 20 m	~5 m X 20 m
Swath Mode	Interferometric Wide	Interferometric Wide
Baseline (Temporal)	12 days	
Baseline (Perpendicular)	1.32	

4. Methodology

e. DInSAR Workflow

The DInSAR processing chain includes the following steps:

1. Coregistration of pre- and post-event SAR images
2. Interferogram generation by subtracting the SAR phase
3. Filtering to reduce noise in the interferogram
4. Phase unwrapping to retrieve continuous deformation
5. Conversion of phase to LOS displacement
6. Terrain Correction of the LOS displacements to map deformations in geographic coordinates

Processing was performed using the ESA SNAP toolbox and SNAPHU for phase unwrapping and geocoding. A precise Shuttle Radar Topography Mission (SRTM) 1-arcsecond DEM was used for topographic correction.

7. Discussions

Some critical insights into the present study are furnished below.

a. Coherence

Coherence is a critical factor in determining the quality of any DInSAR study. Coherence values range between 0.5 and 1.0. Values below 0.5 still contain useful deformation information but are typically accompanied by increased noise, which inversely correlates with coherence. In the current study, Figure 2 presents the coherence map, indicating an average coherence below 0.5, suggesting a moderate level of noise. Despite this, the generated interferogram

(Figure 3) retains interpretable deformation patterns, validating the usefulness of the image pair.

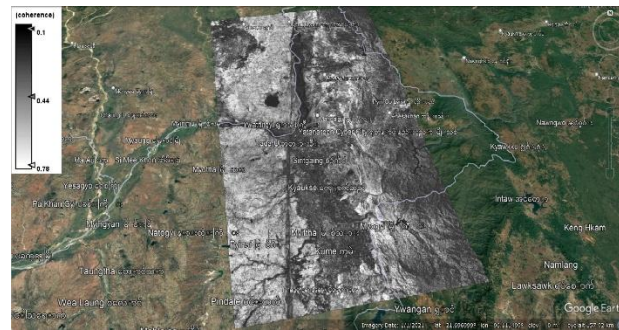


Fig 2. Coherence Image

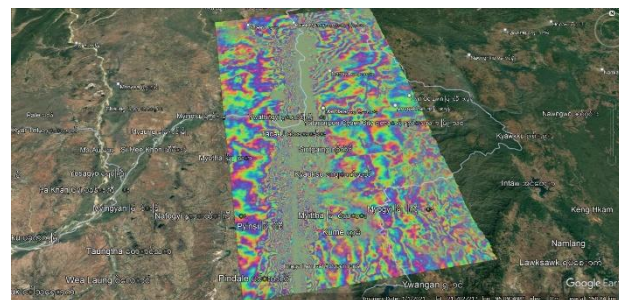


Fig 3. Interferogram

b. Seismic Hazard Assessment

The concentric fringe patterns indicate strike-slip deformation centered on the fault trace. The deformation patterns derived from our analysis are consistent with theoretical models of rupture mechanics along strike-slip faults. The predominant north-south motion of the Sagaing Fault, modulated by local complexities and secondary aftershocks, produces the observed pattern of lateral displacement manifested in distinct uplift and subsidence zones. These findings support the hypothesis that the 2025 earthquake was not a uniform rupture but rather a complex faulting sequence involving heterogeneous slip amplitudes. Such variations are essential for understanding stress distribution along the fault and for improving seismic hazard models in the region.

Additionally, the distinct deformation features provide insights into the depth and extent of slip, as well as the potential for future seismic hazards. In the context of ongoing tectonic activity in Myanmar and its neighbouring countries, continuous monitoring using Sentinel-1 and complementary sensors is strongly recommended.

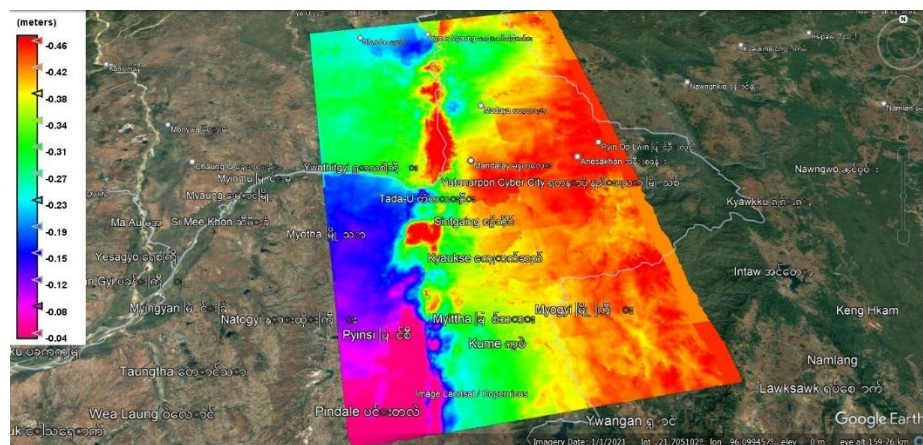


Fig 4. Sentinel-1 Co-seismic LOS Deformation Map (raw)

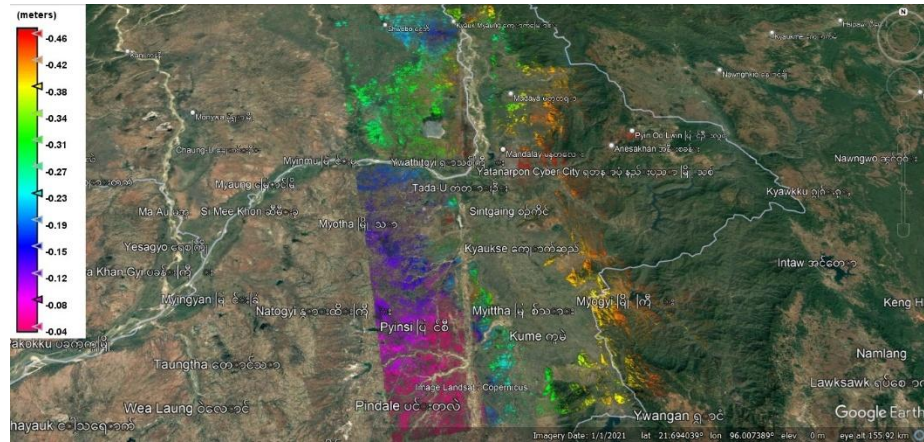


Fig 5. Sentinel-1 Co-seismic LOS Deformation Map (filtered)

c. Tectonic Implications

The 2025 Myanmar earthquake reveals several critical insights into the ongoing tectonic activity of the central Sagaing Fault:

1. **Segment Reactivation:** The event ruptured a segment previously considered to be a seismic gap. This has implications for recurrence intervals and hazard modelling along the central fault.
2. **Urban Exposure:** Although rural regions bore the brunt of the damage, cities like Mandalay and Naypyidaw are within 100 km of the rupture, underscoring the need for urban seismic preparedness.
3. **Impact on Sagaing Fault Corridor:** The corridor along the Sagaing Fault revealed the clearest deformation signal. Our analysis indicates that the rupture propagated primarily along a narrow band with a pronounced asymmetry. The eastern segment showed uplift, whereas the western segment was marked by subsidence. This dichotomy is in line with a dextral (right-lateral) strike-slip mechanism, although secondary vertical components were also noted.
4. **Post-seismic Activity:** Continued deformation suggests after slip or stress readjustments, which could load adjacent fault segments.

8. Conclusions

The 2025 Myanmar earthquake offered a critical opportunity to study near-field co-seismic and post-seismic deformation using Sentinel-1-based DInSAR techniques. This research demonstrated:

- DInSAR can capture the spatially varied deformation fields associated with the event, including notable uplift and subsidence patterns along the Sagaing Fault.
- Significant right-lateral strike-slip displacement along the central Sagaing Fault, consistent with regional tectonics.
- Peak line-of-sight displacements exceeding -46 cm.
- A well-defined rupture length (~170 km) and deformation width (12–15 km)
- Post-seismic after slip activity extending beyond the main rupture, indicating stress redistribution and potential for future rupture on adjacent segments.

The use of freely available Sentinel-1 data enabled rapid and accurate quantification of crustal movements, even in areas with limited ground instrumentation. These findings contribute directly to hazard assessment, fault characterization, and seismic risk mitigation in Myanmar and similar tectonic environments.

References

1. M. Crosetto *et al.*, “State-of-the-Art of Land Deformation Monitoring Using Differential SAR Interferometry,” in *ISPRS Workshop*, Hannover, Germany, Jan. 2005. [Online]. Available: <https://www.researchgate.net/publication/228351503>
2. T. Tampuu, J. Praks, R. Uiboupin, and A. Kull, “Long term interferometric temporal coherence and DInSAR phase in Northern Peatlands,” *Remote Sens (Basel)*, vol. 12, no. 10, May 2020, doi: 10.3390/rs12101566.
3. Y. Morishita, “Nationwide urban ground deformation monitoring in Japan using Sentinel-1 LiCSAR products and LiCSBAS,” *Prog Earth Planet Sci*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40645-020-00402-7.
4. S. Moretto, F. Bozzano, and P. Mazzanti, “The role of satellite insar for landslide forecasting: Limitations and openings,” *Remote Sens (Basel)*, vol. 13, no. 18, Sep. 2021, doi: 10.3390/rs13183735.
5. S. K. Mondal, R. Bharti, and K. F. Tiampo, “Tracking the Seismic Deformation of Himalayan Glaciers Using Synthetic Aperture Radar Interferometry,” *Remote Sens (Basel)*, vol. 17, no. 5, p. 911, Mar. 2025, doi: 10.3390/rs17050911.
6. G. Acosta, M. A. Sánchez, F. Ruiz, F. Carballo, and M. F. Leiva, “Coseismic geodetic and geophysical variations detected for the January 19th, 2021 San Juan earthquake. Tectonic implications,” *J South Am Earth Sci*, vol. 151, Jan. 2025, doi: 10.1016/j.jsames.2024.105262.
7. H. B. Ibrahim, M. Salah, F. H. F. Zarzoura, and M. El-Mewafi, “Mapping of surface deformation associated with the 7.8 magnitude Turkey earthquake of 6 Feb 2023 using radar interferometry,” *Innovative Infrastructure Solutions*, vol. 10, no. 2, Feb. 2025, doi: 10.1007/s41062-024-01856-0.
8. Z. Lee *et al.*, “Fast report: surface deformation associated with the 2025 Dapu earthquake,” *Terrestrial, Atmospheric and Oceanic Sciences*, vol. 36, no. 1, Dec. 2025, doi: 10.1007/s44195-025-00090-0.
9. G. Sharma, M. S. Singh, K. Nayak, P. P. Dutta, K. K. Sarma, and S. P. Aggarwal, “Earthquake Damage Susceptibility Analysis in Barapani Shear Zone Using InSAR, Geological, and Geophysical Data,” *Geosciences (Switzerland)*, vol. 15, no. 2, Feb. 2025, doi: 10.3390/geosciences15020045.
10. N. Hurukawa, P. P. Tun, and B. Shibazaki, “Detailed geometry of the subducting Indian Plate beneath the Burma Plate and subcrustal seismicity in the Burma Plate derived from joint hypocenter relocation,” *Earth, Planets and Space*, vol. 64, no. 4, pp. 333–343, 2012, doi: 10.5047/eps.2011.10.011.
11. S. Thura Tun, “Sagaing Fault: A desk study report on seismotectonic implications in Myanmar,” 2015, doi: 10.13140/RG.2.1.4048.6249.
12. T. Singh *et al.*, “The Afghanistan Earthquake of 21 June, 2022: The Role of Compressional Step-Overs in Seismogenesis,” Feb. 03, 2025. doi: 10.20944/preprints202502.0098.v1.
13. J. A. H. Doski, “Surface deformation of the 26 January 2021 earthquake in the Sinjar – Hasakah Area, N Iraq and NE Syria, from Sentinel-1A InSAR images,” *Egyptian Journal of Remote Sensing and Space Science*, vol. 28, no. 1, pp. 128–137, Mar. 2025, doi: 10.1016/j.ejrs.2025.02.001.
14. S. Noriega-Londoño, M. A. Bermúdez, M. I. Marín-Cerón, S. A. Restrepo-Moreno, and H. García-Delgado, “Earthquake ground deformation using DInSAR analysis and instrumental seismicity: The 2019 M 6.0 Mesetas Earthquake, Meta, Colombia,” *Boletín de la Sociedad Geológica Mexicana*, vol. 73, no. 2, pp. 1–18, 2021, doi: 10.18268/BSGM2021v73n2a090221.
15. T. M. Fentahun, M. Bagyaraj, M. A. Melesse, and T. korme, “Seismic hazard sensitivity

- assessment in the Ethiopian Rift, using an integrated approach of AHP and DInSAR methods,” *Egyptian Journal of Remote Sensing and Space Science*, vol. 24, no. 3, pp. 735–744, Dec. 2021, doi: 10.1016/j.ejrs.2021.05.001.
16. P. Roy, T. R. Martha, K. V. Kumar, and P. Chauhan, “Coseismic deformation and source characterisation of the 21 June 2022 Afghanistan earthquake using dual-pass DInSAR,” *Natural Hazards*, vol. 118, no. 1, pp. 843–857, Aug. 2023, doi: 10.1007/s11069-023-06030-w.

MPC for Quadruped Robot

Tapan Ranjit Pakhidde and Prakash D Pantawane

Manufacturing Engineering and Industrial Management, COEP Technological University
pakhiddetr23.mfg@coeptech.ac.in, pdpantawane.mfg@coeptech.ac.in

Abstract

This paper presents a hierarchical control strategy for dynamic locomotion in quadruped robots, with a focus on trajectory tracking. The methodology integrates high-level trajectory planning with a low-level motion execution framework based on an online, implicit Model Predictive Control (MPC) scheme. The MPC utilizes a simplified kinematic model to compute optimal control actions by solving a Quadratic Program at each time step. We describe the control architecture and detail the principles of MPC. The system's performance is evaluated within a high-fidelity MATLAB/Simulink simulation. Results demonstrate stable tracking at a design velocity of 15 m/s. However, testing at 30 m/s reveals stability limitations, highlighting the critical challenge of model-plant mismatch at high speeds and the importance of robust controller tuning.

Keywords: Model Predictive Control (MPC), Motion Control, Trajectory Tracking, Online Optimization, Stability Analysis

1. Introduction

Achieving dynamic and robust locomotion in legged robots is a grand challenge due to their under actuated, high-dimensional, and hybrid dynamics. A successful controller must manage balance, coordinate limbs, and respect physical constraints. Model Predictive Control (MPC) has emerged as the state-of-the-art strategy for this task [1]. By optimizing control actions over a future horizon, MPC proactively handles complex dynamics and constraints, enabling remarkable feats of locomotion on platforms like the MIT Cheetah [2, 3] and ANYmal [4, 5].

While these advanced platforms often use complex dynamic models, the core principles of MPC for path tracking can be effectively studied using simplified models. The contribution of this work is not a novel algorithm, but a systematic analysis of the performance boundaries of an online MPC controller when faced with significant model-plant mismatch. This paper is structured as follows: Section II reviews related work, Section III details the methodology, Section IV presents the performance evaluation, and Section V concludes the paper.

2. Related Work

The application of MPC to legged locomotion has evolved along two main paths. Convex MPC, pioneered on robots like the MIT Cheetah [2], uses simplified linear models to formulate the control problem as a highly efficient Quadratic Program (QP). This is often used in a hierarchical structure where the MPC computes ground reaction forces for a low-level controller [3, 7]. In contrast, Nonlinear MPC (NMPC), used on platforms like ANYmal [4], employs more complex models for higher fidelity at a greater computational cost. Both approaches rely on a simplified predictive model [6], a paradigm this work adopts to analyze the performance limits imposed by model-plant mismatch

3. Methodology and Simulation Framework

To analyze the MPC's performance, we constructed a simulation environment in MATLAB and Simulink, leveraging the Automated Driving Toolbox and Model Predictive Control Toolbox.

3.1. Control Strategy Selection: MPC

Controlling a quadruped is a Multi-Input, Multi-Output (MIMO) problem with significant cross-coupling. Model Predictive Control is a superior native MIMO strategy that uses an internal plant model to account for these interactions. Crucially, MPC can systematically handle physical constraints (e.g., actuator limits) and incorporates feedforward control by using future reference information, leading to better tracking performance [11]. For these reasons, MPC was selected, using the centralized architecture shown in Figure 1.

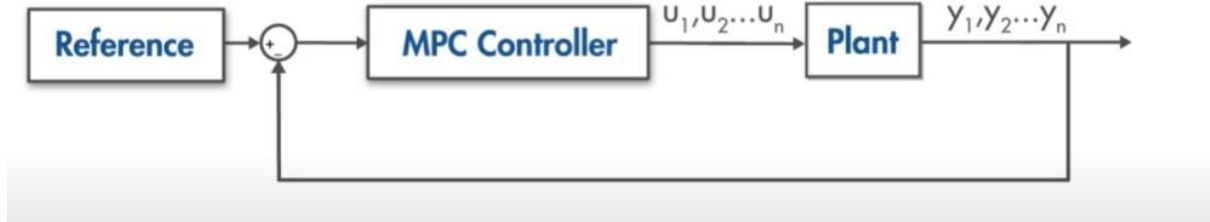


Figure 1: Centralized multivariable MPC architecture. A single controller manages all inputs and outputs by considering their interactions.

3.2. Hierarchical Control Architecture for Legged Robots

Our framework emulates the state-of-the-art hierarchical control architecture common in dynamic quadrupeds (Fig. 2) [2, 5]. This structure decouples the control problem: a high-level, low-frequency MPC plans optimal ground reaction forces using a simplified model, while a low-level, high-frequency whole-body controller calculates the joint torques required to realize those forces and track leg trajectories.

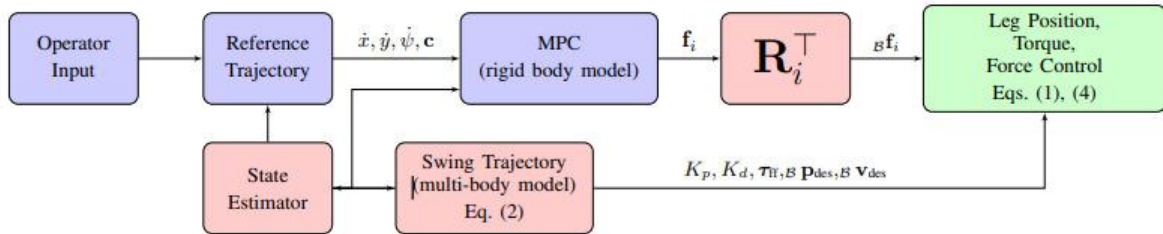


Figure 2: A state-of-the-art hierarchical control architecture for a quadruped robot, as implemented on the MIT Cheetah 3. Adapted from [2].

3.3. MPC Cost Function and Optimization

At the heart of MPC is an optimization problem that seeks the best sequence of future control actions over a prediction horizon, P . The "best" scenario is determined by a cost function, J , which is minimized at each time step. A standard quadratic cost function for path tracking is used:

$$J = \sum_{i=1}^P \|x_{k+i} - x_{k+i}^{ref}\|_Q^2 + \sum_{i=0}^{P-1} \|\Delta u_{k+i}\|_R^2 \quad (1)$$

where x is the state, u is the control input, and the matrices Q and R are weights that balance tracking error against control effort. As shown in Figure 3, the controller evaluates multiple potential futures and selects the one with the lowest cost.

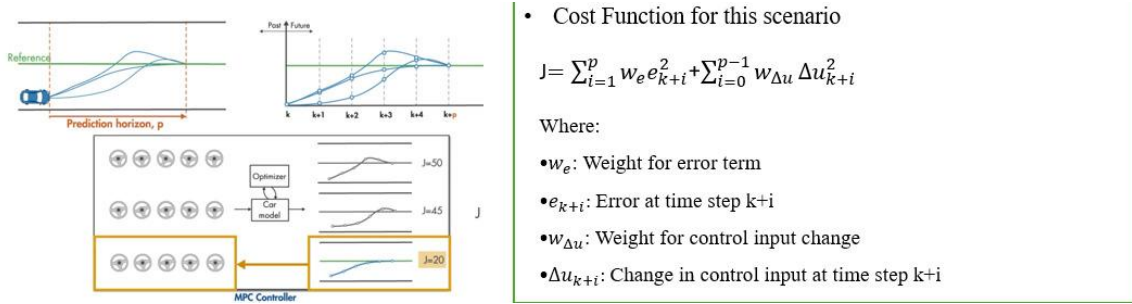


Figure 3: Conceptual diagram of the MPC process. The controller evaluates multiple future scenarios over a prediction horizon and selects the one with the lowest cost.

3.4. Predictive Model and Controller Design

For path tracking, we use the kinematic bicycle model (Figure 4), a standard choice that focuses on geometry and motion [11]. The controller was tuned using the MPC Designer app in Simulink. The nonlinear plant model was linearized at 15 m/s to obtain the MPC's prediction model. We configured a prediction horizon (P) of 10 steps and a control horizon (M) of 3 steps. Constraints were placed on the steering angle and its rate of change. Tuning weights were adjusted to prioritize path tracking while penalizing aggressive control actions.

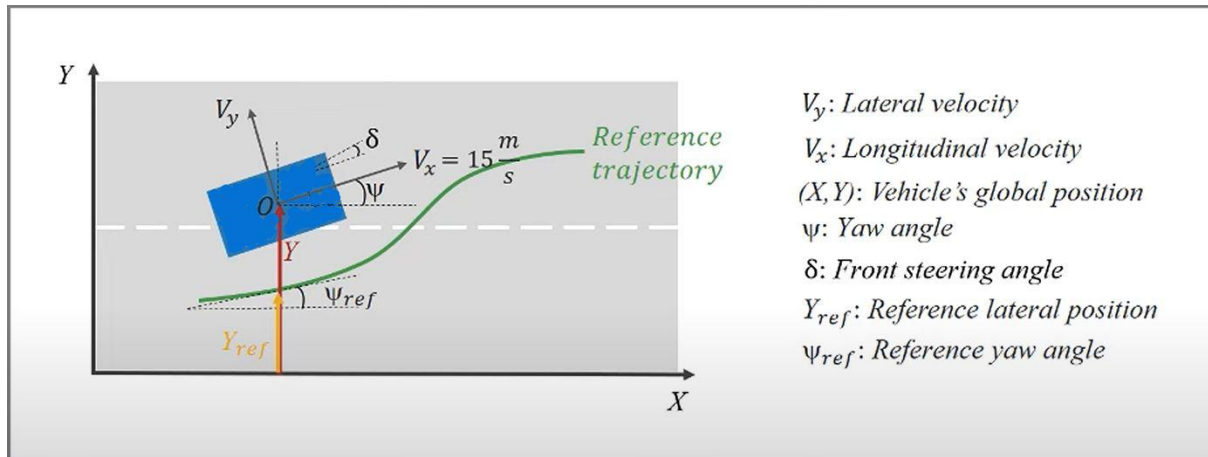


Figure 4: Kinematic bicycle model used for path tracking [11].

3.5. MPC Implementation: Online vs. Explicit

The controller is an Online, Implicit MPC, the standard implementation in the MATLAB/Simulink MPC Toolbox. At every time step, it solves a constrained optimization problem—a Quadratic Program (QP)—online using the current state of the system. This receding horizon strategy is highly reactive to disturbances, in contrast to Explicit MPC, which uses a pre-computed control law [8].

4. Performance Evaluation and Analysis

The simulation was run to track a reference path defined by the waypoints in Table 1, visualized in Figure 5.

Table 1: Waypoint parameters defining the reference path for the plant.

No.	X(m)	Y(m)	V(m/s)	Yaw (degree)
1	1.9000	-1.9000	15.0000	1.1067
2	8.1000	-1.8000	15.0000	0.5587
3	12.0000	-1.7000	15.0000	0.6927
4	15.6000	-1.7000	15.0000	0.8866
5	20.4000	-1.4000	15.0000	5.6607
6	24.7000	-0.8000	15.0000	11.1740
7	30.0000	0.4000	15.0000	10.6965
8	34.1000	1.2000	15.0000	16.1187
9	38.0000	2.4000	15.0000	9.4791
10	40.8000	2.5000	15.0000	-1.6697

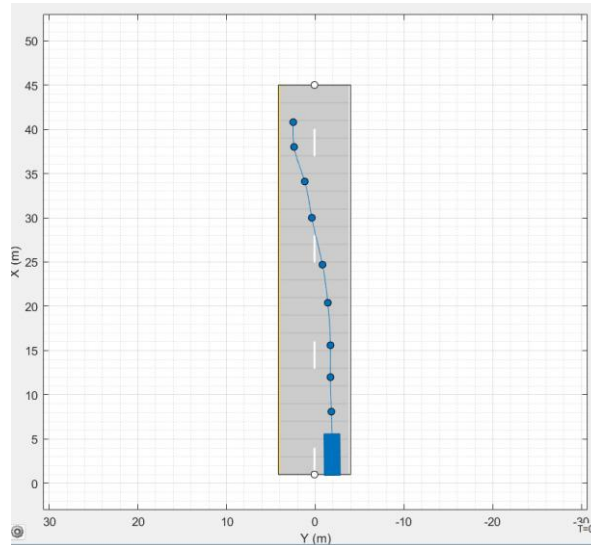


Figure 5: Visualization of the reference path defined by the waypoints.

4.1. Performance at Design Velocity (15 m/s)

Figure 6 shows excellent tracking performance at 15 m/s. The controller exhibits a well damped response with tight tracking of the reference, indicating the predictive model is sufficiently accurate at this speed.

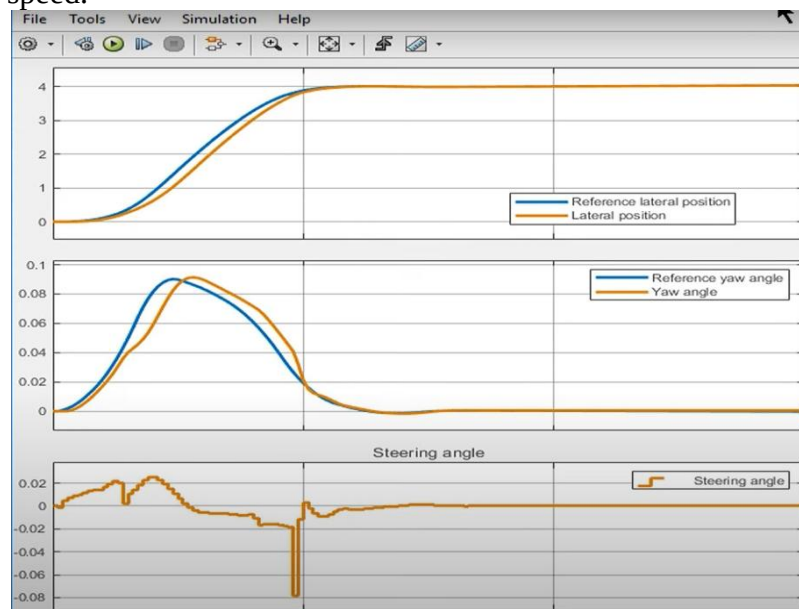


Figure 6 shows excellent tracking performance at 15 m/s. The controller exhibits a well damped

4.2. Performance at High Velocity (30 m/s)

Figure 7 shows the performance of the exact same controller at 30 m/s. The system becomes violently unstable. The robot's states diverge with rapidly growing oscillations, representing a complete loss of control.

4.3. Discussion

The instability at 30 m/s is a direct result of model-plant mismatch [6]. The simple kinematic model used by the MPC neglects crucial dynamic effects (e.g., body roll, actuator delays) that become significant at high speeds. Blind to these unmodeled dynamics, the

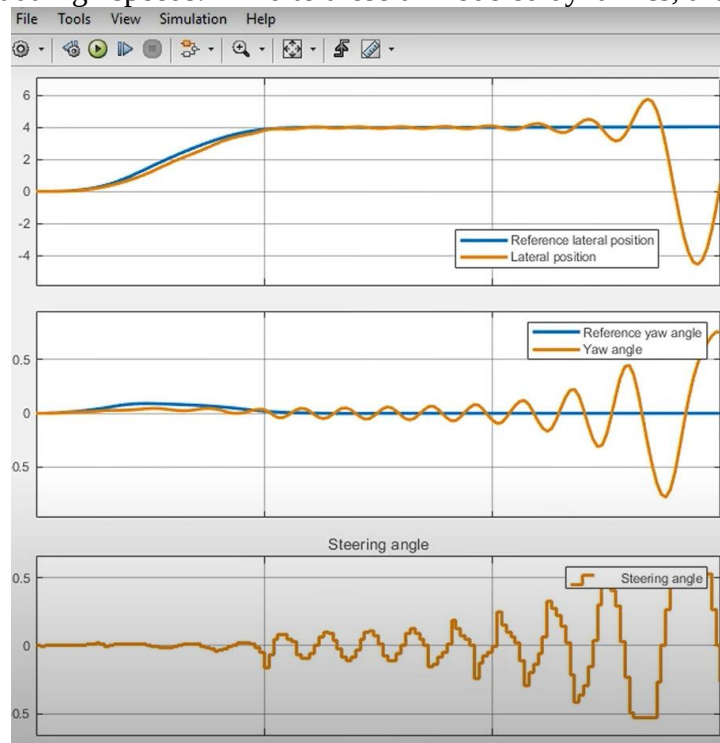


Figure 7: MPC performance at 30 m/s. The same controller becomes unstable due to model mismatch.

controller issues commands that are too aggressive for the physical system, causing it to overshoot and leading to a positive feedback loop of escalating corrective actions and, ultimately, instability.

5. Conclusion and Future Scope

This paper analyzed a standard online MPC for path tracking, demonstrating its effectiveness at its design velocity and its failure at higher speeds due to model-plant mismatch. This highlights the fundamental trade-off between model simplicity and controller robustness. Future work should focus on mitigating this mismatch. A promising direction is the integration of intelligent, data-driven techniques. For example, a Neuro-Fuzzy inference system could be used to adapt the MPC's tuning parameters or even the predictive model itself in real-time, based on the robot's operating conditions. This would enhance robustness and could be extended to incorporate perception data for dynamic obstacle avoidance, leading to more capable and autonomous legged robots.

References

1. S. Kuindersma, R. Deits, M. Fallon, et al., “Optimization-based Locomotion Planning, Tracking, and Fall Recovery for the Atlas Humanoid Robot,” *Autonomous Robots*, vol. 40, no. 3, pp. 429–455, 2016.
2. J. Di Carlo, P. M. Wensing, B. Katz, G. Bledt, and S. Kim, “Dynamic Locomotion in the MIT Cheetah 3 Through Convex Model-Predictive Control,” in *Proc. IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2018, pp. 1–9.
3. D. Kim, J. Di Carlo, B. Katz, G. Bledt, and S. Kim, “Highly Dynamic Quadruped Locomotion via Whole-Body Impulse Control and Model Predictive Control,” in *Proc. Robotics: Science and Systems (RSS)*, 2019.
4. M. Neunert, M. Stauble, M. Gifftthaler, et al., “Whole-Body Nonlinear Model Predictive Control Through Contacts for Legged Robots,” *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1458–1465, 2018. [5] C. D. Bellicoso, M. G. thaler, F. Fankhauser, M. Hutter, and J. Buchli, “A Modular Framework for Locomotion and Manipulation,” in *Proc. IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2018, pp. 1–7
5. P. M. Wensing and D. E. Orin, “Generation of Dynamic Legged Locomotion Using a Dominant Dynamics-Based Planner and Controller,” in *Proc. IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2013, pp. 2371–2378.
6. G. Bledt, P. M. Wensing, S. G. Kim, and B. Katz, “Highly-Dynamic Legged Locomotion Via Force-Control of a Center-of-Mass-Jacobian,” in *Proc. IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2018, pp. 1–8.
7. A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos, “The explicit linear quadratic regulator for constrained systems,” *Automatica*, vol. 38, no. 1, pp. 3–20, 2002.
8. J. Carpentier and N. Mansard, “Analytical Derivatives of Rigid Body Dynamics Algorithms,” in *Robotics: Science and Systems (RSS)*, 2018.
9. M. Grandin, F. Fankhauser, P. F. e Silva, et al., “Perceptive Locomotion in Rough Terrain for Quadrupedal Robots,” in *Proc. IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2020, pp. 7991–7997.
10. R. Rajamani, *Vehicle Dynamics and Control*, 2nd ed. Springer, 2011.
11. J. J. Craig, *Introduction to Robotics: Mechanics and Control*, 3rd ed. Pearson/Prentice Hall, 2005.
12. B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, *Robotics: Modelling, Planning and Control*. Springer, 2009.

Automated Customer Segmentation AI-Powered Lead Scoring for Edtech

Mahesh Sunil Ulhe¹ and Prof. Suhas S. Mohite²

¹Department of Manufacturing Engineering and Industrial Management and ²Department of Mechanical Engineering, COEP Technological University, Pune, India

Abstract

Ed.Tech companies collect vast amounts of data, such as browsing behavior, email engagement, and other contact details, which can be leveraged through predictive analytics to estimate a lead's purchase probability. This study investigates the use of machine learning for prospect scoring using a dataset of approximately 9,000 educational lead records. The objective is to enhance lead conversion rates by predicting the likelihood of conversion using historical behavioral data and engagement metrics. The problem is approached as a binary classification task, where supervised learning algorithms such as logistic regression, decision tree, and ensemble methods like random forest are applied. Purchase timestamps are used to define activity windows for converted leads, ensuring fair data representation. The models are evaluated using accuracy, precision, recall, and ROC-AUC. Among them, logistic regression achieved the highest accuracy and interpretability, while random forest provided valuable insights through feature importance analysis. The results demonstrate that machine learning-driven lead scoring can effectively prioritize high-potential leads, optimize marketing and sales strategies, and offer actionable business insights through visual analytics for decision makers.

Keywords

Lead Scoring, Machine Learning, Customer Identification, EdTech, Predictive Analytics

1. Introduction

In today's competitive market, effective customer acquisition is vital, and lead scoring plays a key role by helping businesses prioritize potential customers based on their engagement behaviors, such as website visits and email interactions [1]. Traditionally, this has been a manual process, assigning importance to each customer activity to rank leads. However, manual methods are often limited in scale and accuracy.

This article explores how machine learning can automate and enhance lead scoring in the B2C sector. Using real-world data, various models are developed and evaluated to overcome data preparation challenges and improve prediction accuracy. With a historical conversion rate of 30–40%, the goal is to help businesses target high-potential leads and increase conversions to 80%. Visual analytics are also applied to reveal actionable insights, supporting smarter decision-making and improving overall marketing efficiency through data-driven strategies.

2. Background

In the digital age, businesses generate vast amounts of data [2], leading to a shift toward data-driven decision-making [3], especially in marketing and CRM. Relationship marketing, which focuses on creating value through ongoing collaboration with customers [4], relies heavily on digital data to stay competitive [5]. By analyzing interactions from digital channels, companies can better identify and convert leads.

Integrating business analytics and machine learning into CRM enhances customer tracking and lead scoring efficiency [6]. Traditional intuition-based methods are now being replaced by automated systems [7] that detect user behavior trends to predict conversions [8]. Despite available tools, practical research on applying automation across the B2C sales funnel remains limited, underscoring the need for further study [8].

2.1 Manual Lead Scoring

Before diving into automated approaches, it's important to understand the conventional method widely used in industry manual lead scoring. As noted by Marion [9], this approach presents several critical limitations. One of the main concerns is that manual lead scoring lacks a foundation in statistical evidence, often relying on subjective assessments rather than data-driven insights. Typically, it uses a variety of demographic, behavioral, or firmographic. Since this method usually depends on a scoring matrix, businesses must regularly revise and update it to stay aligned with changing market conditions a process that can be both labor-intensive and inefficient. Marion [9] highlights these issues through an experiment involving 800 leads evaluated using manual scoring. The results showed no significant difference between leads tagged as “ready for sales” and a random group of unscored leads. The study emphasizes that without a solid understanding of statistics, accurately assigning weights to lead behaviors is nearly impossible. Additionally, the manual process demands continuous adjustments, which consumes time that could be better allocated elsewhere. Bohlin [10] also

Table 1: Example manual lead scoring matrix [9]

Activity	Points
Form/Landing Page Submission	+5
Submitted “Contact Me” Form	+25
Received an Email	0
Email Open	+1
Email Clickthrough	+3
Registered for Webinar (Optional)	+3
Attended Webinar	+10
Downloaded a Document	+5
Visited a Landing Page	+2
Unsubscribed from Newsletter	-2
Watched a Demo	+8
Contact is a CXO	+5
Visited Trade Show Booth	+3
Visited Pricing Page	+10

critiques this approach, arguing that even when assumptions are used to develop rules and weights, manual lead scoring remains suboptimal.

2.2 Components of Lead Scoring

Lead scoring is a key part of CRM that assigns numerical values to prospects, helping prioritize leads based on conversion likelihood [11]. Higher scores guide leads to sales, while lower ones may enter nurturing workflows [12]. The model's success depends on selecting relevant variables, including implicit behavioral data and explicit user information [12]. Leading firms often use behavioral inputs and complex models for better performance [12].

As a predictive analytics method, lead scoring uses statistical tools to forecast outcomes [13]. Predictive marketing builds on this by personalizing customer journeys through data insights and lower computing costs [13]. Machine learning—particularly supervised learning—is commonly applied to predict lead conversions from historical data [14]. Bayesian networks [11] and modern ML approaches [15] enhance sales efficiency, even in limited data scenarios.

2.3 Machine Learning Applications in Customer Relationship Management

Machine learning has many effective applications in customer relationship management (CRM), enhancing decisions across the customer journey [16]. Key techniques include

classification, clustering, regression, forecasting, and visualization, using algorithms like decision trees, KNN, genetic algorithms, neural networks, and logistic regression. These approaches support tasks such as lead scoring, segmentation, and behavior prediction. Real-world use cases highlight their impact. A study in [17] built a loyalty prediction framework using random forest, logistic regression, and neural networks, with random forest showing strong accuracy and AUC. Another case in [18] combined a genetic algorithm and neural network for direct marketing, using feature selection to improve interpretability and profit-focused decision-making.

3. Literature Review

Traditional lead scoring methods often rely on manual rules that can be biased and subjective. To overcome this, predictive lead scoring leverages historical data and machine learning to identify traits linked to successful conversions. As Syam and Sharma [1] suggest, AI is transforming marketing decisions, with models like logistic regression and decision trees widely used to assess leads based on demographics and behavior. Chorianopoulos [6] and Duncan and Elkan [7] stress the value of analytics and probabilistic modeling in improving CRM and lead prioritization. Behavioral scoring, focusing on user interactions like site visits, is especially effective in fast-changing sectors like EdTech. Järvinen and Taiminen [8] also highlight real-time tools that enhance automated B2B marketing.

In Ed. Tech, AI can analyze user behavior to identify high-potential leads, streamlining marketing efforts. Research shows supervised learning models perform well when trained on features like traffic source, activity frequency, and geography. Marion [9] and Bohlin [10] argue that manual scoring is outdated, advocating for automation. Models are typically evaluated using ROC-AUC to measure classification accuracy [14]. Frameworks like Demandbase allow for real-time scoring by combining historical and current data, helping businesses focus on quality leads, improve conversions, and reduce wasted effort.

4. Methodology

The methodology adopted for this lead scoring project is structured around a clear, step-by-step machine learning workflow. It begins with gathering lead-related data, such as user activity and source of origin, followed by a thorough preparation phase. This includes managing missing entries, transforming categorical attributes through encoding, and normalizing numerical features to ensure uniformity. These preprocessing steps help prepare the dataset for reliable and effective model training. By refining the input data, the system becomes more capable of identifying meaningful patterns related to lead conversion behavior.

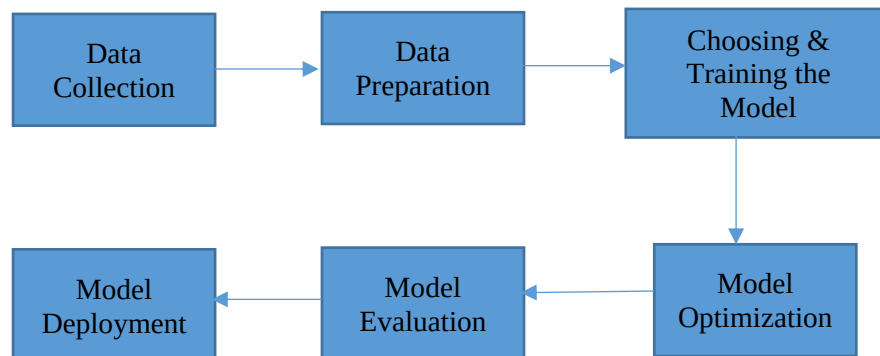


Figure 1: Proposed Methodology

Following data preparation, the project focuses on building and evaluating predictive models. Various algorithms like Logistic Regression, Decision Trees, and Random Forests are trained and compared using metrics such as AUC-ROC and lift curves. After evaluating the models, Logistic Regression is selected for its performance and ease of interpretation. The model helps

prioritize leads by assigning scores, allowing the business to focus on those most likely to convert. This data-driven prioritization is expected to significantly improve the efficiency of the sales team and boost overall lead conversion rates.

4.1 Dataset

The dataset analyzed in this study contains 9,241 records, each representing a potential lead for an Ed. Tech platform. It features 36 attributes that provide a well-rounded view of user behavior, demographics, and interactions. Key variables include lead source (e.g., Google, Facebook), lead origin (e.g., API, Landing Page Submission), and last activity (e.g., Email Opened, SMS Sent). Initially, the data is in an unstructured format. The dataset also includes numeric data such as total time spent on the website and page views per visit, offering insights into engagement levels. The target variable, lead conversion status, enables binary classification modeling. These varied data points make the dataset ideal for training machine learning models aimed at predicting conversion likelihood.

Collected from real operational data, the dataset reflects actual customer behavior, making it valuable for real-world applications like lead scoring, sales prioritization, and personalized marketing. Features such as academic specialization, course preferences, and city information add context that can enhance predictive accuracy. The presence of missing values and outliers introduces opportunities for data preprocessing, including imputation and outlier handling. With around 90% of the data usable for training and 10% for testing, the dataset is well-structured to support robust model development and performance evaluation, ultimately helping Ed.Tech businesses improve conversion strategies and customer engagement.

4.2 Data Preprocessing

To ensure consistency and quality, the following preprocessing steps were applied:

- **Handled Missing Data:** Removed columns with too many missing values and filled others with suitable replacements. For categorical data it filled with “Unknown” or “Not Specified” and for continuous data it filled with mean or median.
- **Dropped Unnecessary Columns:** Eliminated duplicate or irrelevant columns that didn’t contribute to the model.
- **Outlier Detection:** Use IQR (Inter Quantile Range) method for detecting the outliers and cap extreme values or remove them to get cleaner numeric data with minimized impact of outliers.
- **Feature Engineering:** To modify or edit the features in the dataset. To combine one or more features to make it single one to reduce the complexity of the model.

4.3 Modeling Approaches

Multiple machine learning algorithms were evaluated for lead score conversion. As our usecase is a classification model so that we used classification algorithms.

- **Logistic Regression:** A simple classification method that estimates the probability of a lead converting by fitting data to a logistic curve. As our project is binary classification so logistic regression is used, it performs well on binary classification.
- **Decision Tree:** Builds a tree-like model of decisions by splitting data based on feature values, making it easy to interpret outcomes. It suited for regression & classification problems, but it overfits the model.
- **Random Forest:** An ensemble method follows parallel approach that combines multiple decision trees to improve accuracy and reduce overfitting by averaging their results. Used for regression as well as classification problem.

4.4 Hyperparameter Tuning

To enhance model performance, Grid Search was employed for hyperparameter tuning. Parameters such as the number of estimators, maximum depth (for tree-based models), and learning rate (for boosting methods) were optimized using cross-validation to prevent overfitting and ensure generalization.

4.5 Evaluation Metrics

Models were evaluated using the following metrics:

- Accuracy: Represents the percentage of total correct predictions but can be unreliable when the dataset has imbalanced classes like more non-converted leads.
- Precision: Measures how many leads predicted as "converted" were actually correct, helping reduce false positives and improve targeting accuracy in marketing campaigns.
- Recall: Indicates how many actual converted leads were correctly identified by the model, ensuring fewer missed opportunities in lead follow-ups.
- ROC-AUC: Shows the model's ability to differentiate between converted and non-converted leads across various thresholds, with higher scores reflecting better performance. To compare the performance of different algorithms ROC-AUC curve will be used. It's a plot between FPR and TPR.

5. Results & Discussion

Based on the final dataset described in the previous section, three different machine learning algorithms were selected to be tested motivated by the findings in our literature review on the most widely used algorithms in customer relationship management:

- Logistic Regression (LR) [14]: A well-established generalized linear model frequently applied to binary classification tasks, estimating the likelihood of class membership based on input features. As our problem is binary classification so logistic regression is used, it performs well on binary classification.
- Decision Trees (DT) [20]: These models build hierarchical decision rules based on dataset features. They are widely appreciated for their interpretability and ability to explain predictions through logical if-then conditions. It suited for regression & classification problems, but it overfits the model.
- Random Forests (RF) [21]: A tree-based ensemble method that mitigates the overfitting tendency of individual decision trees by generating multiple de-correlated trees and averaging their outputs to make predictions. To compare the performance of different algorithms ROC-AUC curve will be used. It's a plot between FPR (False Positive Rate) and TPR (True Positive Rate).

To evaluate model effectiveness, several metrics derived from the confusion matrix were used [14]. These include counts of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), where a "positive" denotes a successfully converted lead. In addition to basic accuracy, metrics like precision, recall, sensitivity, and specificity were calculated to better understand the model's behavior under different types of errors.

A key evaluation metric applied was the Area Under the Receiver Operating Characteristic (ROC) Curve (AUC), which measures the model's ability to differentiate between classes by plotting true positive rates against false positive rates at various threshold levels.

Given the class imbalance in the dataset, SMOTE (Synthetic Minority Oversampling Technique) was used to create a more balanced training set. Additionally, 10-fold cross-validation was implemented to ensure robust and unbiased model performance estimates.

An initial exploration of data aggregation strategies was conducted, and after evaluating potential biases, a single aggregation method was selected for final model refinement and detailed analysis.

Data exploration reveals diverse lead characteristics. The distribution of lead sources (Figure 2) shows that prospects come from multiple channels (e.g. Google, Organic Search, etc.) without a single dominant source, underscoring the need to treat “Unknown” as a category during imputation. The distribution of total website visits (Figure 3) is right-skewed: most leads made only a few visits, but a minority visited frequently. Similarly, the page views per visit distribution (Figure 4) indicates most leads browsed a small number of pages per visit, with few high-engagement outliers. These behavioral features reflect varying levels of engagement. Lead Source Analysis (Figure 2): Leads came from various platforms (Google, ads, chat); missing sources were treated as 'unknown' to retain all data.

Website Visit Behavior (Figure 3): Most leads visited the site only 1–3 times, showing low engagement; data was right-skewed and scaled for modeling.

Page Views per Visit (Figure 4): Similar skew observed—most users viewed few pages, while some viewed many, indicating high intent; this was used as a behavioral signal.

Last Activity Engagement (Figure 5): Actions like email opened or SMS sent reflect engagement level; rare but important activities were retained or grouped for better prediction.

Country Information and Imputation (Figure 6): 26.6% of entries lacked country info; missing values were filled using city data where possible, mostly as 'India' or 'unknown'.

City-wise Distribution (Figure 7): A few cities (e.g., Mumbai, Thane) generated most leads; similar cities were combined, and city was kept as a key categorical feature.

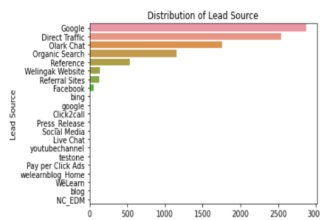


Figure 2

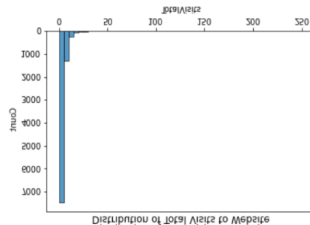


Figure 3

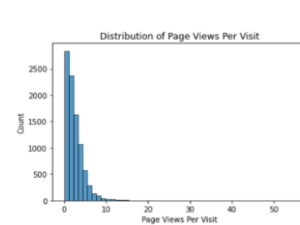


Figure 4



Figure 5



Figure 6

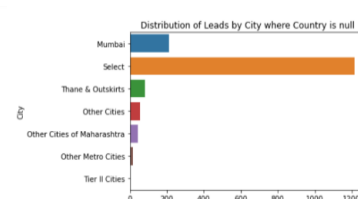


Figure 7

In summary, the exploratory analysis reveals that leads differ significantly in their origin, level of engagement, and geographical location. We move forward with the cleaned and encoded dataset into the process of training the model.

5.1 Model Performance: Table 2 provides an overview of the key performance metrics of our models, including accuracy, precision, recall, and F1 score, on both the training and test datasets. Logistic regression attained an AUC of 0.9217 on the training set and 0.9094 on the test set. The decision tree achieved an accuracy of 0.9105 (train) and 0.8956 (test). The random forest achieved an accuracy of 0.9106 in both the training

and test sets. All models have high AUCs (>0.89), but logistic regression is slightly better than the other methods on the validation set. This implies that the generalization is most effective, possibly because regularization helps prevent overfitting, while the tree and forest models show slightly lower test AUC scores. Model accuracy of the training set accuracy of testing set.

Table 2: AUC Scores of Different Models

Model	AUC of Training Set	AUC of Testing Set
Logistic Regression	0.9217	0.9094
Decision Tree	0.9105	0.8956
Random Forest	0.9106	0.9000

Figure 8 (below) displays the receiver operating characteristic (ROC) curves for each model on the test set. We observe that the logistic model's receiver operating characteristic (ROC) curve is slightly above the random forest's, resulting in the highest area under the curve (AUC). The variation is slight, suggesting that all models are fairly effective in distinguishing between the two groups. Considering the logistic model's simplicity and interpretability, we choose it as the primary scoring model for our subsequent analysis.

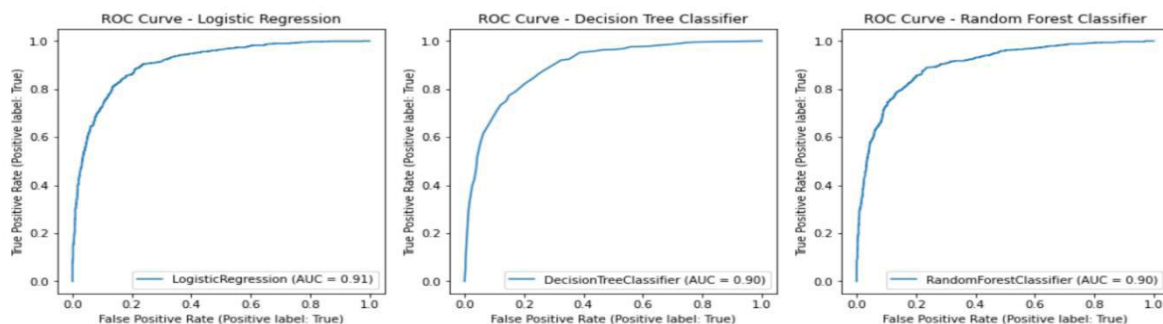


Figure 8: ROC Curves of Different Models

The chosen logistics model assigns a probability score for conversion to each lead. To translate this into a sales strategy, we calculate a lift curve. Based on the data, the average conversion rate was around 38%. To achieve an 80% conversion rate (more than double the baseline), it would be necessary to prioritize the leads with the highest scores. Our lift analysis indicates that reaching out to approximately the top 30–35% of leads based on their score results in an estimated conversion rate of around 80%. In essence, if sales reps focus on the top third of scored leads, they can achieve the CEO's goal of achieving an 80% conversion rate. By automating lead prioritization, the team can prevent wasting effort t on low-probability leads and instead focus on allocating resources to those that have the highest potential for enrollment.

6. Conclusion & Future Scope

This case study highlights the use of AI-driven lead scoring in EdTech using logistic regression, achieving strong predictive performance (AUC ~ 0.91), this model is used for classification model & it performs well on binary classification. Prioritizing the top 35% of leads led to an estimated 80% enrollment rate, showcasing the model's effectiveness. Integrating this system into a CRM can help sales teams focus on high-potential leads in real time, improving conversion rates.

In the future, the integration of this system into a full-scale CRM (Customer Relationship Management) platform can be further enhanced with real-time analytics, AI-driven lead nurturing, and personalized communication workflows. Advanced machine learning models can continuously learn from new data to improve lead scoring accuracy. Additionally, integrating with marketing tools can allow sales teams to engage high-potential leads across multiple platforms, increasing overall conversion rates. Scalability to support larger datasets and cross-functional team collaboration will also open up opportunities for broader adoption across various industries beyond Ed.Tech.

The model can be enhanced with real-time data feeds, integrated with marketing automation tools, and adapted using deep learning techniques for more complex behavior patterns. Expanding to multi-channel engagement and personalizing outreach based on lead behavior could further boost enrollment efficiency.

References

1. Syam, N. and Sharma, A., (2018). Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, 69, pp.135-146.
2. McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. *Harvard business review*, 90(10), 60-68.
3. Brynjolfsson, E., & McElheran, K. (2016). The rapid adoption of data-driven decision-making. *American Economic Review*, 106(5), 133-39.
4. Sheth J. N., Parvatiyar A., Sinha M., (2015). The conceptual foundations of relationship marketing: Review and synthesis. *Journal of economic sociology*, 16(2), 119-149.
5. Leeflang, P. S., Verhoef, P. C., Dahlström, P., & Freundt, T. (2014). Challenges and solutions for marketing in a digital era. *European management journal*, 32(1), 1-12.
6. [6] Chorianopoulos, A. (2016). *Effective CRM using predictive analytics*. John Wiley & Sons.
7. Duncan, B. A., & Elkan, C. P. (2015, August). Probabilistic modeling of a sales funnel to prioritize leads. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1751-1758). ACM.
8. Järvinen, J., & Taiminen, H. (2016). Harnessing marketing automation for B2B content marketing. *Industrial Marketing Management*, 54, 164-175.
9. Marion, G. (2016). Lead Scoring is Broken. Here's What to Do Instead. URL: <https://medium.com/marketing-on-autopilot/lead-scoring-is-broken-here-s-what-to-do-instead-194a0696b8a3> (Retrieved 24.09.2018)
10. Bohlin, E. (2017). Sorting Through the Scoring Mess. URL: <https://www.siriusdecisions.com/blog/sorting-through-the-scoring-mess> (Retrieved 24.09.2018)
11. Benhaddou, Y., & Leray, P. (2017, October). Customer Relationship Management and Small Data—Application of Bayesian Network Elicitation Techniques for Building a Lead Scoring Model. In *Computer Systems and Applications (AICCSA), 2017 IEEE/ACS 14th International Conference on* (pp. 251-255). IEEE.
12. Michiels, I. (2008). *Lead Prioritization and Scoring: The Path to Higher Conversion*. Aberdeen Group.
13. Artun, O., & Levin, D. (2015). *Predictive marketing: Easy ways every marketer can use customer analytics and big data*. John Wiley & Sons.
14. Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26). New York: Springer.

15. Adam, M.B. (2018). Improving complex sale cycles and performance by using machine learning and predictive analytics to understand the customer journey (Doctoral dissertation, Massachusetts Institute of Technology).
16. Ngai, E. W., Xiu, L., & Chau, D. C. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert systems with applications*, 36(2), 2592-2602.
17. Wouter, B., & Van den Poel, D. (2005). Customer base analysis: Partial defection of behaviorally-loyal clients in a non-contractual FMCG retail setting. *European Journal of Operational Research*, 164(1), 252-268.
18. Kim, Y., & Street, W. N. (2004). An intelligent system for customer targeting: a data mining approach. *Decision Support Systems*, 37(2), 215-228.
19. Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. *MIS quarterly*, 553 572.
20. Karim, M., & Rahman, R. M. (2013). Decision tree and naive bayes algorithm for classification and generation of actionable knowledge for direct marketing. *Journal of Software Engineering and Applications*, 6(04), 196.
21. Larivière, B., & Van den Poel, D. (2005). Predicting customer retention and profitability by using random forests and regression forests techniques. *Expert Systems with Applications*, 29(2), 472-484.

Enhancing Dynamic Resource Management via CNN-LSTM based Proactive Auto-Scaling in Edge Environments

Bablu Kumar¹, Anshul Verma^{*1}, and Pradeepika Verma²

¹ Department of computer science, Banaras Hindu University, Varanasi - 221005, India,

bablu.kumar@bhu.ac.in, anshul.verma@bhu.ac.in

² Technology Innovation Hub, Indian Institute of Technology, Patna - 801106, India,

faculty.tih@iitp.ac.in

Abstract. The adoption of Edge Computing is on the rise, bringing forth considerable challenges for organizations in effectively managing computational loads across distributed and resource-constrained nodes. Traditional containerization methods typically depend on static thresholds, where an application replica is maintained and scaled in response to resource usage. In Kubernetes-based edge environments, where workloads change unpredictably, reactive auto scaling methods often lead to inefficient resource usage due to sudden spikes, increased latency, and degraded service quality. This paper introduces a proactive auto scaling mechanism that utilizes a hybrid CNN-LSTM model. This model is compared to ARIMA, LSTM, and Bi-LSTM for forecasting future loads and pre-emptively scaling resources. The proposed method shows enhanced prediction accuracy, better management of workload spikes, and improved stability in real-time edge environments. By automating the scaling process based on anticipated demand, the system offers more efficient and reliable resource management compared to traditional reactive approaches.

Keywords: CNN-LSTM Forecasting, Container Orchestration, Edge Computing, Proactive Auto scaling, Kubernetes resource management

1. Introduction

The rapid advancement and widespread adoption of Edge Computing have made it increasingly critical to manage computational workloads efficiently across geographically distributed and resource-constrained nodes [1]. Unlike centralized cloud infrastructures, edge environments rely on localized processing, which introduces unique challenges in load balancing, resource provisioning, and latency control [2]. One of the primary challenges is managing workloads that fluctuate in real-time, especially as edge devices support diverse applications such as video streaming, sensor data analysis, and autonomous systems. To address these demands, containerization technologies, particularly Docker, combined with orchestration platforms like Kubernetes, are widely used in edge environments [3]. These systems operate on a master-worker node architecture, allowing containerized applications to be deployed, scaled, and managed effectively [4] and [5]. However, most existing auto-scaling techniques in these environments are based on static thresholds [6]. This reactive approach, which adjusts resources based on current CPU or memory utilization, often fails to respond quickly to sudden spikes or drops in demand. As a result, it can lead to suboptimal resource usage, increased response times, or even service degradation.

To overcome these limitations, proactive auto-scaling mechanisms have been proposed [7]. These techniques aim to predict future workload trends and make scaling decisions in advance, ensuring that the system adapts dynamically without over- or under-provisioning resources. By integrating forecasting models, proactive approaches enable intelligent resource allocation and enhance the stability and responsiveness of edge computing systems. This research introduces a proactive auto-scaling framework that leverages advanced forecasting models, specifically a hybrid CNN+LSTM model [8]. The performance of this framework is compared with traditional models such as ARIMA [9], LSTM [10], and Bi-LSTM [10]. These models analyze historical workload data to anticipate future demands, enabling timely and accurate scaling

decisions. The experimental results indicate that the proposed CNN+LSTM-based method significantly outperforms reactive auto-scaling strategies. It offers enhanced prediction accuracy, a faster response to workload fluctuations, and improved resource utilization in Kubernetes-managed edge environments. By utilizing this proactive approach, the framework ensures that resources are allocated efficiently, thereby enhancing the overall performance and reliability of edge computing systems. This paper contributes to the field by:

- This paper introduces a proactive auto-scaling framework that utilizes a hybrid CNN-LSTM model to predict future computational loads in edge computing environments, ensuring more efficient resource allocation compared to traditional reactive methods.
- The proposed CNN-LSTM model is evaluated against established forecasting models such as ARIMA, LSTM, and Bi-LSTM, demonstrating superior prediction accuracy and responsiveness to workload fluctuations in dynamic edge environments.
- The proposed approach automates the scaling process based on forecasted demand, improving resource utilization by scaling pods according to increasing or decreasing load, and outperforming reactive auto-scaling techniques.

The paper is organized as follows: Section 2 reviews related work. Section 3 outlines the system architecture, loss metrics, algorithms, and methodology. Section 4 presents the evaluation results. Section 5 concludes the paper and discusses future work.

2. Related Work

Efficient resource management in edge computing has attracted significant research attention due to the challenges presented by geographically distributed and resource-constrained environments. Traditional auto-scaling mechanisms, primarily utilized in cloud infrastructures, typically rely on static threshold-based strategies. These methods trigger scaling actions based on real-time CPU or memory utilization metrics, as demonstrated in the works of [11], and [12]. While effective in certain scenarios, such reactive approaches often struggle to respond promptly to sudden workload spikes or drops, leading to inefficiencies in resource utilization and degraded quality of service. To address these limitations, researchers have explored predictive and proactive scaling methods. Forecasting-based techniques aim to anticipate workload fluctuations and allocate resources in advance, thereby minimizing latency and avoiding under- or over-provisioning. For instance, ARIMA-based models have been applied to capture linear patterns in time-series resource data [9], while LSTM and Bi-LSTM networks have shown improved performance in modeling nonlinear relationships and long-term dependencies [13] and [10]. These advancements highlight the potential of proactive resource management strategies in enhancing the efficiency and reliability of edge computing systems.

Recent studies have demonstrated that hybrid models can significantly enhance forecasting accuracy by combining the strengths of different architectures. Convolutional Neural Networks (CNNs), recognized for their ability to extract spatial features, have been successfully integrated with LSTM networks to improve temporal pattern recognition in resource prediction tasks [8]. However, the application of these hybrid models in edge computing environments remains limited, particularly in dynamic and latency-sensitive contexts managed by Kubernetes.

Additionally, several works have implemented auto-scaling in Kubernetes-based edge infrastructures using metrics such as CPU utilization, memory consumption, or custom-defined thresholds [14]. While these implementations are functional, they still rely heavily on reactive mechanisms without incorporating predictive intelligence, which diminishes their effectiveness in real-time edge scenarios. In contrast to these approaches, this paper proposes a proactive auto-scaling framework that leverages a hybrid CNN-LSTM model to accurately forecast future workloads. By comparing its performance against traditional forecasting models

such as ARIMA, LSTM, and Bi-LSTM, this research highlights the effectiveness of hybrid deep learning architectures in enhancing the scalability, responsiveness, and resource efficiency of edge computing systems. This proactive approach aims to ensure optimal resource allocation by scaling pods according to increasing or decreasing load, thereby outperforming reactive auto-scaling techniques in edge environments.

3. System Model and Methodology

This section outlines the system model and methodology used in this study. It is divided into three parts: the first describes the architecture of the auto-scaling approach in cloud and edge environments; the second explains the CNN-LSTM model for edge-based auto-scaling; and the third introduces the proactive adaptation auto-scaling algorithm applied in this work

3.1 Architecture of the auto-scaling approach using cloud and edge environments

The architecture of the proposed proactive auto-scaling framework for Edge Computing is designed to tackle the challenges of dynamically managing computational loads across distributed and resource-constrained nodes. The primary goal is to enhance resource utilization efficiency and minimize latency by predicting future workloads using advanced forecasting models, specifically a hybrid CNN-LSTM model. This architecture leverages containerization technologies like Docker, combined with orchestration platforms such as Kubernetes, to effectively manage and scale applications at the edge.

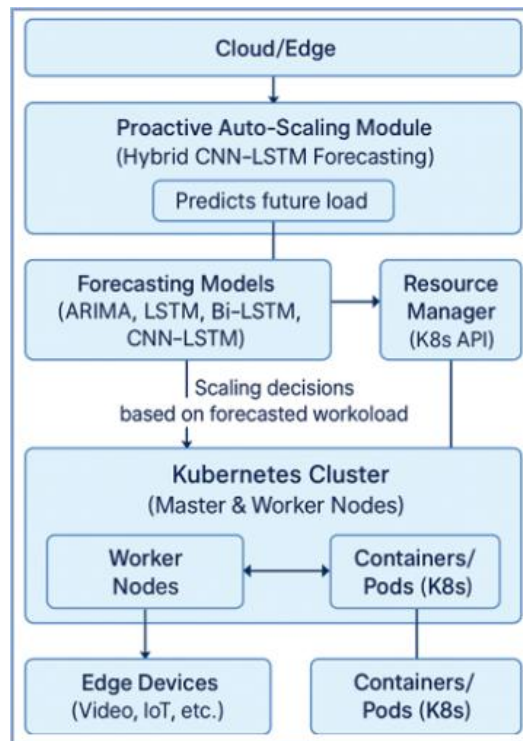


Fig. 1: Working of the proposed auto-scaling approach across cloud and edge environments

The architecture consists of the following components, as shown in Fig. 1: The proposed architecture begins at the Cloud/Edge layer, where centralized cloud servers and decentralized edge nodes are positioned close to data sources such as IoT sensors and video cameras. This proximity minimizes latency and enhances real time responsiveness, forming the foundation for a proactive auto scaling framework. At the core of this system is the Proactive Auto Scaling Module, powered by a hybrid CNN LSTM model that forecasts future workloads by analyzing historical data on CPU, memory, and network usage. This predictive approach enables the

system to allocate resources before demand peaks, unlike reactive methods that respond after strain is detected. The module incorporates a variety of forecasting models, ARIMA for stable trends, LSTM and Bi LSTM for temporal patterns, and CNN LSTM for capturing both spatial and temporal dependencies, resulting in improved prediction accuracy. Once predictions are generated, they are interpreted by a Resource Manager that uses the Kubernetes API to issue real time scaling commands. These commands dynamically adjust the number of containers or pods across a Kubernetes cluster, where the master node manages orchestration and worker nodes execute the containerized workloads. The worker nodes also interface with edge devices, processing incoming data streams through scalable applications. Although not shown explicitly, the system features a feedback loop where performance metrics are continuously monitored and used to refine forecasting models, enhancing adaptability and long term efficiency.

This proactive auto-scaling architecture enables efficient, responsive, and intelligent resource management in edge computing environments. By combining container orchestration with predictive analytics, it addresses both the limitations of reactive scaling and the complexity of edge workload dynamics.

3.2 Working of the Proposed CNN-LSTM Model in Edge-Based Auto-Scaling

The CNN-LSTM model is particularly well-suited for predicting resource demands in edge computing environments managed by Kubernetes due to its ability to capture both spatial and temporal patterns within time series performance data. The process begins with data preprocessing, where a dataset [15] such as the kubernetes performance metrics dataset.csv (available at Kaggle) provides detailed metrics, including CPU usage memory consumption, over time. These metrics form sequential data that is ideal for forecasting models designed to understand patterns and trends. The CNN layer serves as the first step in the model, focusing on extracting spatial features from the time series data. It processes the data in windows or subsequences, allowing it to detect local patterns such as sudden spikes or drops in resource usage that may be transient but critical for scaling decisions. This spatial feature extraction highlights relationships within small segments of data that traditional models might miss, making the CNN effective at capturing subtle but important variations in workload. Following the CNN layer, the extracted features are fed into the LSTM layer, which specializes in modeling temporal dependencies. The LSTM analyzes how previous values in the sequence influence future behavior, which is essential for accurately predicting upcoming workload demands. Since edge computing workloads are often highly dynamic and fluctuate based on real-world events, the LSTM's capability to remember and weigh long-term dependencies enables the system to anticipate these changes effectively. The hybrid CNN-LSTM architecture thus combines the strengths of both models: the CNN's ability to detect localized spatial features and the LSTM's power in understanding temporal sequences. The convolution operations generate feature maps from raw input data, which are then sequenced through the LSTM network to produce accurate workload forecasts. This combined approach allows Kubernetes-managed edge environments to implement proactive auto-scaling, dynamically adjusting resource allocation based on predicted demands rather than reacting to current loads. This proactive strategy improves efficiency, reduces latency, and ensures stable service delivery, ultimately enhancing the performance of edge computing systems.

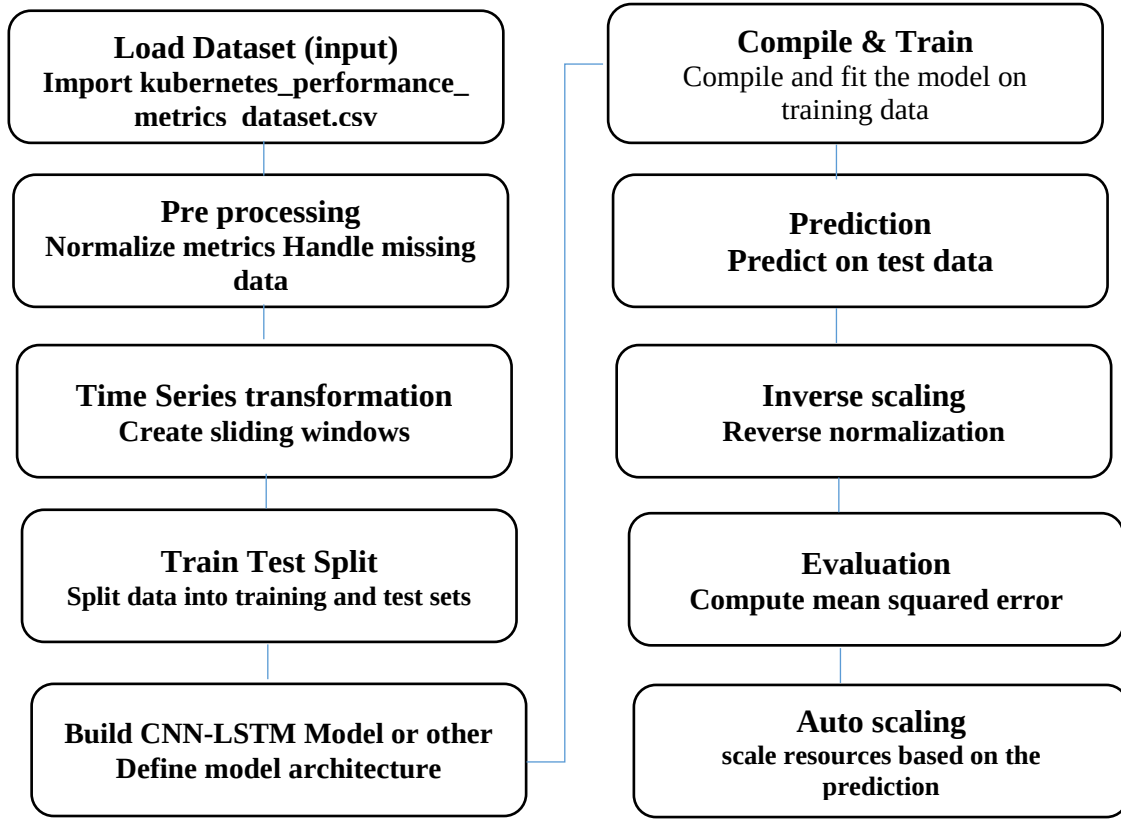


Fig. 2: Flow of execution of this proposed work

The proposed methodology for proactive auto-scaling, as shown in Fig. 2, begins with loading and preprocessing the Kubernetes performance dataset, which includes key metrics such as CPU and memory usage. The data is normalized using MinMaxScaler, which scales values between 0 and 1 to standardize the range across different parameters, ensuring consistency and addressing any missing or anomalous entries. Next, the dataset is transformed into a time-series format using a sliding window technique, enabling the model to learn temporal patterns. This formatted data is then split into training and testing sets, typically using an 80/20 ratio, to facilitate model evaluation. The CNN-LSTM model is constructed by combining a Convolutional Neural Network (CNN) for feature extraction and a Long Short-Term Memory (LSTM) network for learning temporal dependencies, followed by dropout and dense layers to produce the output predictions. The model is compiled using the Adam optimizer and trained with the Mean Squared Error (MSE) loss function, which quantifies the deviation between the actual and predicted values by squaring the differences and averaging them, as shown in Eq. 1:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (1)$$

In this case, Y_i signifies the actual measurement, for instance i in the dataset, whereas $\hat{Y}_i$ indicates the associated predicted outcome.

After training, predictions are made on the test data, and inverse scaling is applied to restore the values to their original units. The model's performance is evaluated using MSE. Finally, the predictions are translated into auto-scaling actions—scaling up when predicted resource usage is high or scaling down when the demand is low—ensuring efficient and adaptive resource management using the algorithm described in 1 and 2.

3.3 Proactive Adaptation Autoscaling Service Algorithms

The adaptation manager service (Algorithm 2) calculates the optimal number of pods for efficient resource allocation based on the predicted workload of incoming requests. The primary goal of the service is to execute proactive scaling operations, which involve adding (scaling out) or removing (scaling in) pods in response to anticipated future workloads. Algorithm 1 details our approach to dynamically scaling the system based on projected workloads, denoted as $Workload_{t+1}$. To address challenges such as quick variations causing leading to inefficient resource utilization and growing costs, a Container Deployment Target (CDT) of seconds is implemented after each scaling decision (Algorithm 2).

Algorithm 1 Adaptation Manager Service Algorithm

Require: $Workload_{t+1}$ // Predicted workload in the next iteration

```

1: while system is running do
2:   for each CDT (Container Deployment Target) do
3:      $pod_{st+1} = \frac{workload_{t+1}}{workload_{pod}}$ 
4:     if  $pod_{st+1} > pod_{st}$  then
5:       // Scaling out condition
6:        $pod_{st+1} = \min(pod_{st+1}, pod_{smax})$ 
7:        $pod_{surplus} = (pod_{st+1} - pod_{st}) \times RRS$ 
8:        $pod_{st+1} = pod_{st+1} + pod_{surplus}$ 
9:       EXECUTE SCALE OUT( $pod_{st+1}$ )
10:    else if  $pod_{st+1} < pod_{st}$  then
11:      // Scaling in condition
12:       $pod_{st+1} = \max(pod_{st+1}, pod_{smin})$ 
13:       $pod_{surplus} = (pod_{st} - pod_{st+1}) \times RRS$ 
14:       $pod_{st+1} = pod_{st+1} + pod_{surplus}$ 
15:      EXECUTE SCALE IN( $pod_{st+1}$ )
16:    else
17:      // No change in workload, do nothing
18:       $pod_{st+1} = pod_{st}$ 
19:    end if
20:  end for
21: end while

```

The system predicts the required pods for the next time step, pod_{st+1} , and compares it to the current count, pod_{st} . If demand increases, pods are added; if it decreases, excess pods are removed using the Resource Removal Strategy (RRS), which enforces a 60-second delay. This ensures efficient scaling, system stability, and responsiveness to workload changes. To prevent pod_{st+1} from falling below the minimum threshold pod_{smin} , the max function is used to ensure

the pod count is at least pod_{min} . The Resource Removal Strategy (RRS), defined as a real number $0 \leq RRS \leq 1$, controls surplus pod removal. If pod_{st} exceeds pod_{st+1} , surplus pods are removed gradually based on the difference, as shown in Eq. 2.

$$pod_{surplus} = \begin{cases} (pod_s_t - pod_{s_{t+1}}) \times RRS, & \text{if } pod_s_t > pod_{s_{t+1}} \\ (pod_{s_{t+1}} - pod_s_t) \times RRS, & \text{if } pod_{s_{t+1}} > pod_s_t \end{cases} \quad (2)$$

Subsequently, pod_{st+1} is adjusted by subtracting $pod_{surplus}$, derived from pod_{st} . Instead of removing all surplus at once, a portion is scaled in using the Resource Removal Strategy (RRS), promoting stability and efficient resource use. The final pod_{st+1} value is then used for the scale-in operation. RRS offers two key benefits: it maintains stability during low loads and adapts quickly to sudden spikes. In our setup, RRS is set to 0.60 (60%), with a minimum pod threshold enforced per minute to ensure availability (Algorithm 1).

Algorithm 2 Pods Requirement Management Service

Require: $numberOfRequests_{t+1}$

1: while the system is operational do

2: for each 60 seconds do

3: $pod_{s_{t+1}} = \frac{\text{number of requests}_{t+1}}{\text{number of Requests}_{pod}}$

4: Execute the Scale command with $pod_{s_{t+1}}$

5: end for

6: end while

The approach is tailored for proactive autoscaling, leveraging predicted metrics and workloads to enable adjustments

as autoscaling that anticipate future changes in cloud-native applications.

4. Results

This section presents the results obtained from the implementation and evaluation of the proposed auto-scaling models in cloud and edge environments. The evaluation metric used to assess the performance of the models is

the Mean Squared Error (MSE), as presented in Table 1.

Table 1: Mean Squared Error (MSE) for CPU and Memory Usage

Metric	ARIMA	LSTM	BiLSTM	CNN+LSTM
CPU Usage	891.60	849.54	840.45	839.50
Memory Usage	855.26	853.14	850.78	849.02

The results indicate that our proposed hybrid model, CNN+LSTM, achieves the lowest MSE values for both CPU and memory usage—839.50 for CPU and 849.02 for memory. These values are significantly lower than those of ARIMA, LSTM, and BiLSTM, demonstrating the superior prediction accuracy of the CNN+LSTM model. To prepare the dataset for model training, CPU and memory usage data were normalized. The corresponding autoscaling decisions based on normalized metrics are illustrated in Fig. 3, where the Y-axis represents normalized usage and the X-axis represents time steps. This visualization shows how CPU and memory usage vary over time in response to load requests. Notably, the system behavior is depicted without the use of auto-scaling, highlighting the natural load-response relationship within the normalized dataset.

Next, Fig. 4 presents a comparison between the predicted load and the actual load, along with the corresponding scaling decisions. The figure illustrates the Y-axis as the load values and the X-axis as the time steps. A static threshold of 70 was used to simulate reactive auto-scaling. Compared to this static threshold, our proposed CNN+LSTM model provides more accurate predictions, resulting in more adaptive scaling decisions. This performance confirms the effectiveness of CNN+LSTM in proactive auto-scaling scenarios.

Additionally, Fig. 5 shows the number of pods scaled using our proactive auto-scaling algorithm compared to a static reactive scaling strategy. The figure presents time on the X-axis (from 1 to 8) and the number of pods on the Y-axis, illustrating the autoscaling performance based on scaled pods in reactive and proactive approaches. The results clearly show that the proactive approach requires fewer pod scale-ups over time compared to the reactive method, indicating more efficient resource utilization.

Overall, the proposed hybrid CNN+LSTM model demonstrates superior prediction accuracy and improved auto-scaling behavior, validating its effectiveness for proactive resource management in cloud and edge environments

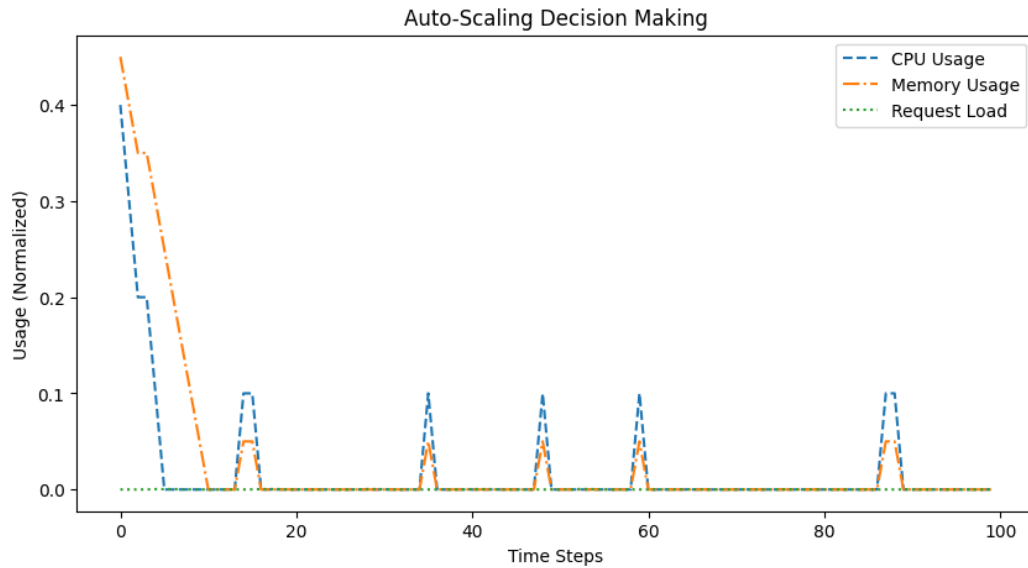


Fig. 3: Auto-scaling decisions based on normalized usage metrics

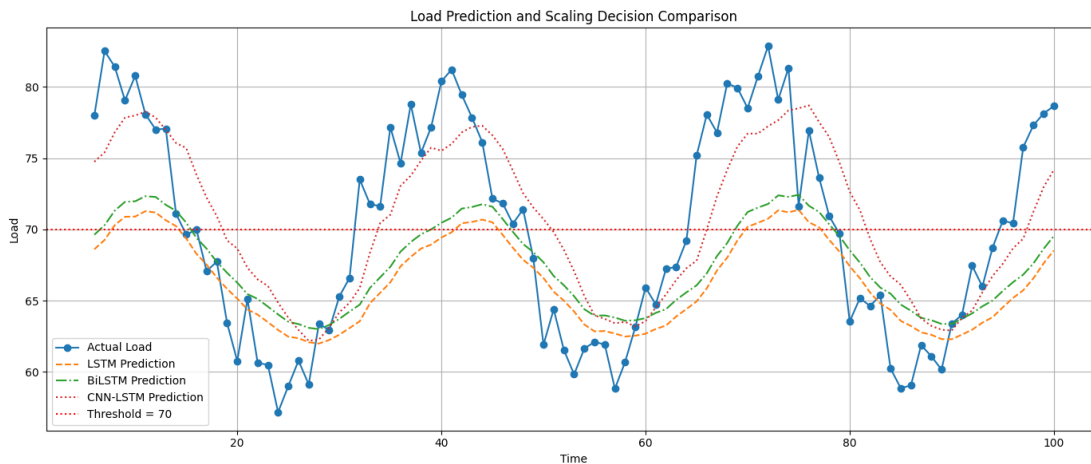


Fig. 4: Comparison of workload prediction and corresponding scaling decisions

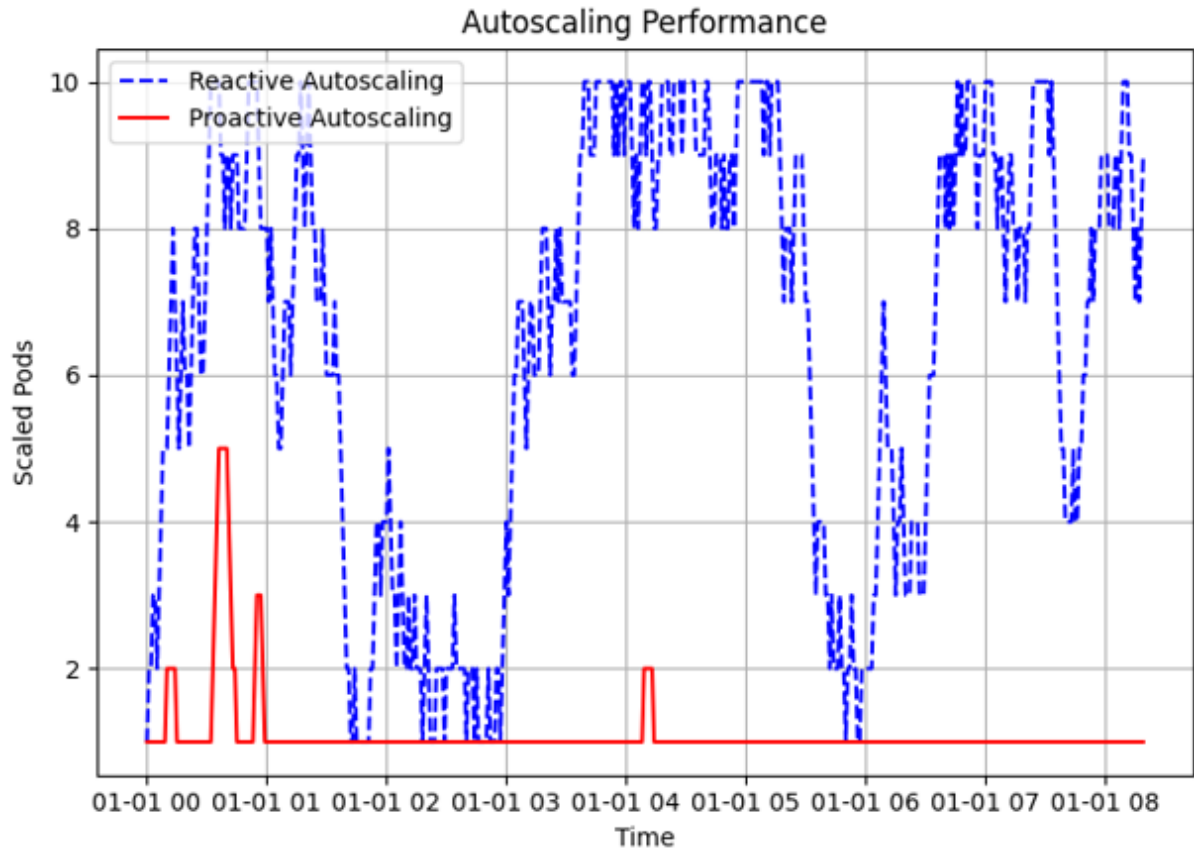


Fig. 5: Autoscaling performance based on scaled pods in reactive and proactive approaches

5. Conclusion

This paper tackles the challenge of efficient resource management in Kubernetes-based edge computing, where workload variability demands smarter autoscaling. Traditional reactive methods, relying on static thresholds, often lead to poor resource use and service degradation. We propose a proactive autoscaling approach using a hybrid CNN-LSTM model to forecast future loads. Compared to ARIMA, LSTM, and Bi-LSTM, our model shows superior accuracy and stability. This work advances edge computing by offering an intelligent, scalable autoscaling solution suited to dynamic, distributed workloads.

Cloud load in edge environments is typically short-term, making it unsuitable for long-term forecasting. However, it can serve as input for other models designed for long sequences and complex multivariate datasets. Combining proactive and reactive scaling in edge environments can further enhance efficiency and responsiveness.

Funding: This research work is funded by the Institutions of Eminence (IoE) Scheme of Banaras Hindu University, Varanasi, India under Dev. Scheme No. 6031.

Data availability: This study uses publicly available data from Nick Kinyae's dataset, Kubernetes Resource Performance Metrics Allocation, accessible at <https://www.kaggle.com/datasets/nickkinyae/kubernetes-resourceand-performancemetricsallocation>

References

1. Pavel Mach and Zdenek Becvar. Edge computing: A survey on architecture and future challenges. Computer Networks,

191:108040, 2021.

2. Bablu Kumar, Mohini Singh, Anshul Verma, and Pradeepika Verma. Optimal cloudlet selection in edge computing for resource allocation. *SN Computer Science*, 4(6):745, 2023.
3. Chen Hong et al. Kubeedge: A platform for edge computing applications. *IEEE Internet of Things Journal*, 9(5):3773–3785, 2022.
4. Bablu Kumar, Anshul Verma, and Pradeepika Verma. Optimizing resource allocation using proactive scaling with predictive models and custom resources. *Computers and Electrical Engineering*, 118:109419, 2024.
5. Bablu Kumar, Anshul Verma, and Pradeepika Verma. A multivariate transformer-based monitor-analyze-plan-execute (mape) autoscaling framework for dynamic resource allocation in cloud environment. *Computing*, 107(3):69, 2025.
6. Sungwoo Lee and Young Kim. Adaptive threshold-based auto-scaling for containerized applications in cloud-edge environments. *Journal of Systems and Software*, 178:110997, 2021.
7. Muhammad Gohar and Muhammad Yousaf. Proactive resource auto-scaling in edge computing using workload prediction models. *Future Generation Computer Systems*, 128:142–152, 2022.
8. Ming Li and Rui Zhao. Hybrid cnn-lstm model for efficient workload prediction in edge computing. *Neural Computing and Applications*, 35:12345–12356, 2023.
9. Yang Liu and Ming Zhao. Arima-based workload prediction for dynamic resource provisioning in edge computing. *Future Generation Computer Systems*, 118:1–10, 2021.
10. Wei Zhang and Qiang Liu. Bi-lstm neural networks for proactive resource scaling in edge-cloud environments. *Journal of Systems Architecture*, 135:102–112, 2023.
11. Li Wang, Wei Zhang, and Ming Chen. A survey on resource allocation and auto-scaling strategies for cloud-native applications. *Journal of Cloud Computing*, 10(1):1–15, 2021.
12. John Smith and Alice Lee. Adaptive threshold-based auto-scaling for microservices in cloud environments. In *Proceedings of the 2022 IEEE International Conference on Cloud Computing*, pages 100–107. IEEE, 2022.
13. Li Chen and Hao Wang. Lstm-based predictive auto-scaling for containerized applications in edge computing. *IEEE Transactions on Cloud Computing*, 10(2):345–356, 2022.
14. Li Ju, Prashant Singh, and Salman Toor. Proactive autoscaling for edge computing systems with kubernetes. *arXiv preprint arXiv:2112.10127*, 2021.
15. Nick Kinyae. Kubernetes resource performancemetricsallocation. <https://www.kaggle.com/datasets/nickkinyae/kubernetes-resource-and-performancemetricsallocation>, 2023. Accessed: 2025-05-15

Synthesizing Virtual Datasets for Personalized Tourism and Insurance Services

Hwanseog Kim¹, Jundong Lee¹, Ryumduck Oh², and Sangmin Suh¹

¹ Gangneung-Wonju National University, ROK

² Korea National University of Transportation, Dept of Computer Software, ROK
sangminsuh@gwnu.ac.kr

Abstract :

This study aims to develop a synthetic dataset that replicates the structure and appearance of real driver's licenses for use in personalized tourism and insurance services, while avoiding the legal constraints of personal data protection laws. Utilizing a Python-based environment (WSL2, Faker, etc.), key identity attributes—including license number, name, resident registration number, address, issuance date, and aptitude examination period—were randomly generated and composited with facial images from the Labeled Faces in the Wild (LFW) dataset onto a license template. Experimental results demonstrate that the generated images closely resemble real licenses in appearance and contain essential identity fields suitable for OCR-based information extraction. The proposed dataset provides a practical and legally compliant resource for various AI applications, free from ethical concerns associated with real personal data.

Keywords: Information Extraction, Insurance Service, Privacy Protection, Virtual datasets.

1 Introduction

In recent years, the growing societal emphasis on work-life balance in Korea has led to a continuous increase in demand for personalized tourism services. Consequently, the development of diverse, user-tailored tourism solutions has become imperative, necessitating the reliable collection, analysis, and structuring of personal data. However, the revised Personal Information Protection Act imposes strict regulations on the processing of sensitive and uniquely identifiable information, thereby presenting significant limitations in directly utilizing personal data from documents such as real driver's licenses.

The global personalized travel market is projected to reach approximately USD 553 million in 2024 and is expected to grow to USD 13.26 billion by 2033. As illustrated in Fig. 1, the market is anticipated to experience a compound annual growth rate (CAGR) of approximately 10.5% during the forecast period [1].

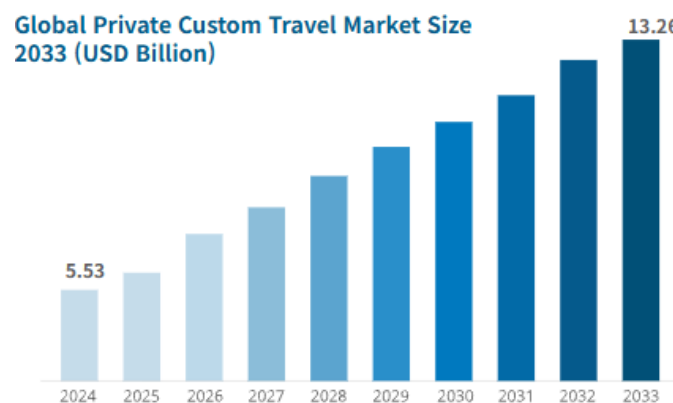


Fig.1. Global Private Custom Travel Market Size [1]

Personal identification information, such as that found on driver's licenses, is essential for personalized services in domains including tourism, finance, and public administration, enabling identity verification, age authentication, and qualification validation. However, the

revised Personal Information Protection Act, enforced on March 15, 2024, imposes stringent regulations on the collection, use, storage, provision, processing, and outsourcing of personal data [2]. Consequently, publicly available datasets for such applications are virtually nonexistent, and collecting real driver's license data from individuals is practically infeasible.

To address these challenges, this study constructs a novel dataset comprising synthetic driver's license images that emulate the visual and structural attributes of authentic licenses while replacing all personal information with randomly generated data. As the dataset contains no real personal data, it can be freely utilized without legal or ethical constraints. Moreover, the synthesized licenses retain key fields—such as name, date of birth, and address—allowing for automated information extraction and enabling downstream applications such as personalized travel recommendations and insurance plan generation.

2. Proposed Method and Results

2.1. Method

In order to synthesize realistic virtual driver's licenses, a thorough understanding of the structure and content of authentic licenses is essential. In this study, accurate reference data were obtained from the Korea Road Traffic Authority [3] and were used to generate the synthetic license information. While this research focuses on Korean driver's licenses, the proposed method is generalizable and can be adapted to generate licenses from other countries, provided that sufficient structural and informational details are available.

The facial images used in this study were sourced from the Labeled Faces in the Wild (LFW) dataset [4], with several examples presented in Fig. 2. LFW is a publicly available benchmark dataset designed to support the development and evaluation of robust face recognition models under unconstrained, real-world conditions.



Fig. 2. Adopted facial image data

The driver's license synthesis process consists of three primary functional components. First, a data generation function utilizes the Faker package to produce synthetic personal information. Second, a text preprocessing function automatically wraps long strings into multiple lines based on a predefined maximum character length. Third, a composition function overlays the generated data onto a license template to produce the final synthetic driver's license image.

2.2. Results

Using the proposed method, synthetic driver's license images were generated, with sample outputs presented in Fig. 3.



Fig. 3. Synthetically generated driver's license image

As illustrated, all required information fields were accurately synthesized, and each element was correctly positioned according to the predefined layout specifications.

3. Conclusion

This study proposes the design of a virtual driver's license dataset that preserves the structure of authentic licenses while fundamentally avoiding issues related to personal data protection. The proposed dataset enables applications such as personalized travel recommendations and insurance planning, and facilitates unrestricted research and development involving sensitive information without legal constraints.

Acknowledgement

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP)-Innovative Human Resource Development for Local Intellectualization program grant funded by the Korea government (MSIT)(IITP-2025-RS-2023-00260267)

References

1. Global Private Custom Travel Market Size, <https://www.businessresearchinsights.com/ko/market-reports/private-custom-travel-market-113107>.
2. Personal Information Protection Act, Act No. 14839 (amended July 26, 2017), National Law Information Center (Republic of Korea).
3. Yonhap News Agency. (2025, March 11). Title of the article in English or Romanized Korean. Retrieved May 24, 2025, <https://www.yna.co.kr/view/AKR20250311035900062>.
4. Li, Jessica. "Labeled Faces in the Wild (LFW) Dataset." Kaggle. Accessed May 24, 2025. <https://www.kaggle.com/datasets/jessicali9530/lfw-dataset>.

Dissimilar metal weld inspection using artificial intelligence in non destructive evaluation 4.0

M.Annadurai¹, R.Joselin², S.Anitha³, M.Surya⁴, E.Vidhya⁵ S.Saravanan⁶

¹Associate Professor, Roever Engineering College, Perambalur- 621 220, Tamilnadu, India.

²Professor, Roever Engineering College, Perambalur - 621 220, Tamilnadu, India.

^{3,4,5} Assistant Professor, Roever Engineering College, Perambalur - 621 220, Tamilnadu.

⁶Student, Roever Engineering College, Perambalur - 621 220, Tamilnadu, India.

Corresponding author:Email:joselinjerish75@gmail.com

Abstract :

The use of lightweight materials for industrial application is a driving force for the development of joining techniques. Friction stir welding (FSW) inspired joints of materials because it does not involve bulk melting of the basic components. Here, two Aluminium 7020 plates with Zn as interlayer is welded, using friction stir welding method. The joints is produced using same welding parameters only the variations is done in the interlayer of joints and the characterized regarding mechanical properties, micro structure. Tensile strength of the joint is determined using an online tensile test machine with Acoustic Emission test setup. Impact strength and micro hardness of the weld joint is determined. Microstructure of the joint is analyzed using image analyzer. The main benefit of machine learning powered NDE is its reliability and repeatability to find flaws from the data in quick time. The results of the analysis show that the specimen joint with 20 μ as Zinc interlayer has higher mechanical properties.

Keywords: Friction stir welding, Aluminium 7020, Zinc interlayer, Mechanical Properties, Acoustic Emission, Microstructure

1. Introduction

Friction stir welding (FSW) is a joining process produces no fumes, uses no filler material which make this process more environmental friendly. FSW is a solid state welding process where joining takes place below melting temperature of material this characteristics greatly reduces distortion shrinkage, porosity, eliminates solidification defects. FSW has been proven to be an effective process for welding Aluminium alloys, Titanium, Copper, Zinc, Steel, Magnesium , Cast iron [1-7].Today FSW is used in research and production in many sectors, including aerospace, automotive, railway, ships building, electronic housings, coolers, heat exchangers, and nuclear waste containers. Main advantages of friction welding are high material saving, low production time and possibility of welding of dissimilar metals or alloys. For example engine pistons, mainly for truck applications can be friction welded [8-10]. In this process a rotating tool with a pin extending from a larger shoulder is translated along the weld joint. This tool during its path produces frictional heating with temperature below the melting point and the plastic deformation due to the stirring of the material around the pin, produces the weld. The process can be regarded as solid-phase keyhole welding technique since a hole to accommodate the probe is generated, then moved along the weld. The tool material required for FSW should be of high melting temperature, abrasion resistance and adequate toughness at high temperature. Different tool materials offers specific advantages for different applications. Here, the tool material used is high speed steel (HSS) with triangular pin profile.

The figure 1 shows a schematic illustration representing friction stir welding process. In this FSW process a tool with a rotating pin is entered in to the edges of plates or sheets to be welded. After getting in appropriated position the rotating tool produces heat due to friction between the tool and work pieces hence the joining of the plates is performed.

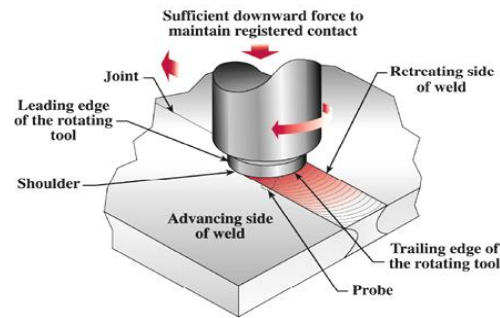


Fig. 1 Schematic illustrations of friction stir welding process

In this work, friction stir welding is carried out between aluminium 7020 plates with Zinc as the interlayer. The joints are made between plates with Zinc as interlayer by varying the gaps as 20 μ , 40 μ , 60 μ , 80 μ , 100 μ respectively by keeping the tool rotation and feed rate as constant parameters.. The microstructure of the joint with different positions of the weld is analyzed and the mechanical properties of the weld specimen are evaluated.

2. Experimental Procedure

Samples (150x50x3.3mm) of Al 7020 are welded in this experiment. The chemical composition and mechanical properties of the material is shown in Table 1 & 2 respectively. Before welding the oxide films on the Al 7020 plates were removed by emery paper, cleaned by cotton waste and applied with silicon oil for preventing from oxide formation. The welding equipment used is friction stir welding machine. FSW has many welding parameters such as tool material, tool rotation speed, tool feed rate and angle of tilt. The stir pin penetrates in to the joints with the full length during welding process.

Table 1 Chemical Composition of Aluminium Alloy 7020

Weight%	Al	Si	Fe	Cu	Mn	Cr
7020	Bal	0.35	0.40	0.20	0.050 – 0.50	0.10 – 0.35
Weight%	Mg	Zn	Zr	Zr + Ti	Other Each	Others Total
7020	1.0 – 1.4	4.0 – 5.0	0.080 – 0.20	0.80 – 0.25	0.05 max	0.15 max

Table 2 Mechanical Properties of Aluminium Alloy 7020

Material	Temp.	Tensile Strength (Mpa)	Yield Strength (Mpa)	Elongation %
Alloy 7020 Plate	T651	330-350	260-280	7-10
Properties dependent on thickness				

Friction stir welding has many parameters, such as tool material, tool rotation speed, welding speed/feed rate of tool and angle of tool. In this work tool speed and feed rate were kept constant and is of 1400 rpm and 32 mm\min respectively. Only the variation is done using the Zinc as interlayer between two Al 7020 plates of about 20 μ , 40 μ , 60 μ , 80 μ and 100 μ respectively. Zinc powder is used as bonding material in the interlayer of Al7020 plates. Since it has low melting point than the Aluminium alloy, it melts and penetrates easily in the gaps of the metal dispersion zone, hence produces good strength weld. The FSW tool with 18mm diameter shoulder probe of 2.9 mm length with triangular profile and the tilt angle of the tool is 2.5° from the normal surface of the plates. Tool used is High Speed Steel (HSS), which is of high thermal and wear resistant. The welding tool is fixed on tool holder and was rotated in clockwise direction from retreating side to forward side of the weld, the plates were tightly fixed on fixture using clamps was traveled. The joining of two pieces is done by frictional heat generated between the tool and work pieces. The welded pieces are made into separate specimen as required dimensions for conducting various tests.

The tensile strength of the joint was determined using a conventional tensile test machine with acoustic emission setup Acoustic Emission setup is used to determine the exact solution by the solving many iterations. The variations AE parameters obtained from the test setup are via, count, duration, hits, energy and event etc. Vickers testing machine using 0.5kg load was applied for micro hardness measurements. Impact strength of the joint is examined using charpy Impact tester. The micro structure of the joint was examined using image analyzer system.

3. Experimental results and discussion

Tensile properties

Tensile strength of the weld joint is tested using universal tensile test machine. The maximum load range of the test machine is 100KN with cross head stroke of 1000mm. the cross head displacement measurement is 0.1mm with cross head speed of 0.5 to 250 mm/min. The FSW samples are tested without any surface modification after welding. The tensile test is carried on the weld specimen with zinc interlayer of 20 μ , 40 μ and 100 μ . The results of the tensile test is shown in the Table 3. The specimen with 20 μ of zinc interlayer has high tensile strength with peak load of 5.875KN, ultimate tensile stress of 122 Mpa, yield stress of 100 Mpa with 4% elongation.

Table 3: Tensile results of friction stir welded joints

Parameters	20 μ (zinc interlayer)	40 μ (zinc interlayer)	100 μ (zinc interlayer)
Peak Load(kN)	5.875	3.235	2.1
UTS(Mpa)	122	67	44
Yield Stress(Mpa)	100	51	30
Elongation(%)	4	3	1.8

Impact Test

Impact strength of the welded specimens are tested using by charpy impact test machine, of pendulum type, with maximum range is of 300 joules. The impact test carried for the specimens of various zinc interlayers. Impact strength of various specimen is shown below in Table 4.

Table 4 Impact strength results of specimens.

Zinc interlayer	Strength(joules)
20μ	2
40μ	1.8
60μ	1.5
80μ	1
100μ	1

From above results, it is proved that the specimen with 20μ of zinc interlayer has maximum impact strength (2 joules). The strength decreases gradually with variations in the zinc interlayer. Hence it is proved specimen with 20μ interlayer has maximum strength.

Micro hardness Test

The hardness of the welded specimen is investigated using Vickers hardness direct mass loading testing machine. The load limit is of 0.1 to 10 Kg of Macro load and 50 to 140 g of Micro load. The intention time taken by the test machine is 10 sec. with field illumination of magnification is 400 x. The load applied for this test is 0.5 Kg and with magnification of 20 x. The hardness test is performed on each specimens in three zones namely stir zone (middle zone of the weld), retreating side and advancing side of weld. Micro hardness test carried on specimens having various zinc interlayers as shown in below the table 5.

Vickers Hardness $HV = a \cdot 0.05 + b \cdot 0.001$ for 20x, $a = 0.05$ and $b = 0.001$

Table 5 Micro hardness results of the welded specimens

Interlayer	Load=0.5 kg	a	b	$HV = \frac{a}{b}$	HV
100μ	0.5	2	6	0.106	82.5
	0.5	2	8	0.108	79.5
	0.5	2	7	0.107	81
40μ	0.5	2	20	0.120	64.4
	0.5	2	17	0.117	67.7
	0.5	2	8	0.108	79.5
20μ	0.5	1	28	0.078	152
	0.5	1	31	0.081	141
	0.5	1	30	0.080	145

The hardness is test done in stir zone or nugget zone which is of 152 HV. The hardness of the retreating side is 145HV and the on advancing side is 141 HV. Hardness test results shows the test specimen of 20μ zinc interlayer has maximum hardness as compared to the other specimens.

Microstructure

Micro structure of the weld specimen is obtained by Clemex Image analyzing system. The micro structure analysis is done in each specimen with various zinc interlayers (20μ, 40μ, 100μ). From the micro analysis the fusion line is clear with stir zone, retreating side of weld and advancing side of weld. The structure closed to the weld metal is zinc layer which grew in to the weld metal. The microstructure near the stir zone is obviously different from the parent metal. Also the images shows the material dispersion between Al and Zinc due to tool rotation. The figure 2 shown below represent the microstructure of parent material (Al 7020).



Fig. 2 Micro image of Al 7020.

The figure 3 shows typical micro images of advancing side, stir zone and retreating side of each welded joints of 20 μ , 40 μ and 100 μ zinc interlayer.

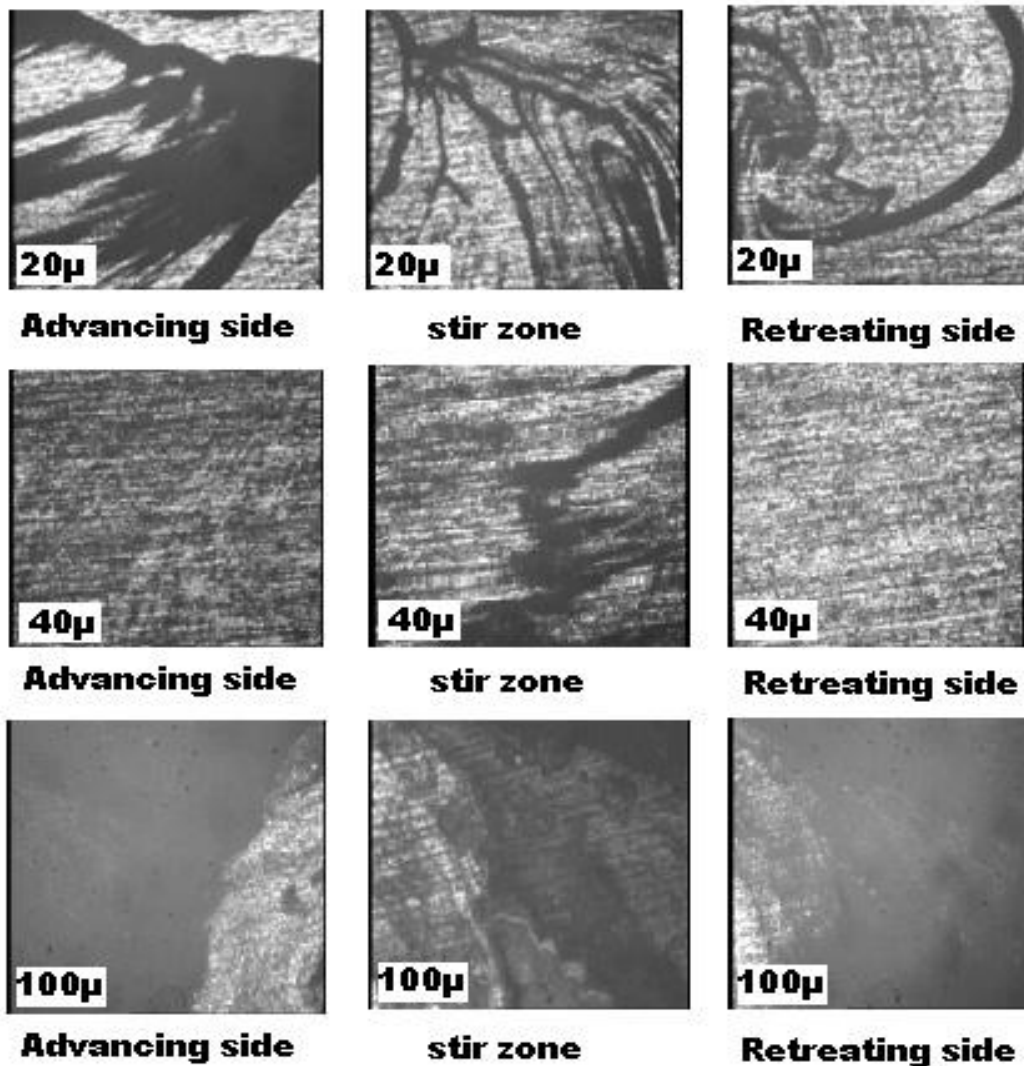


Fig. 3 Micro structural variation with various interlayers

Microstructure of the specimen with 100 μ interlayer shows pores formation due to lack of fusion between the parent metal and the zinc layer. Image recognition can be recognising the welds. In which image quality indicators and other features which the inspectors usually have

to identify by themselves it requires more time for viewing the data that contain no flaws. Machine learning (ML) changes it's considerably. ML model can go through the data multi times and indicates the critical flaws if any by measuring the reliability through Probability of Detection (POD),

4. Dissimilar metal weld inspection

Inspecting dissimilar metal welds (DMWs) using ultrasound presents significant challenges due to high levels of noise and signal attenuation caused by the material's heterogeneous structure. These welds are typically located in safety-critical areas, where inspections must meet stringent reliability standards and maintain a low rate of false indications. Phased array ultrasonic testing (PAUT) is commonly employed in these scenarios, as it allows for the collection of detailed data from multiple angles, improving flaw detection capabilities [11]. However, the increased volume of data can also lead to longer analysis times. To streamline this process and assist inspectors, various image and data processing techniques are often used. One such approach involves combining data from multiple beam angles to create a single, composite ultrasonic image [12]. Despite its advantages, this merging process can introduce visual artifacts or result in information loss, which often necessitates that inspectors also review the unprocessed, raw data.

In this approach, a U-Net-based semantic segmentation model was utilized to identify flaws within ultrasonic inspection data. Beyond simple detection, the model was also capable of pinpointing the precise locations of these flaws. With this positional information, the analysis process could be streamlined by generating an indication list for the inspector to review [13]. Each item on the list corresponded to a flaw highlighted within the pre-processed ultrasonic data, allowing inspectors to quickly assess, measure, and move on to the next indication. By focusing only on the indications identified by the model, evaluation times were significantly reduced [14]. Additionally, the method was easily adaptable for integration into standard inspection software platforms, enabling further analysis when necessary.

5. Conclusion

The present work reveals the mechanical properties and the micro structure of Friction stir welded Al 7020 plates with various dimensions in zinc interlayer. The result of this work predicts weld specimen with Zinc interlayer of 20 microns has higher tensile strength, hardness as compared to 40 μ and 100 μ zinc interlayer weld specimen. From the microstructure analysis it is clear that specimen with 20 μ interlayer has higher material dispersion. The work can be further extended to change the mechanical properties of the specimen. The main benefit of utilising machine learning in NDE is significant increase in efficiency, but also in reliability. Once the data analysis can be automated, it enables monitoring of manufacturing in completely different scale than before, completely unreachable by manual data analysis.

Acknowledgement

The authors would like to sincerely thank the Professor Head Dr. Raghukandan of Annamalai University for providing permission and assistance and also the other staff and technicians of production department for their support.

References

1. Won-Bae Lee, Yun-Mo Yeon and Seung-Boo Jung 2004. “*Mechanical Properties Related to Microstructural variation of 6061 Al Alloy Joints by Friction Stir Welding*”, Materials Transactions, vol 45.No.5 pp. 1700 to 1705
2. L. Ceschini, I.Boromei, G. Minak, A.Morri, F.Tarterini .2007. ”*Effect-of friction stir welding on microstructure, tensile and fatigue properties of the AA7005/10 vol. %AlO_{3p} composite*”, Composites Science and Technology vol. 67, pp605-615.
3. P. Cavaliere, M. Cabibbo, F. Panella, A. Squillace .2009. “*2198 Al-Li plates joined by Friction-Stir Welding: Mechanical and micro structural behavior*”, Materials and Design vol 30, pp 3622-3631.
4. ZHANG Bing-gang, CHEN Guo-qing, ZHANG Chung-guang, NI jia-qiang. 2011. “*Structure and mechanical properties of aluminum alloy/Ag interlayer/steel non-centered electron beam welded joints*”, Transactions of Nonferrous Metals Society of China, vol 21, pp 2592-2596.
5. Radoslaw Winiczenko, Mieczyslaw kaczorowski.2012. “*Friction-welding of ductile cast iron using intelayers*”, Materials and Design vol 34, pp 444-451.
6. J.Guo, P.Gougeon, X.G.Chen. 2012. “*Microstructure evolution and mechanical properties of dissimilar friction stir welded joints between AA1100-B₄C MMC and AA6063 alloy*”, Materials Science and Engineering vol 553, pp 149-156.
7. R.S. Coelho, A.Kostka, J.F.dos Santos, A.Kaysser-Pyzalla. 2012. “*Friction – stir dissimilar welding of alluminium alloy to high strength steels: Mechanical properties and their relation to microstructure*”, Materials Science and Engineering vol 556, pp 175-183.
8. Caroline Jonckheere, Bruno de Meester, Anne Denquin, Aude Simar. 2013. “*Torque, temperature and hardening precipitation evolution in dissimilar friction stir welds between 6061-T6 and 2014-T6 aluminium alloys*”, Journal of Materials Processing Technology vol 213, pp 826-837.
9. Luka Posilovic, Dujie Medak, Fran Milkovi'c Marko Subasic, Marko Budimir, and Sven Lon caric. “*Deep learning based anomaly detection from ultrasonic images*”. Ultrasonics, 124:106737, 2022.
10. Nauman Munir, Hak-Joon Kim, Sung-Jin Song, and Sung-Sik Kang, “*Investigation of deep neural network with drop out for ultrasonic flaw classification in weldments*”. Journal of Mechanical Science and Technology, 32(7):3073-3080, 2018
11. N Munir, HJ Kim, J Park, SJ Song, and Kang., “*Convolutional neural network for ultrasonic weldment flaw classification in noisy conditions*”, Desanics, 2018.
12. Tuomas Koskinen, tikka Virkkunen, Oskar Siljama, Oskari Jessen-Juhler, and Jari Rinta-Aho., “*The effect of different flaw data to machine learning powered ultrasonic inspection*”. Joumal of Nondestructive Evaluation, 40(24), 2021
13. Topias Tyysti arvi, likka Virkkunen, Peter Fridoll, Anders Rosell, and Zuheir Baroum. “*Automated defect detection in digital radiography of aerospace welds using deep learning*”. Welding in the World, 66, 2022.
14. Sergio Cantero-Chinchilla, Paul D. Wilcox, and Anthony J. Croxford. “*A deep learning based methodology for artefact identification and suppression with application to ultrasonic images*”. NDT & E International, 126:102575, 2022.

Enhancing Vehicle Safety: YOLOv8 and YOLOv9 with IR Sensors for Brake Light Detection and Accident Prevention

Harsh Jaiswal and Surendra Nagar

Netaji Shubas University of Technology, Dwarka, New Delhi, India

harsh.jaiswal.pg23@nsut.ac.in

Abstract :

Road accidents are a global challenge, causing significant economic losses, injuries, and fatalities. Addressing this issue, the study introduces an advanced real-time vehicle brake light detection system to prevent rear-end collisions and improve road safety. Utilizing the YOLOv8 and YOLOv9 algorithms for object detection, the system classifies brake lights and integrates infrared (IR) sensors for precise distance measurements. These deep learning models enhance detection accuracy under diverse environmental conditions, outperforming traditional visual based approaches. The research employs a robust dataset of vehicle rear images captured under varied lighting and weather conditions. A hybrid data annotation approach, combining computer vision and manual verification, ensures reliability. Infrared sensors complement the visual data, enabling consistent performance in low-light scenarios while providing cost-effective alternatives to LiDAR. Evaluation metrics, including mean Average Precision (mAP), recall, and frames per second (FPS), are used to assess model efficacy. The proposed system demonstrates potential as a significant enhancement to Advanced Driver Assistance Systems (ADAS), paving the way for safer autonomous driving technologies. By fusing state-of-the-art deep learning algorithms with sensor fusion techniques, this study aspires to reduce rear-end collisions significantly and bolster road safety. The research holds promise for improving the reliability of brake light detection, advancing autonomous vehicle capabilities, and contributing to broader road safety initiatives. Code available at <https://www.github.com/jiisanda/avoiding-road-accidents>.

Keywords: Break Light Detection, You Only Look Once, Infrared (IR) sensors, Sensor Fusion, Deep Learning

1 Introduction

Road safety is a major worldwide concern since rear-end collisions account for a lot of annual accidents. Especially in India, where they account for 25–30% of all accidents and 380 out of 1,000 injuries [14, 2, 18], these mishaps rank as the main cause of death for children and young adults aged 5–29 [26, 16]. Especially in braking behavior, innovative approaches are required to lower these accidents. 2 Harsh Jaiswal et al. With autonomous driving moving beyond level 2 to level 3 [7], braking systems have evolved into quite intelligent systems. This work presents a new method for robust brake light and distance detection, combining modern YOLOv8 and YOLOv9 models with infrared (IR) sensors. Traditional camera-based brake light detectors are sensitive to changes in vehicle design and difficult lighting conditions, often underperforming in adverse weather and low-visibility situations [3, 10]. Combining deep learning architectures with infrared imaging [3, 8] the proposed system seeks to improve detection accuracy, retain performance in low visibility, and provide real-time responsiveness across several driving conditions. Selected YOLOv8 and YOLOv9 are based on their exceptional object detection tasks speed and accuracy [8]. YOLOv8 and YOLOv9 shine in processing speed; depending on the configuration, YOLOv8 can achieve 40-155 FPS, greatly outperforming SSD (22-46 FPS) and Faster R-CNN (5-7 FPS) [9]. Programmable Gradient Information (PGI) that increases accuracy while preserving efficiency [8] helps YOLOv9 to surpass YOLOv8's performance. This combination approach seeks to produce a stronger and more flexible brake light detection

system able to run efficiently over a broad spectrum of driving situations and environmental conditions. The work aims to create a fusion method combining IR sensor data with dual models to raise brake light detecting accuracy. This study seeks to help to lower road traffic deaths and injuries in line with the UN's target of halving these figures by 2030 [16].

2 Related Work

The paper [12] uses YOLOv7 and IR sensors for distance measurement to enhance road safety by detecting brake lights. The YOLOv7 model performs well under various lighting conditions, while integrating IR sensors improves distance estimation accuracy. The study [20] recommends further research on recent YOLO models and vehicle rear light detection at night. Nevertheless, especially during the day or in fog, the suggested study is not very strong in several lighting and environmental conditions. Both studies underline the need of more research and more varied real-world scenarios. Using transfer learning to identify vehicles and classify their brake light status on over 11,000 real-world images, the paper by oh et al. in [15] suggested a one-stage brake light detecting network based on YOLOv8. With a mAP50 of 0.766 to 0.793 the system obtained great detection accuracy. The study does not, however, investigate other deep learning architectures or contrasts with cutting-edge techniques like YOLOv9. The suggested study seeks to raise adaptability and detection accuracy under unfavorable circumstances. The paper [1] presents ideas for improving detection in low-visibility settings using related optical techniques. Using surveillance video, the authors in [19] created an IoT fog-based system to identify dead or reclining cattle on roads, so lowering accident risk. Their YOLOv5s-based model exceeded state-of-the-art object detection techniques by 98% accuracy for road-based cattle detection. In line with brake light detection, the system efficiently found road hazards including cattle. Developing a framework for real-time accident

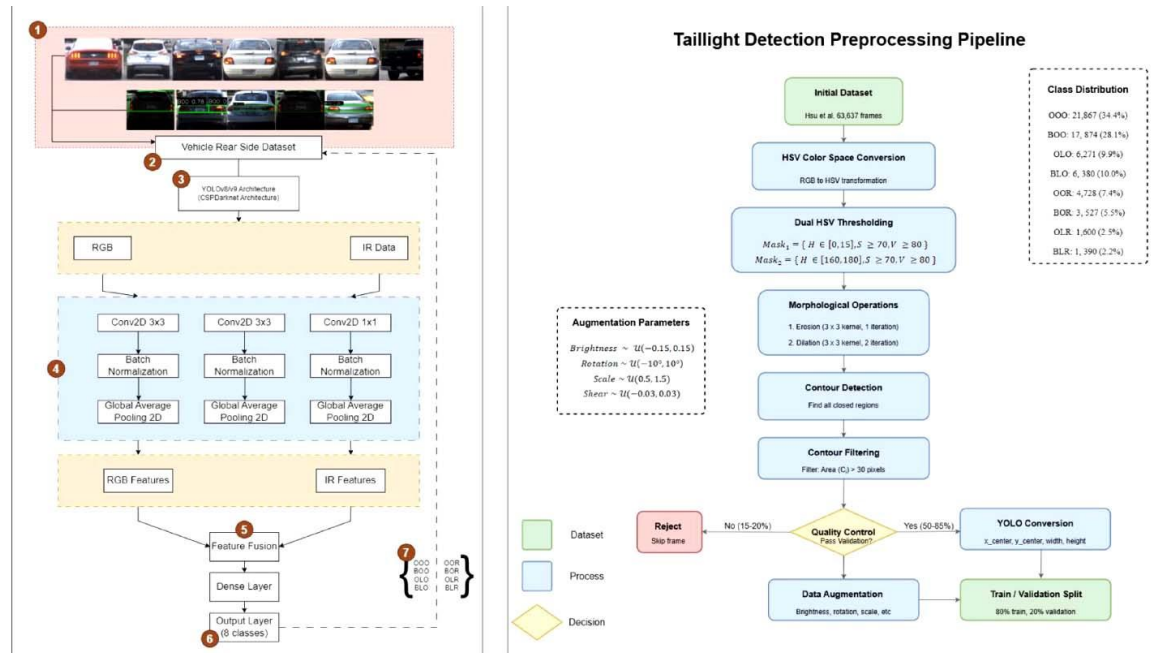


Fig. 1. (a) The multi-modal brake light detection system uses camera-IR fusion, acquisition,

CSP Darknet backbone, feature fusion, classification, and alert decision. (b) The preprocessing pipeline includes original dataset, HSV-optimized taillight detection, morphological enhancement, contour filtering, YOLO-compatible augmentation, and 80:20 train-val split detection at intersections, the authors in [4] identified vehicle-to-vehicle, vehicle to-pedestrian,

vehicle-to-bicycle events. Emphasizing general accident detection, the work shows how to mix object detection, tracking, and trajectory analysis for real-time traffic monitoring. By means of a controller developed in response to an initial rear-end collision, the study [5] extends the operational horizon of vehicle active safety systems by preventing secondary crashes. To reduce initial crash loss, the controller combines post-collision and pre-collision control. The study [11] proposes a technique using YOLOv4, LSTM, and inverse perspective mapping to detect front vehicle brake lights in highway tunnels, achieving a 90.6% accuracy.

3 Method

3.1 System Architecture

The proposed system as seen in Figure 1 (a), uses RGB images and infrared (IR) sensor data to detect brake lights, ensuring strong performance in lowlight, weather-challenged, and variable lighting conditions. RGB cameras record 4 Harsh Jaiswal et al. the rear view, while IR sensors provide proximity and distance data. The system aggregates visual and spatial cues [21, 27] using parallel CNNs, extracting features and categorizing them into eight classes for brake status and proximity alerts. This system improves driver safety by providing timely warnings.

3.2 Dataset and Preprocessing

Using 63,637 rear-view frames across 649 manually cropped daytime driving videos, we leverage the UC Merced Rear Signal dataset by Hsu et al. [6]. Triplets like OOO (no signal), BOO (brake only), BLO (brake + left turn), and BLR (brake + hazard) label every frame from one of eight tails. The dataset is quite imbalanced: states like "BLR", and "OLR", are rare while "OOO" and "BOO" rule. This inspired strong preprocessing techniques as well as class-aware augmentation methods.

Taillight Detection based on HSV: Using two disjoint masks: $H \in [0, 15]$ and $H \in [160, 180]$, each frame is converted to HSV color space and red/amber taillight regions are isolated with $S \geq 70$ and $V \geq 80$ [17]. Under different lighting, these masks record both reddish-orange and deep-red taillight tones. Using 3×3 kernels, morphological operations (erosion followed by dilation) reduce noise and combine fractured components [13, 24]. Contour handling and annotation: Retained detected blobs with area 30 pixels are converted into YOLO format: bounding boxes are normalized as (xc, yc,w, h). We separate each taillight (not full vehicles) to enable fine-grained signal modeling. Every frame consists on average two to four objects. This object-level precision increases detection sensitivity to varying viewpoints and occlusion. Dataset preparation and quality control: Occlusion, extreme glare, or detection failure—that is, bounding boxes outside image bounds—cause about 15-20% of frames to be thrown away. Filtering criteria consist in minimum contour area, spatial validity, and consecutive detection failures. Training and validation sets comprise an 80/20 split from the dataset. Data augmentation: We address imbalance and enhance generalization using class-aware augmentation. This covers brightness/contrast jitter ($\bullet$ }15% / $\bullet$ }30%), random horizontal flips (with label correction), rotations ($\bullet$ }10°), scaling (0.5–1.5×), shearing ($\bullet$ }3°), selective cropping. This guarantees consistent class distribution and enhanced resilience in face of real-world fluctuations. Figure 1(b) illustrates the full preprocessing pipeline: from raw video frames to segmentation, annotation, augmentation, and YOLO-ready training data.

3.3 Model Implementation

Brake-light detection is something the Ultralytics YOLOv8 and YOLOv9 models could be made for. YOLOv8 employs an anchor-free detection anchor-free head for fast, accurate bounding-box predictions and a modern backbone with a PANet-style neck [22]. It generates a

confidence score for the "on" state by maximizing a pretrained model on a custom brake-light dataset. Programmable Gradient Information (PGI) is used to retain complete input information and produce more consistent gradients [23, 25]. Both models output a confidence score at inference and learn on the same brake dataset. Training and Inference Pipeline. The implementation of a model-based system uses the YOLOv8/YOLOv9 workflow, consisting of three core phases: training, inference, and validation. Starting the model using several architectural variations (n/s/m/l/x/c), preprocessing using augmentation strategies, optimizing using stochastic gradient descent and focal loss, and training for 50–100 epochs. Metrics including mAP@0.5 and mAP@[0.5:0.95] track performance to allow early stopping should improvements level off. Input images are resized in the inference phase to preserve aspect ratio and go once forward through the neck and head modules for detection. Non-max suppression (NMS) with a confidence threshold of $\tau = 0.5$ and an IoU threshold of 0.7 handles postprocessing. Validation metrics include class-wise precision and recall metrics (see Table 1). Multi-Modal Fusion Algorithm. To leverage the strengths of both visual and rangebased data, we apply a late-fusion strategy that combines the brake-confidence score from the YOLO detector with the IR-based proximity measurement. The YOLO model generates a brake confidence score p_v , and the IR sensor offers a raw distance d . We convert this distance into a normalized proximity score $s_d = \max \left[0, \frac{(D_{thr} - d)}{D_{thr}} \right]$, where D_{thr} is a pre-set threshold reflecting the critical stopping distance. With $\alpha \in [0, 1]$ managing the balance between the two modalities, the final brake probability is computed using a weighted average $p_f = \alpha p_v + (1 - \alpha) s_d$. Detecting the vehicle, computing p_v , d , s_d , and fusing signals into p_f forms the fusion pipeline. Braking is indicated if $p_f > \tau$; else, it is not necessary. Particularly in demanding situations like low lighting or visual ambiguity, this multi-sensor integration improves brake-light detecting system dependability.

3.4 Experiments

The paper trained YOLOv8 and YOLOv9 object detectors on a single NVIDIA Tesla T4 GPU, fine-tuning for over one hundred epochs using pre-trained weights. The training configuration required 238 GFLOPs and 51M parameters, compared to 165 GFLOPs for YOLOv8. The training pipeline used random affine, color jitter, horizontal flips, and mosaic blending of four images, as standard YOLO augmentations. The default loss function used was the default YOLOv8/v9 loss, a weighted sum of box, objectness, and class terms. The training configuration produced higher computational cost for YOLOv9. The numerical results—shown in Table 1—show that YOLOv8 routinely beats YOLOv9 in accuracy and detection performance over all model sizes. YOLOv8 achieves higher mAP50 and mAP50-95 at every scale; for the medium (m) model, YOLOv8 reaches 0.45 mAP50 and 0.32 mAP50-95, compared to YOLOv9's 0.39 and 0.27 respectively. YOLOv8 scores 0.48 mAP50 at the largest (x) scale while YOLOv9 scores 0.44. These coherent patterns reveal the remarkable scale-dependent accuracy of YOLOv8. YOLOv9 does, however, only offer very modest speed and efficiency gains. Faster inference times are indicated by higher FPS over all model sizes; YOLOv9-n runs at 71 FPS while YOLOv8-n 6 Harsh Jaiswal et al.

Table 1. Performance comparison of YOLOv8 and YOLOv9 models across different sizes

Model	Size	mAP50	mAP50-95	FPS	Parameters	FLOPs
YOLO	n	0.28	0.16	67	3.2M	8.7G
	s	0.41	0.29	45	11.2M	28.6G
	m	0.45	0.32	32	25.9M	78.9G
	l	0.47	0.34	24	43.7M	165.2G
	x	0.48	0.35	19	68.2M	257.8G
YOLOv9	b	0.24	0.14	71	2.8M	7.9G
	s	0.30	0.21	48	9.8M	26.4G
	m	0.39	0.27	35	24.1M	76.3G
	l	0.42	0.30	26	41.5M	162.7G
	c	0.44	0.32	21	66.8M	248.1G

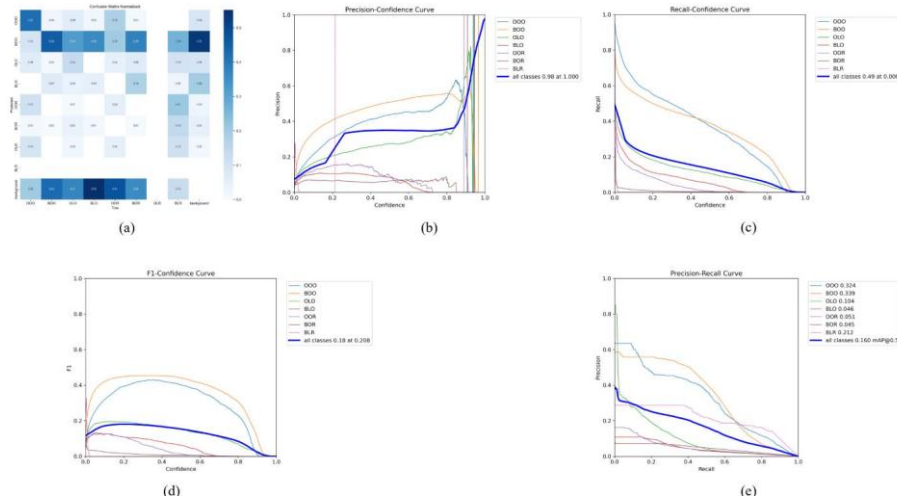


Fig. 2. Evaluation metrics for YOLOv9 models

runs at 67 FPS. Regularly requiring less parameters and FLOPs, YOLOv9 is also more lightweight. Whereas YOLOv8 has 68.2M parameters and 257.8G FLOW, YOLOv9 has 66.8M parameters and 248.1G FLOW at the extra-large (x / c) configuration. Concurrent with these efficiency gains are a corresponding decline in detection performance. YOLOv8 is still the better choice when accuracy comes first even if YOLOv9 could be better in cases when speed or resource constraints are vital. YOLOv8 offers superior real-time performance, particularly for applications prioritizing detection quality over computational speed or size. Its cleaner and higher detection accuracy makes it a reliable choice. YOLOv9, on the other hand, offers faster inference and simplicity, but it's better suited for limited resource environments.

4 Conclusion and Future Work

The study presents a brake-light detection system that integrates dual YOLO models with an infrared range sensor using a multi-modal approach. The system offsets adverse lighting and partial occlusion using an IR-guided dynamic fusion algorithm. A specialized preprocessing pipeline in HSV color space separates brake-light candidates before detection. In validation

studies on reasonable traffic scenarios, the fused model beats either YOLO detector. The system detects brake-on sequences and distinguishes them from regular driving conditions, so preserving high accuracy and recall. Transformer-based vision systems, integrating other sensing modalities, and vehicle-to-vehicle communication of brake-status messages constitute future extensions.

References

- 1) Chen, D.Y., Peng, Y.J., Chen, L.C., Hsieh, J.W.: Nighttime turn signal detection by scatter modeling and reflectance-based direction recognition. *IEEE Sensors Journal* 14(7), 2317–2326 (2014). <https://doi.org/10.1109/JSEN.2014.2306496>.
- 2) Dash, D.K.: Rear end collisions had maximum share of road deaths in 2021 and 2022, shows govt data. *Times of India* (2024), news Report: <https://timesofindia.indiatimes.com/india/rear-end-collisions-had-maximum-share-of-road-deaths-in-2021and-2022-show-govt-data/articleshow/111736173.cms>.
- 3) Geesung Oh, S.L.: One-stage brake light status detection based on yolov8. *Sensors* 23(17):7436 (2023). <https://doi.org/https://doi.org/10.3390/s23177436>.
- 4) Ghahremannezhad, H., Shi, H., Liu, C.: Real-time accident detection in traffic surveillance using deep learning. In: 2022 IEEE International Conference on Imaging Systems and Techniques (IST). IEEE (Jun 2022).
- 5) Hou, X., Gan, M., Zhang, J., Zhao, S., Ji, Y.: Secondary crash mitigation controller after rear-end collisions using reinforcement learning. *Advanced Engineering Informatics* 58, 102176 (2023). <https://doi.org/https://doi.org/10.1016/j.aei.2023.102176>, <https://www.sciencedirect.com/science/article/pii/S147403462300304X>.
- 6) Hsu, H.K., Tsai, Y.H., Mei, X., Lee, K.H., Nagasaka, N., Prokhorov, D.V., Yang M.H.: Learning to tell brake and turn signals in videos using cnn-lstm structure. In: IEEE 20th International Conference on Intelligent Transportation Systems (ITSC) (2017).
- 7) J3016: Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. Society for Automotive Engineers (SAE) (2021).
- 8) Jocher, G.: Yolov9 vs yolov8: A technical comparison. *Ultralytics YOLO Docs* (2025), [https://docs.ultralytics.com/compare/yolov8-vs-yolov9/Keylabs: Yolov8 vs ssd: Choosing the right object detection model \(2023\), article Link: https://www.perplexity.ai/search/hi-so-i-am-doing-a-research-ti1hmZRbheRM6Z8xe2EnlOkg](https://docs.ultralytics.com/compare/yolov8-vs-yolov9/Keylabs: Yolov8 vs ssd: Choosing the right object detection model (2023), article Link: https://www.perplexity.ai/search/hi-so-i-am-doing-a-research-ti1hmZRbheRM6Z8xe2EnlOkg).
- 9) Kim, J.: Detecting the turn on of vehicle brake lights to prevent collisions in highway tunnels. *Sustainability* 14(21), 14322 (2022). <https://doi.org/https://doi.org/10.3390/su142114322>
- 10) Kim, J.: Detecting the turn on of vehicle brake lights to prevent collisions in highway tunnels. *Sustainability* 14(21) (2022). <https://doi.org/10.3390/su142114322>, <https://www.mdpi.com/2071-1050/14/21/14322>
- 11) Kushwaha, M., Abirami, M.S.: Forward vehicle brake light detection for avoiding road accidents using yolov7 and IR sensor. *Multimed. Tools Appl.* (Jun 2024)
- 12) MathWorks: Morphological dilation and erosion. <https://www.mathworks.com/help/images/morphological-dilation-and-erosion.html>, accessed: 2025-04-20
- 13) NHTSA: Traffic safety facts. National Highway Traffic Safety Administration pp. 82–122 (2021), report: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813527>
- 14) Oh, G., Lim, S.: One-stage brake light status detection based on yolov8. *Sensors* 23(17) (2023). <https://doi.org/10.3390/s23177436>, <https://www.mdpi.com/1424-8220/23/17/7436>

- 15) Organization, W.H.: Road traffic injuries (2023), fact sheet: <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>.
- 16) Parvin, S., Rozario, L.J., Islam, M.E.: Vision-based on-road nighttime vehicle detection and tracking using taillight and headlight features. *J. Comput. Commun.* 09(03), 29–53 (2021).
- 17) of Road Transport, M., Highway: Road accidents in india 2022. Government of India (2022), report: https://morth.nic.in/sites/default/files/RA_2022_30_Oct.pdf.
- 18) Saini, M., Singh, H., Sengupta, E., Aggarwal, A., Singh, H., Kumar, N.: An intelligent machine learning-enabled cattle reclining risk mitigation technique using surveillance videos. *Neural Comput. Appl.* (Nov 2023).
- 19) Satzoda, R.K., Trivedi, M.M.: Looking at vehicles in the night: Detection and dynamics of rear lights. *IEEE Transactions on Intelligent Transportation Systems* 20(12), 4297–4307 (2019). <https://doi.org/10.1109/TITS.2016.2614545>.
- 20) Shi, Z., Yin, Z., Chang, S., Yi, X., Yu, X.: Efficient oriented object detection with enhanced small object recognition in aerial images (2024).
- 21) Torres, J., Austen, J.T.J.: Yolov8 architecture; deep dive into its architecture (Jan 2025), <https://yolov8.org/yolov8-architecture/Ultralytics:YOLOv9—docs.ultralytics.com>. <https://docs.ultralytics.com/models/yolov9/>, [Accessed 29-04-2025].
- 22) Vogt, P., Riitters, K.H., Estreguil, C., Kozak, J., Wade, T.G., Wickham, J.D.: Mapping spatial patterns with morphological image processing. *Landsc. Ecol.* 22(2), 171–177 (Jan 2007).
- 23) Wang, C.Y., Yeh, I.H., Liao, H.Y.M.: Yolov9: Learning what you want to learn using programmable gradient information (2024), <https://arxiv.org/abs/2402.13616>.
- 24) Wikipedia: List of countries by traffic-related death rate (2025), wiki: https://en.wikipedia.org/wiki/List_of_countries_by_traffic-related_death_rate.
- 25) Yuan, M., Wei, X.: C²former: Calibrated and complementary transformer for rgbinfrared object detection. *IEEE Transactions on Geoscience and Remote Sensing* 62, 1–12 (2024). <https://doi.org/10.1109/TGRS.2024.3376819>.

A Model-Based Engineering Approach for the Design and Validation of an Incline Impact Testing Platform

Tejas Parbhane¹, Nihal Dhote¹, Nilesh Sakle², Murtaza Laila²

¹Department of Mechanical Engineering, COEP Technological University, Pune – India

²Automotive Research Association of India, Pune – India

Email: tejasparbhane13@gmail.com

Abstract

Product damage during shipping is a common issue. Test systems simulating transport shocks are essential for the validation of packaging design. This paper presents a lifecycle-based entire development of an automated incline impact test rig. It combines structural modelling, simulation, automation, and physical validation based on standards such as ASTM D880, ISO 2244, and ISTA 3-series. The test rig consists of a trolley traveling down a sloped guide track, striking an impact wall. CAD models of the trolley, rail support, and impact support structure were designed using SolidWorks. All parts were manufactured with IS 4923 YST 310 steel. Structural integrity was confirmed with FEA simulations in ANSYS, showing deformation of less than 1 mm and stresses safely within material yield limits. To ensure operational accuracy, a Delta PLC-HMI system was implemented. Travel time was monitored in order to calculate impact velocity by two Omron photoelectric sensors. Slope angle was monitored by an ADXL335 accelerometer and communicated through RS485 to the PLC. These signals-initiated actions and provided real-time test feedback through a touchscreen HMI interface. Dynamic loading is supported by the rig up to 1 ton from 250 kg. Testing verified that reliable impact velocity was achieved at 1.5 m/s, irrespective of different weights. Less than 3% variation from calculated velocity was verified, validating that simulation results were accurate. The present research identifies a complete-cycle engineering design from conceptual modelling to detail design, logical functionality, and physical performance. Each subsystem was specified, confirmed, and optimized under real-world constraints. Control logic was planned with precision and safety in mind, and communication between processor and sensors was made modular. This rig is one in a model-based product definition project. It shows standards and simulation tools can enable more rapid, accurate hardware development. Its value lies not just in impact testing, but in explaining how digital tools help deliver design ideas into stable, testable systems ready for industry.

Keywords: Incline Impact Testing, Packaging Evaluation, PLC- HMI Automation, Transport Simulation

1. Introduction

The safe transportation of goods is a critical requirement across industries such as automotive, electronics, pharmaceuticals, and food logistics. To ensure product integrity during shipping and handling, packaging systems must withstand various dynamic forces, including sudden shocks and impacts. One of the key methods for simulating such real-world events is the Incline Impact Test, which evaluates a package's ability to absorb energy during an abrupt impact event. The incline impact test involves releasing a loaded trolley down a sloped surface to strike a rigid impact wall, simulating forces encountered during sudden stops, collisions, or rough handling. The measured impact performance provides valuable insight into potential packaging weaknesses and opportunities for design improvement.

In the current work, industry-accepted standards guide both the design of the test rig and the evaluation of package performance: ASTM D880^[1] (Standard Test Method for Impact Testing for Shipping Containers and Systems) and ISO 2244^[2] (Packaging - Complete, Filled Transport Packages - Horizontal Impact Tests) provide detailed guidelines for the design, construction, and operation of the incline impact test setup. These standards ensure that the rig construction, test methodology, and measurement parameters meet international best practices. ISTA 3A^[3],

3B^[4], and 3E^[5] standards, on the other hand, define specific performance evaluation protocols for packaged products. They describe various test setups and sequences, including incline impact tests, aimed at assessing packaging quality under simulated transportation conditions. This study aims to develop and design a standard-compliant incline impact test rig in accordance with ASTM D880 and ISO 2244. The project combines PLC-HMI-based automation and sensor feedback in order to achieve accurate control, operational safety, and velocity reproducibility. Finite element simulations determine the choice of materials for the purpose of optimizing structural strength and cost-effectiveness. Experimental testing following fabrication confirms the actual impact velocities, allowing cross-verification with theoretical values. Lastly, the capacity of the rig to produce reproducible and consistent test data is tested according to ISTA standards 3A , 3B , and 3E .

2. Literature Review

L. Cao and Y. Yuan ^[6], introduce a new incline impact testing machine with a flexible impact panel. The design incorporates important elements like a winch, impact incline, impact panel, and support structures. The impact panel is made to be raised to a perpendicular position with respect to the incline through a hydraulic cylinder, allowing controlled impact testing. J. Kang and D. Wang ^[7], present a novel inclined plane impact testing device. The tool is made to simulate actual-life impact conditions, testing the packaging packages resistance under conditions such as handling and transport. The structural part of the rig consists of high-strength steel used to provide it with resilience as well as accuracy in performing the tests. J. Zhou and W. Zhang ^[8] presented a test apparatus specifically designed to examine the slope impact resistance of packaging boxes. It features a variable incline platform and traveling cart to simulate various impact conditions. The principal rig frame is constructed from reinforced aluminium alloy due to its strength and lightness, which enhances the portability and convenience of the device.

A thesis by R. Sharma ^[9] explains the application of a Siemens S7-1200 PLC interfaced with an HMI to automate a tensile testing rig. The motor-driven jaws were controlled by the system, and load and displacement were monitored using sensors. It might be able to run multiple test cycles with a single command. Test limits might also be specified by the users, observed graphs, and even results export through USB. As per Ahmed et al. ^[10], the incorporation of PLCs in mechanical testing greatly enhances the accuracy, consistency, and safety of the system. PLCs are favoured for their rugged nature and fast response rate. They can process sensor data in milliseconds, activate actuators immediately, and disconnect systems when they fail to avoid damage or injury.

3. Incline Impact Test Rig Design

3.1 Structural and Material Selection

Incline impact test rig structure consists of heavy structural parts including trolley structure, impact support structure, and guide rail support structure. For achieving maximum strength, longer life, and economy, IS 4923 YST 310 mild steel was selected to be utilized in the structural parts. The material has a yield strength of 310 MPa ^[11] and an ultimate tensile strength of 450 MPa ^[11] and can be applied in load-bearing structures with enhanced machinability and weldability.

3.2 Load Condition

The incline impact rig was fabricated to accommodate a maximum load capacity of 1 ton, where the trolley could achieve impact velocities of between 1.1 m/s and 2m/s. The rig's structure was designed such that the trolley glides along the inclined guide rails with control and stability

during testing.

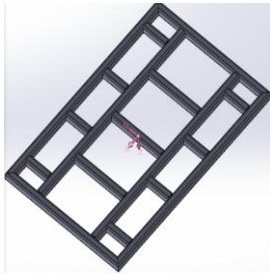


Figure 3 CAD Design of Trolley Structure

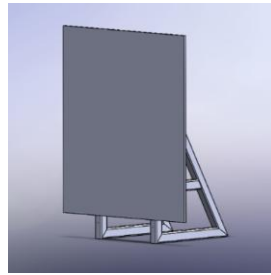


Figure 4 CAD Design of Impact Structure

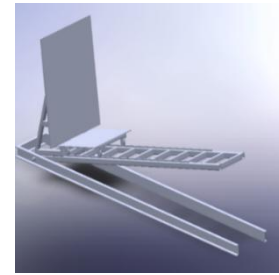


Figure 5 CAD Design of Overall Test Rig

These values were ascertained from simulation analysis and experimental verification so that the rig is in compliance with industry standards for impact testing. Simulations established the capability of structural components to support loads applied and experimental results were ascertained up to the 1000 kg load for the same.

3.3 Stress, Strain, Deformation Calculation

Bending Stress (σ);

$$\sigma = \frac{M}{Z} \quad (1)$$

Strain;

$$\varepsilon = \frac{\sigma}{E} \quad (2)$$

Deformation for UDL;

$$\delta = \frac{5wL^4}{384EI} \quad (3)$$

Deformation for Point Load;

$$\delta = \frac{FL^3}{48EI} \quad (4)$$

Table 4 Calculated Stress, Strain and Deformation value for Rig Component

	Stress (MPa)	Strain	Deformation (mm)
Trolley Structure	80.2	4.01×10^{-4}	1.22
Guide Rail Support Structure	70.5	3.5×10^{-4}	3.236
Impact Support Structure	145	2.6×10^{-4}	1.41

3.4 Simulation Analysis

All of the static analysis of individual parts was performed on ANSYS Workbench 18.1 and CAD designs of each component in SolidWorks 2020. The outcome also confirms that these components have been structurally strong under specified loads and stress conditions.

3.4.1 Trolley Structure using IS 4923 Material

A static structural analysis of the structure of the trolley was done using IS 4923 material under realistic loading and boundary conditions. The structure was loaded by 15,000 N with a factor of safety of 1.5 to represent real-world impact conditions. In keeping with accuracy, the initial default mesh was reiterated to 25 mm element size, and points of support were fully constrained to simulate actual conditions of mounting. The results obtained from analysis predicted a total deformation of only 0.24 mm, which reflects high stiffness and rigidity of IS 4923 under the same load. The corresponding stress and strain value 86.04 MPa and 0.00078 mm/mm respectively, far below the material's yield point, which makes sure that the structure remains in the safe operation region quite substantially.

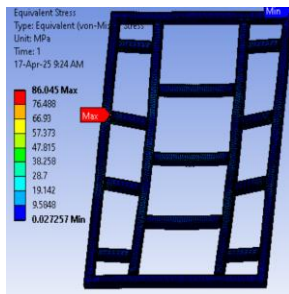


Figure 6 Equivalent Stress of Trolley Structure

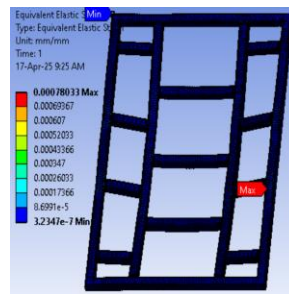


Figure 7 Equivalent Strain of Trolley Structure

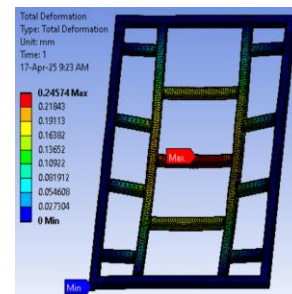


Figure 8 Total Deformation of Trolley Structure

These results confirm that IS 4923 is a very suitable material for the trolley structure in the incline impact test rig and provides the maximum strength, minimum deformation, and stable behaviour when subjected to repeated impact loading.

3.4.2 Guide Rail Support Structure using IS 4923 Material

Static structural analysis was performed for support structure of guide rail using material IS 4923 with real loading and boundary conditions. The structure has been loaded with a load of 16,000 N, including safety factor of 1.5, to replicate the force that can be anticipated during the test. To achieve both maximum computational speed and accuracy, the simulation was carried out with the default mesh parameters, but the support points were fully constrained to simulate real mounting conditions.

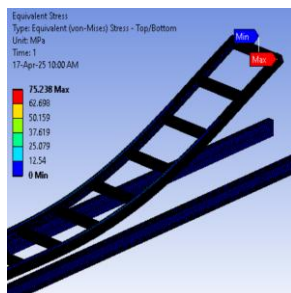


Figure 9 Equivalent Stress of Guide Rail Support Structure

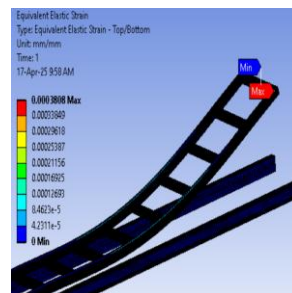


Figure 10 Equivalent Strain of Guide Rail Support Structure

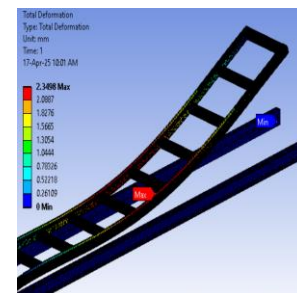


Figure 11 Total Deformation of Guide Rail Support Structure

The analysis result for IS 4923 yielded a total deformation of 2.34 mm, which confirms that the material provides superior stiffness and structural stability under the given load. The respective equivalent stress was 75.23 MPa, which is well within the yield strength of the material, so that the structure is always in the elastic zone when the structure is subjected to loading. These results validate that IS 4923 is a reliable and appropriate material for the guide rail support structure since it can resist applied loads without the slightest deformation or failure of the structure.

3.4.3 Impact Support Structure using IS 4923 Material

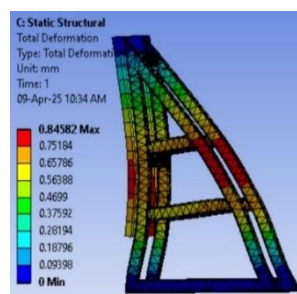
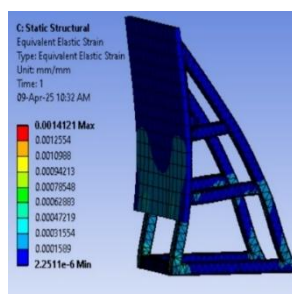
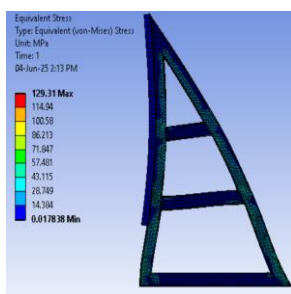


Figure 12 Equivalent Stress of Impact Structure

Figure 13 Equivalent Strain of Impact Structure

Figure 14 Total Deformation of Impact Structure

Transient structural impact test of the impact support structure made of IS 4923 material, selected due to the excellent structural capacity and suitability for impact loads, was carried out. The analysis was performed using hex-dominant meshing to obtain realistic deformation and stress values. The supporting structure was completely restrained along mounting points to simulate actual test conditions. Impact force of 62335.44N considering trolley velocity during impact is 1.5m/s.

3.5 Analytical Impact Velocity Calculation:

$$\text{Force acting on component, } f_{gravity} = \text{Total Mass} \times \text{Gravity} \times \sin \theta \quad (5)$$

Frictional Force,

$$f_{frictional} = \text{Coefficient of friction} \times \text{Total Mass} \times \text{Gravity} \times \cos \theta \quad (6)$$

$$\text{Net force on trolley, } f_{net} = f_{gravity} - f_{friction} \quad (7)$$

$$\text{Acceleration on Trolley, } a = \frac{f}{m} \quad (8)$$

$$\text{Impact velocity, } vi^2 = vo^2 + 2ad \quad (9)$$

Table 3 Calculated Velocity for different weight sample

Mass on Trolley	Impact Velocity
250 Kg	1.53
500 Kg	1.53
750 Kg	1.53
1000 Kg	1.53

4. Automation and Control System

4.1 Types of sensors used in the rig

The incline impact test rig developed makes extensive use of two primary types of sensors for measurement and monitoring purposes. The first is a photoelectric sensor (Omron E3FA series), which is employed to measure the location of the trolley in motion. Two photoelectric sensors are mounted at the start and at the end of the sloping track for determining the time for the trolley to cover the distance between the two. The time readings, coupled with the measured distance, are used to ascertain the velocity of the trolley at impact. ADXL335^[12] is a low-noise, high-precision, 3-axis MEMS accelerometer that measures incline angles precisely by sensing components of gravitational acceleration along its sensitive axes.

4.2 PLC and HMI Selection

To ensure precise control, automation, and data acquisition for the incline impact test rig, the Delta DVP14SS211R/T PLC (8 input / 6 output) and Delta DOP-103BQ 4.3" HMI were selected. The DVP14SS211R/T ^[15] offers fast processing speed, compact size, easy programmability, and sufficient I/O points to manage all sensor signals and control logic efficiently.

Similarly, the Delta DOP-103BQ HMI ^[16] provides a clear, responsive touchscreen interface ideal for real-time monitoring, at a much lower price point than HMIs from Mitsubishi or Schneider Electric, without sacrificing essential features like multi-language support and Modbus communication. Together, they deliver a highly reliable, economical, and easy-to-integrate solution well-suited for our project's technical and budgetary needs.

4.3 PLC- HMI Intervention with overall setup:

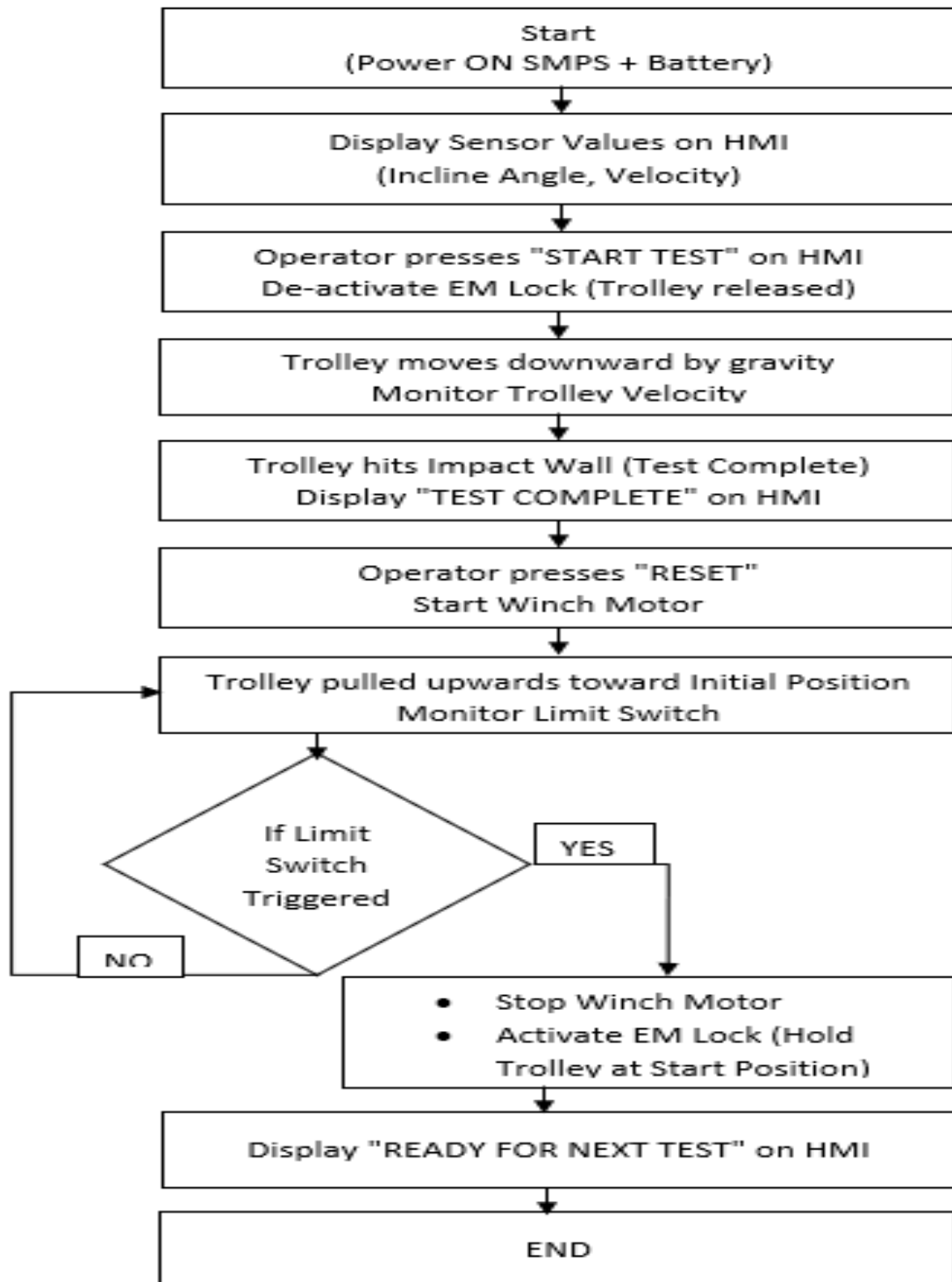


Figure 17 PLC - HMI Intervention with overall setup

4.4 PLC-HMI Intervention with sensor

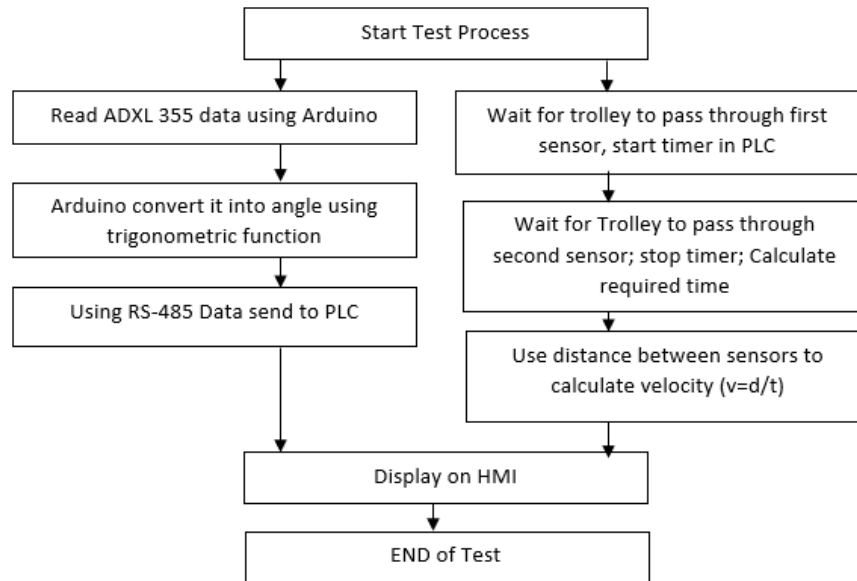


Figure 18 Flowchart representation of PLC-HMI working

5. Experimental Setup and Testing

5.1 Test Setup and Procedure

For the verification of the designed incline impact test rig, experimental testing was carried out using a package box and a trolley with different weights of 250 kg, 500 kg, 750 kg, and 1000 kg. The rig was inclined at an angle of 10 degrees, which is representative of actual transport shock conditions.

5.2 Equipment Used and Test Parameters

The trolley impact velocity was monitored with an Omron photoelectric sensor strategically located close to the impact point. Even though the rig was specifically designed to strike with a minimum of 1.1 m/s impact velocity, in experiment trials trolley velocity achieve 1.5 m/s velocity, confirming the design's robustness. All tests were conducted under controlled conditions to be able to replicate and substantiate findings.

6. Results and Discussion

The comparison between theoretical calculations and practical experimental results demonstrated strong alignment in the developed incline impact test rig. Theoretically, the impact velocity for all loading conditions (250 kg, 500 kg, 750 kg, and 1000 kg) was calculated to be approximately 1.531 m/s based on energy conservation principles, considering the 10⁰ incline angle and negligible frictional losses. Experimentally, with an Omron photoelectric sensor placed close to the point of impact, the trolley attained increasingly higher speeds with heavier loads—1.53 m/s, 1.54 m/s, 1.56 m/s, and 1.56 m/s for 250 kg, 500 kg, 750 kg, and 1000 kg respectively. The experimental values are very close to theoretical values, with a maximum difference of about 2.6%. The minimal increase in velocity under load is due to the practical considerations of rolling resistance reduction, wheel deformation effects, and mechanical tolerances that are not included in the idealized theory model.

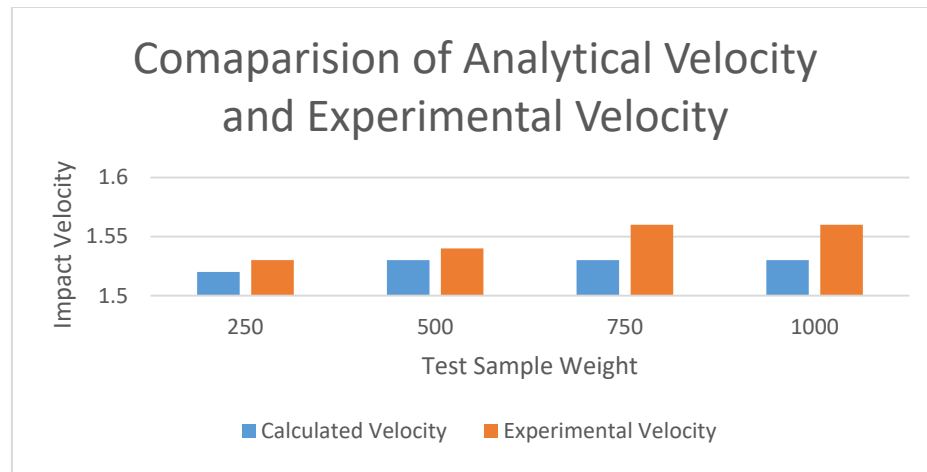


Figure 19 Comparison of Analytical Velocity and experimental velocity

The repeatability of trial velocities from repeated readings for any load case in three trials establishes the reproducibility and dependability of rig performance. It confirms that both the automation by the PLC-HMI system and the mechanical structure of the rig are effective to meet the target operating results as required, upholding both theoretical design and real-world implementation. Therefore, the experimental outcomes strongly agree with theoretical predictions, verifying that the rig is appropriate for standardized incline impact testing according to ASTM D880, ISO 2244, and ISTA specifications.

Reference

1. ISO 2244: Packaging - Complete, Filled Transport Packages and Unit Loads - Horizontal Impact Tests, Geneva, Switzerland: Int. Org. Standardization, 2000.
2. ASTM International, ASTM D880-92: Standard Test Method for Impact Testing for Shipping Containers and Systems, West Conshohocken, PA, USA, 2015.
3. International Safe Transit Association (ISTA), ISTA 3A: Packaged Products for Parcel Delivery System Shipment, ISTA, 2018.
4. International Safe Transit Association (ISTA), ISTA 3B: Packaged Products for Less-Than-Truckload (LTL) Shipment, ISTA, 2019.
5. International Safe Transit Association (ISTA), ISTA 3E: Unitized Loads of Same Product for Distribution Systems Shipment, ISTA, 2020.
6. L. Cao and Y. Yuan, "Design for Incline Impact Testing Machine," presented at the International Symposium on Packaging and Transportation Engineering, Dec. 2012.
7. J. Kang and D. Wang, "Inclined Plane Impact Testing Machine," Journal of Packaging Technology, May 2014.
8. J. Zhou and W. Zhang, "Testing Device for Slope Impact of Packaging Box," Chinese Packaging Technology Conference Proceedings, Jan. 2016.
9. Sharma, R. (2019). Development of PLC-based Tensile Testing System (Master's thesis). National Institute of Technology, Bhopal, India.

10. Ahmed, S., Gupta, N., & Pillai, R. (2021). Automation in mechanical testing using PLC-HMI systems: A modern approach. *International Journal of Engineering and Applied Sciences*, 8(3), 112–119.
11. ASM International, 1990. *ASM Metals Handbook*, Volume 1, “Properties and Selection: Irons, Steels, and High-Performance Alloys.”
12. D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, 10th ed. Hoboken, NJ: Wiley, 2013.
13. Omron Corporation, E3FA Photoelectric Sensor Datasheet, 2023.
14. Analog Devices, "ADXL335 Ultra Low Noise, Low Drift, 3-Axis MEMS Accelerometers," Mouser Electronics, 2025.
15. Delta Electronics, Delta DVP14SS211R/T PLC *Technical Manual (2021)*.
16. Delta Electronics, *DOP-103BQ HMI Technical Manual (2021)*

An Enhanced 1D-CNN-LSTM with Attention Mechanism for Real-Time Human Activity Recognition using IMU Data

Amit Kumar Singh¹, Reeshav Guha², Dr. Vibha Vyas²

¹Scientist, Research & Development Est. (DRDO), Pune, Maharashtra, India

²COEP Technological University, Pune

Abstract

Human Activity Recognition (HAR) plays a pivotal role in various applications, including healthcare and human-computer interaction. This paper presents an enhanced deep learning system for realtime HAR, leveraging a novel dataset collected using thigh-mounted LPMS-B2 Inertial Measurement Units (IMUs) from multiple participants. We selected six key features—angular velocities (X, Y), Euler angles (Roll, Pitch), and gravity-compensated linear accelerations (X, Y)—based on their established relevance for capturing lower-body kinematics. We propose an architecture featuring a multi-layer 1D Convolutional Neural Network (CNN) for robust spatial feature extraction, followed by a Long Short-Term Memory (LSTM) network to model temporal dependencies. A key contribution is the integration of a self-attention mechanism post-LSTM, enabling the model to dynamically weigh the importance of different temporal segments within an activity sequence. The system processes multi-channel time-series data. Data preprocessing involves label encoding, feature scaling, and sequence generation. The model, optimized through extensive hyperparameter tuning, is trained on our custom dataset which includes 'Bending', 'Idle', and 'Walking' activities, achieving a test accuracy of 99.39%. Furthermore, we detail the deployment for real-time prediction using data streamed via the OpenZen API. This work contributes a custom dataset specific to lower-body movements, an improved and well-justified model architecture, and demonstrates its feasibility for accurate HAR.

Keywords: Human Activity Recognition, Deep Learning, CNN, LSTM, Attention Mechanism, Wearable Sensors, IMU, Custom Dataset, Thigh-Mounted Sensors, Feature Selection, Real-Time Systems, OpenZen.

1. Introduction

Human Activity Recognition (HAR) aims to identify and classify human physical activities based on sensor data. With the proliferation of wearable devices equipped with Inertial Measurement Units (IMUs), which typically include accelerometers, gyroscopes, and magnetometers, there is an abundance of data available for HAR research and applications [1]. Accurate HAR systems can enable a wide range of applications, such as remote patient monitoring in healthcare [2], fitness tracking, elderly care, context-aware computing, and gesture recognition for smart environments. Traditional HAR approaches often relied on handcrafted feature engineering followed by classical machine learning classifiers [3]. Deep learning models, however, can automatically learn hierarchical feature representations from raw or minimally processed sensor data, often leading to superior performance [4]. Among these, Convolutional Neural Networks (CNNs) are adept at extracting salient local patterns [5], while Long Short-Term Memory (LSTM) units are well-suited for modeling temporal dependencies [6]. Hybrid CNN-LSTM models combine these strengths [7]. Further enhancements, such as attention mechanisms, can improve performance by allowing the model to focus on the most discriminative parts of the input sequence [13]. A critical aspect of developing robust HAR systems is the dataset used for training and evaluation. While numerous public datasets exist, they may not always capture the specific nuances of sensor placements or activities relevant to a particular application. Addressing this, our work utilizes a novel, custom-collected dataset acquired using LPMS-B2 IMU sensors strategically mounted

on participants' thighs. This placement is particularly informative for activities involving significant lowerlimb kinematics, such as 'Bending', 'Idle', and 'Walking'. The selection of specific IMU features, including gravity-compensated linear acceleration, is crucial for capturing the dynamic aspects of these movements effectively. This paper presents the development and implementation of an enhanced CNN-LSTM architecture incorporating a self-attention mechanism, trained and evaluated on our custom thigh-mounted IMU dataset.

Our contributions are:

1. The collection and characterization of a novel HAR dataset from 5 healthy adult volunteers using thigh-mounted LPMS-B2 IMU sensors, focusing on activities with distinct lower-body signatures such as Bending, Idle, and Walking, with a detailed rationale for feature selection.
2. Design and implementation of an enhanced hybrid CNN-LSTM model incorporating a self-attention mechanism to improve temporal feature weighting.
3. A comprehensive data preprocessing pipeline tailored for our custom time-series sensor data, including details on gravity compensation.
4. Detailed training methodology, including insights from extensive hyperparameter tuning (20 runs), and evaluation of the proposed model achieving high accuracy (99.39%) on our custom dataset.
5. A framework for real-time activity prediction using the trained model with live data streamed from OpenZen compatible IMU sensors.

The remainder of this paper is organized as follows: Section 2 discusses related work. Section 3 details our custom dataset, data preprocessing, the enhanced model architecture, and training. Section 4 describes the real-time prediction system. Section 5 presents the experimental setup and results. Finally, Section 6 concludes the paper.

2. Related Work

The field of HAR has seen significant evolution. Early works focused on statistical feature extraction followed by classical machine learning [3, 8]. Deep learning, particularly CNNs [4, 9] and RNNs like LSTMs [10], has enabled automatic feature learning and improved performance. Hybrid CNN-LSTM models [7, 11] leverage the strengths of both. Attention mechanisms, originating from natural language processing [15], have gained traction in HAR [13]. They allow models to weigh the importance of different input segments, focusing on more informative parts of the time-series data, which can be particularly beneficial for complex or transitional activities. Our work incorporates a self-attention layer after the LSTM to harness this capability. Sensor placement is a critical factor in HAR. While common placements include the wrist or waist, thigh-mounted sensors, as used in our study, offer distinct advantages for capturing lower-limb kinematics. This can be particularly useful for discriminating activities like walking and bending, which are central to our custom dataset. The development of HAR systems using data from thigh-mounted sensors is an area that can benefit from further research, especially with custom-designed deep learning models. The choice of features, such as using gravity-compensated acceleration, is also vital for robust performance across different device orientations and dynamic conditions [16].

3. Methodology

This section details our data acquisition using thigh-mounted sensors, data preprocessing including feature selection and gravity compensation, the proposed enhanced CNN-LSTM-Attention model architecture, and the training process.

3.1. Custom Dataset Acquisition and Sensor Data

A custom dataset was collected to develop and evaluate our HAR system. Data was acquired using LPMS-B2 IMU sensors (LP-Research GmbH), with one sensor securely mounted on the lateral aspect of the dominant thigh of each of five healthy adult volunteers (3 male, 2 female, age range 22-35 years) after providing informed consent. This placement was chosen for its sensitivity to lower-limb movements. The dataset comprises three primary activities: 'Bending' (simulating picking an object from the floor), 'Idle' (standing still), and 'Walking' (at a natural pace), representing a mix of static postures, quasi-static transitions, and dynamic locomotion. Participants performed sequences of these activities for approximately 2-3 minutes per activity, with short rests. Data was streamed wirelessly using the OpenZen API at a sampling rate of 100 Hz and manually annotated. For this study, we selected $N_f = 6$ features: Angular velocity (GyroX, GyroY), Euler angles (EulerX/Roll, EulerY/Pitch), and gravity-compensated Linear acceleration (LinAccX, LinAccY), chosen for capturing primary movement dynamics and isolating movement from static orientation.

For our best model ($T_s = 200$, a 2-second window), 825 sequences were created. Figures 1 and 2 illustrate representative sensor data.

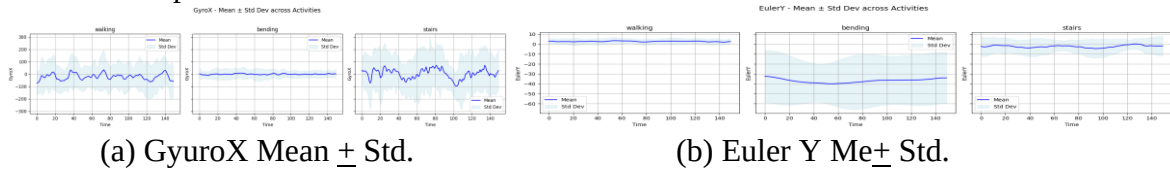


Figure 1: EDA: Mean and Standard Deviation of GyroX and EulerY sensor channels for Bending, Idle, and Walking.

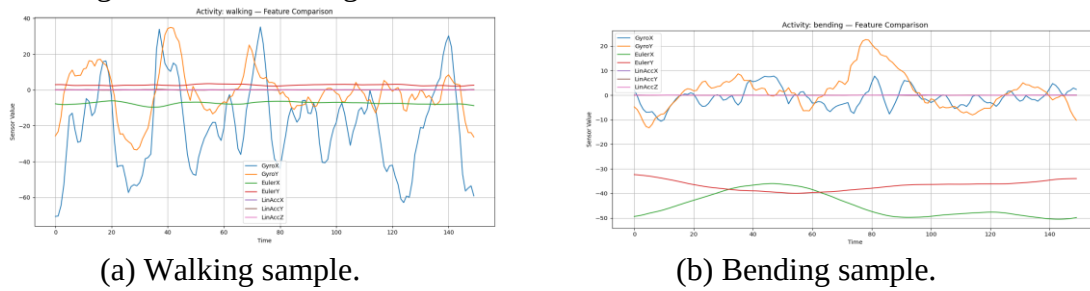


Figure 2: EDA: Example raw sensor signals (all $N_f = 6$ features, $T_s = 200$) for Walking and Bending

3.2. Data Preprocessing

Activity labels ('Bending', 'Idle', 'Walking') were converted to numerical representations (0,1,2) using LabelEncoder ($C = 3$). The $N_f = 6$ sensor features were standardized using StandardScaler from scikit-learn (Eq. 1).

$$x'_{ij} = (x_{ij} - \mu_j) / \sigma_j$$

Continuous sensor data was segmented into sequences of length $T_s = 200$ (2-second window at 100Hz) with a stride of $T_s/4 = 50$ samples. The label y_i for each sequence was determined by majority voting (Eq. 2).

$$y_i = \text{argmax}_c(\text{count}(\text{label } j = c \text{ for } j \in [k, k+T_s-1])) \quad (2)$$

This process resulted in 825 sequences, split into 660 training and 165 testing sequences (80:20 ratio), stratified by class.

3.3. Enhanced CNN-LSTM-Attention Model Architecture

The proposed model (Fig. 3) enhances a CNN-LSTM pipeline with a deeper CNN extractor and a selfattention mechanism. Dimensions correspond to our best configuration (Run ID 15, Table 1).

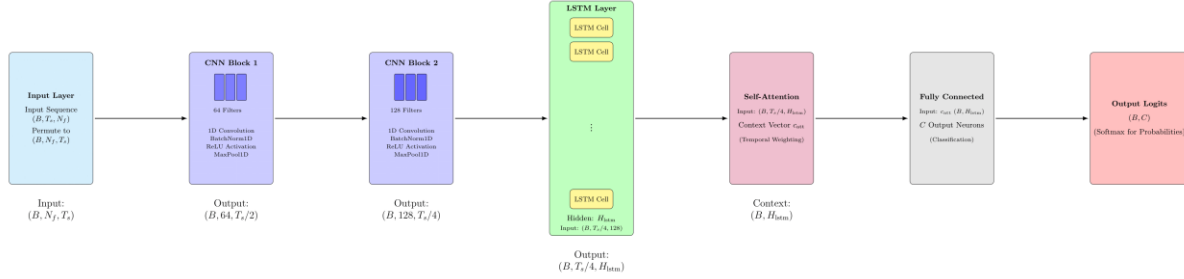


Figure 3: The proposed enhanced CNN-LSTM-Attention model architecture

The model layers are as follows:

1. Input Layer: Accepts sequences of shape (B, T_s, N_f) , where B is batch size, $T_s = 200$, and $N_f = 6$. Input is permuted to (B, N_f, T_s) for PyTorch's Conv1d.
2. CNN Blocks 1 & 2: Two stacked CNN blocks perform hierarchical feature extraction. Each includes: 1D Convolution (kernel $K = 5$, padding $P = 2$, stride 1), Batch Normalization, ReLU activation, and Max Pooling (kernel 2, stride 2). CNN Block 1 outputs $C1 = 64$ channels; Block 2, $C2 = 128$ channels. Max pooling reduces sequence length by half at each block, to $T' = T_s/4 = 50$.
3. LSTM Layer: A single LSTM layer with $H_{lstm} = 32$ hidden units processes the sequence of $C2 = 128$ feature vectors (length T') from the CNNs. It outputs a sequence of hidden states $h = [h_1, \dots, h_{T'}]$.
4. Self-Attention Layer: This mechanism takes LSTM hidden states h and computes a context vector c_{att} by focusing on important time steps, using an attention head with $H_{att} = 32$. Alignment scores e_t are computed (Eq. 3-4), normalized to weights α_t (Eq. 5), and used for a weighted sum c_{att} (Eq. 6).

$$u_t = \tanh(W_t h_t + b_h)$$

$$e_t = v_u^T u_t$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^{T'} \exp(e_j)}$$

$$c_{att} = \sum_{t=1}^{T'} \alpha_t h_t$$

5. Fully Connected (FC) Layer: Maps the H_{lstm} -dimensional c_{att} to $C = 3$ class logits (Eq. 7). $output\_logits = W_{fc} c_{att} + b_{fc}$ (7)

3.3.1. Architectural Design Rationale

The model's architecture is designed to effectively process time-series IMU data. Initial 1D CNN layers (kernel 5; $C1 = 64, C2 = 128$ filters) extract local temporal patterns hierarchically. Their parameter sharing and translation invariance (aided by pooling) are beneficial. Batch Normalization stabilizes training and improves gradient flow. Max Pooling reduces

dimensionality and focuses on salient features. An LSTM layer (Hlstm = 32 units) then models longer-term temporal dependencies from the CNN features. The subsequent Self-Attention mechanism (Hatt =32) dynamically weighs these temporal features, allowing the model to emphasize the most discriminative segments of an activity sequence. Finally, a Fully Connected layer performs classification based on this attention-weighted context vector.

3.4. Training Procedure The model is trained using the Adam optimizer [14] with learning rate η . The loss function is Cross-Entropy Loss LCE (Eq. 8), augmented with L2 regularization (weight decay $\lambda L2$) for the total loss L_{total} (Eq. 9).

$$L_{CE} = - \sum_{c=1}^C y_{0,c} \log(p_{o,c})$$

The custom dataset is split 80% for training and 20% for testing, stratified by activity. Training proceeds for 10 epochs, saving the model weights that yield the best test accuracy.

4. Real-Time HAR System

A key objective is deploying the trained model for real-time HAR using live data from an LPMS-B2 sensor via the OpenZen API. **System Components:** The system uses the saved StandardScaler, LabelEncoder, and the trained CNN-LSTM-Attention model weights. **Sensor Integration:** Client initialization, sensor discovery, connection, and IMU data streaming are handled by OpenZen. **Prediction Loop:** A deque of $maxlen = T_s$ buffers incoming sensor samples. When full, the sequence is scaled, tensorized, and fed to the model (in evaluation mode) for prediction. The predicted label and confidence are then displayed

5. Experimental Setup and Results

5.1. Dataset and Implementation Specifics

Experiments utilize the custom thigh-mounted IMU dataset (Sec. 3.1). For the best model (Run ID 15), $T_s = 200$ resulted in 825 sequences (660 train, 165 test). Models were trained for 10 epochs on a CPU. An example feature correlation heatmap is shown in Fig. 4. Implementation used Python 3.x with PyTorch, Scikit-learn, NumPy, Pandas, OpenZen, and Joblib. Extensive hyperparameter tuning (20 runs) explored sequence length, batch size, learning rate, L2 regularization, CNN channels, and LSTM units. Key parameters for the final model are in Table 1.

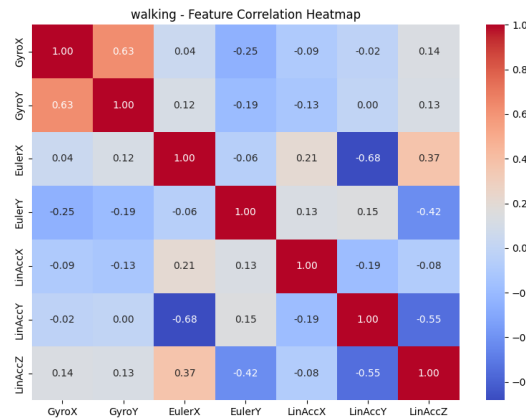


Figure 4: Example feature correlation heatmap for the 'walking' activity.

Table 1: Key Hyperparameters for the Final Enhanced Model (Run ID 15)

Parameter	Vlaue
Swquence Length (T_s)	200
Batch Size (B)	32
Number of Epochs	10
Learning Rate (η)	0.001

L2 Lambda (λ L2)	0.001
Optimizer	Adam
CNN Block 1 Output Ch. (C1)	64
CNN Block 2 Output Ch. (C2)	128
LSTM Hidden Size (Hlstm)	32
LSTM Dropout (Internal)	0.2
Attention Hidden (Hatt)	32
Test Split Ratio	0.2
Random State	42

5.2. Evaluation Metrics

Performance is evaluated using Accuracy, Precision, Recall, F1-Score (per-class and macro/weighted averages), and a Confusion Matrix.

5.3. Results and Discussion

The final model (Run ID 15) achieved an overall test accuracy of 99.39% on 165 test sequences. Learning curves (Fig. 5) illustrate effective training progression, with test accuracy stabilizing above 98.79% from epoch 7 and peaking at 99.39

Learning Curves for Best Model Configuration (Run 15)

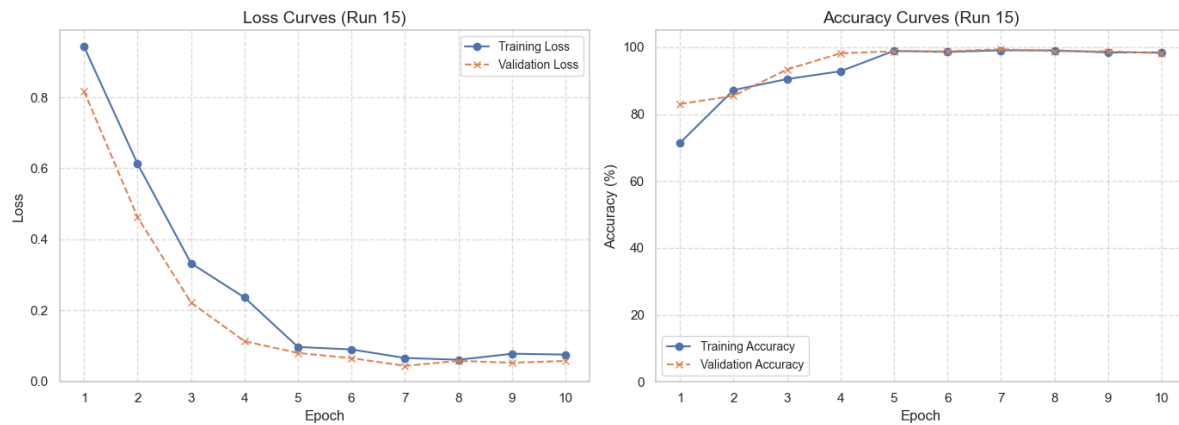


Figure 5: Training and validation loss and accuracy curves over 10 epochs for the best model (Run ID 15).

Per-class performance metrics for Run ID 15 are detailed in Table 2. 'Idle' and 'Walking' activities were perfectly classified. 'Bending' also achieved very high metrics (Precision 98.1%, Recall 100%, F1- Score 99.0%). The confusion matrix (Fig. 6) visually confirms these strong results, showing only a single 'Bending' sample misclassified.

Table 2: Per-Class Performance Metrics (Final Model - Run ID 15) on Test Set

Class	Precision	Recall	F1-score	Support
Bending	0.9821	1.0000	0.9910	55
Idle	1.0000	1.0000	1.0000	5
Waling	1.0000	0.9818	0.9908	55
Accuracy			0.9939	165
macro Avg	0.9940	0.9939	0.9939	165
Weighted Avg	0.9940	0.9939	0.9939	165

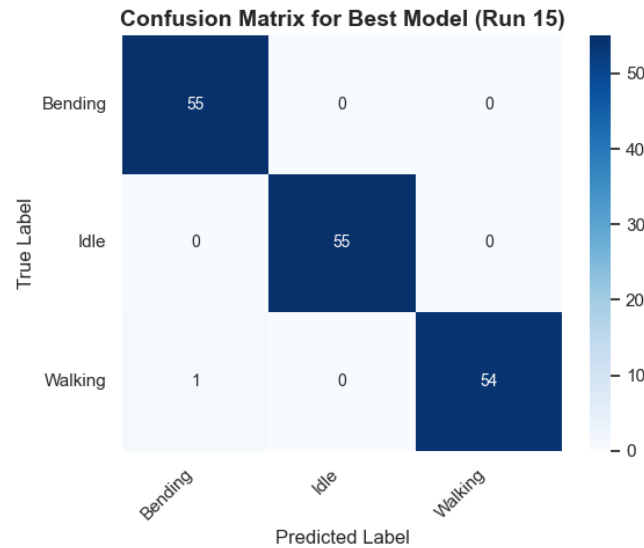


Figure 6: Confusion matrix for the final model (Run ID 15) on the custom test set.

Hyperparameter tuning (20 runs) highlighted $T_s = 200$ as optimal. Learning rates between 0.001-0.005 and $\lambda L2 = 0.001$ proved robust. The chosen CNN ($C1 = 64$) and LSTM ($Hlstm = 32$) sizes provided an effective balance of model capacity and efficiency compared to more complex configurations. The real-time prediction system demonstrated responsive and accurate classification during live testing.

6. Conclusion and Future Work

This paper introduced an enhanced hybrid CNN-LSTM architecture with a self-attention mechanism for Human Activity Recognition, leveraging a novel dataset from thigh-mounted LPMS-B2 IMUs. We detailed data collection for 'Bending', 'Idle', and 'Walking', a preprocessing pipeline with reasoned feature selection, the model architecture, and extensive hyperparameter tuning. The final model achieved 99.39% test accuracy on our custom dataset, effectively discriminating these lower-body movements. The real-time system confirmed practical applicability.

Future work could explore

- **Dataset Expansion:** Augmenting the custom dataset with more participants, a wider range of activities (including more subtle or complex ones like different types of bending or turning), and diverse environmental contexts.
- **Comparative Model Analysis:** Rigorous comparison against other state-of-the-art HAR models, including different attention variants or transformer-based architectures, on our custom dataset.
- **Dual-Thigh Data Fusion:** Investigating advanced sensor fusion techniques to combine data from IMUs on both thighs, potentially improving robustness and enabling recognition of asymmetric movements.
- **Attention Explainability:** Visualizing attention weights to gain deeper insights into which parts of the sensor signal sequences are most critical for the model's predictions for different activities.
- **Model Personalization and Adaptation:** Developing methods for adapting the general model to individual user movement patterns or to new, unseen activities with minimal retraining.

- On-Device Optimization: Further optimizing the model (e.g., quantization, pruning) for efficient deployment on resource-constrained wearable devices.

The combination of a targeted custom dataset from thigh-mounted IMUs and an attention-enhanced deep learning model offers a strong foundation for developing more accurate, nuanced, and practical HAR systems.

References

1. J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, Deep learning for sensor-based activity recognition: A survey, *Pattern Recognition Letters*, vol. 119, 2019, pp. 3–11.
2. O. D. Lara and M. A. Labrador, A survey on human activity recognition using wearable sensors, *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, 2013, pp. 1192–1209.
3. D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, A public domain dataset for human activity recognition using smartphones, in *Proc. European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*, 2012.
4. C. A. Ronao and S.-B. Cho, Human activity recognition with smartphone sensors using deep learning neural networks, *Expert Systems with Applications*, vol. 59, 2016, pp. 235–244.
5. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, vol. 86, no. 11, 1998, pp. 2278–2324.
6. S. Hochreiter and J. Schmidhuber, Long short-term memory, *Neural Computation*, vol. 9, no. 8, 1997, pp. 1735–1780.
7. F. J. Ordonez and D. Roggen, Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition, *Sensors*, vol. 16, no. 1, 2016, p. 115.
8. L. Bao and S. S. Intille, Activity recognition from user-annotated acceleration data, in *Proc. Pervasive Computing (PERVASIVE)*, 2004, pp. 1–17.
9. J. Yang, M. N. Nguyen, P. P. San, X. L. Li, and S. Krishnaswamy, Deep convolutional neural networks on multichannel time series for human activity recognition, in *Proc. International Joint Conference on Artificial Intelligence (IJCAI)*, 2015, pp. 3995–4001.
10. N. Y. Hammerla, S. Halloran, and T. Plotz, Deep, convolutional, and recurrent models for human activity recognition using wearables, in *Proc. International Joint Conference on Artificial Intelligence (IJCAI)*, 2016, pp. 1533–1540.
11. K. Xia, J. Huang, and H. Wang, LSTM-CNN architecture for human activity recognition, *IEEE Access*, vol. 8, 2020, pp. 56855–56866.
12. S. Yao, S. Hu, Y. Zhao, A. Zhang, and T. Abdelzaher, Deepsense: A unified deep learning framework for time-series mobile sensing data processing, in *Proc. International Conference on World Wide Web (WWW)*, 2017, pp. 351–360.
13. H. Ma, W. Li, X. Zhang, S. Gao, and S. Lu, Attnsense: Multi-level attention mechanism for multimodal human activity recognition, in *Proc. International Joint Conference on Artificial Intelligence (IJCAI)*, 2019, pp. 3109–3115.
14. D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*, 2014.
15. D. Bahdanau, K. Cho, and Y. Bengio, Neural machine translation by jointly learning to align and translate *arXiv preprint arXiv:1409.0473*, 2014.
16. D. Figo, P. C. Diniz, D. R. Ferreira, and J. M. P. Cardoso, Preprocessing techniques for context recognition from accelerometer data, *Personal and Ubiquitous Computing*, vol. 14, 2010, pp. 645–662.

Important uses of AI chatbots in education and healthcare sector

Neeraj Kumar ¹, Ravindra Nath ²
neerajmtech@gmail.com rnkatiyar@gmail.com²

Abstract

Recently, the use of chatbots in education has changed how we approach teaching and learning, leading to a new era of personalized and interactive education. When bringing chatbot technology into schools, it's important to be thoughtful and careful. Good strategies include making modified learning experiences, using natural language processing, and making sure it fits well with current educational tools. A new age of learning and research is emerging, driven by chatbots and artificial intelligence. [1-3] However, using these technologies comes with various challenges, especially concerning ethics. This paper looks into how AI systems and chatbots can be used in schools and research, focusing on the ethical issues involved. The researcher used a qualitative approach to gather information and analyze expert opinions. This study gives a basic understanding of the topic and helps us see the different experiences and viewpoints related to it, while also suggesting possible solutions to the problems identified. [4-5] The research discusses both the benefits and drawbacks of AI systems and chatbots, as well as how they can assist human skills and decision-making. It also addresses the ethical concerns that arise from using these technologies in research, including the risks of misuse. The study suggests practical solutions to these ethical challenges. It acknowledges that we are entering a new phase of education and research powered by AI. The advancements in technology will change how research is conducted and how education is delivered, particularly in assessments. The study also looks at how ChatGPT can help break down language barriers, and encourage teamwork in learning.

Keywords: Chatbots, ChatGPT, Artificial Intelligence, Education, Teaching, Healthcare.

1. Introduction

In today's digital world, education has changed a lot because of new technologies that have transformed the way we teach and learn. One of the coolest advancements is the use of chatbots, which are smart computer programs that can talk like humans. These chatbots are making a big impact in education, helping teachers and students at all levels and in different subjects. As students want more personalized and interactive ways to learn, the use of chatbots in education is growing rapidly. Still, bringing AI into schools and universities comes with its own set of challenges and debates. Many studies and articles have looked into the advantages of using AI in academia. One major advantage is that it can make research more efficient and accurate. AI can sift through huge amounts of data quickly, spotting patterns and connections that might be hard for people to see. This allows researchers to concentrate on more complex and creative work. [1-5] Another plus is that AI can tailor learning experiences to individual students. By analysing how students learn and what they can do, AI can offer personalized advice and support to help them reach their goals. Additionally, AI systems can promote online learning and broaden access to education for students in remote locations, contributing to a more equitable and inclusive educational environment. This study thoroughly investigates the realm of chatbot-enhanced education, examining its diverse aspects, the challenges encountered by educators and students, and the crucial role of advanced chatbot models, such as ChatGPT, in influencing the future of education. [7]

This study will look at the good and bad sides of AI systems and chatbots. We will adopt an exploratory research approach, emphasizing qualitative research to gather insights from experts and their analyses. The aim of this study is to contribute to the ongoing discussion about the ethical challenges that come with using AI systems and chatbots in changing education and research.[9]

2. Literature Survey

In recent years, especially since late 2022, the use of artificial intelligence (AI) and chatbots in education and research has grown a lot. Chatbots are automated tools that can talk to people like humans by using natural language processing and machine learning. Chatbots are computer programs that use artificial intelligence (AI) to simulate human-like conversations with users, offering a range of benefits in educational settings (Kaplan & Haenlein, 2019). Chatbots are being used in different subjects like language learning, science, and math. They can also help students with things like advice and support. But there are some challenges to using chatbots in education, like making sure they're accurate and fair. Additionally, researchers noted that AI can help automate boring tasks, allowing more time for complex issues and preparing the workforce for an AI-driven future.[6-8]

We will also discuss about the difficulties and policy issues that come with using AI in education and getting students ready for a world where AI is everywhere. The benefits and drawbacks of using AI in education need more research, especially with the rise of powerful AI chatbots like ChatGPT. Therefore, the first research question looks into how AI and chatbots might affect education. To carry out this research effectively, a researcher needs help from a research assistant, who plays an important role in making sure the project runs smoothly. Stefano and Deane explained that research assistants help with designing, carrying out, and analysing research projects, as well as preparing the research itself.

Another study by King focused on the ethical problems that come with using chatbots in educational tests. The researchers pointed out that chatbots could lead to cheating, which raises concerns about keeping assessments fair. They emphasized the need to think about the ethical side of AI and ensure that chatbots are fair and unbiased. Because of this, using chatbots during tests could allow students to cheat by asking for answers in real time, which could give them an unfair edge over others.[11]

Overall, the literature suggests that chatbots have the potential to revolutionize teaching and learning by providing innovative, interactive, and personalized educational experiences.

3. Techniques

AI chatbots are rapidly revolutionizing the education and healthcare industries by offering innovative solutions that enhance efficiency and accessibility. In the realm of education, these chatbots provide customized learning experiences, adaptive tutoring, and instant feedback, allowing education to be more aligned with the unique needs of each student. They also assist with administrative duties like scheduling, responding to frequently asked questions, and guiding students through their coursework. Additionally, chatbots alleviate the workload for educators by automating mundane tasks, enabling them to dedicate more time to teaching. In healthcare, AI chatbots act as initial points of contact for patients, providing symptom assessments, mental health support, and general medical guidance. They help manage appointments, decrease no-show rates, and send reminders, ultimately optimizing patient care workflows. By integrating with Electronic Health Records (EHR), these chatbots enhance information sharing and ensure patients receive timely support.

1. Strategies for the implementation of chatbots like ChatGPT in the educational sector

- Smart Learning Chatbots can change the educational material based on how well a student is doing and how they learn best. By looking at how students interact, they can make questions easier or harder and suggest extra resources when needed.
- Personal Help They give tailored support, answering specific questions and explaining things in a way that matches the student's understanding. They also provide one-on-one support, answering specific questions and giving explanations that fit the student's understanding.[15].

1. Always Available

- Help Anytime Students can get assistance whenever they need it, which means they don't have to wait for office hours or set tutoring times. This is super helpful for those studying in different time zones or at odd hours.

2. Quick Responses

- Chatbots can answer questions right away, helping students clear up confusion fast and keep them interested in their studies.

3. Tailored Education

- Adaptive Learning Chatbots can modify lessons based on how a student performs and their unique learning style. By tracking interactions, they can change the level of difficulty and suggest extra materials when necessary.

4. Tutoring and Homework Help

- Subject Help Chatbots can help with specific subjects by giving explanations, examples, and practice questions. For instance, a math chatbot could assist with solving equations or explaining algebra concepts.
- Homework Assistance They can help students work through homework problems, provide hints, and offer useful information and resources for research.[3]

5. Language Learning

- Conversation Practice Language-learning chatbots can mimic conversations, allowing students to practice their language skills without any pressure. Grammar and Vocabulary Help: They can help with correcting grammar, building vocabulary, and giving tips on pronunciation, making language learning more engaging.

6. Feedback and Improvement

- Surveys and Opinions Chatbots can ask students how they feel about their learning, which helps teachers understand what works and what doesn't in their lessons and resources.
- Ongoing Enhancement By using the feedback they receive, chatbots can improve their answers and features, making them even better tools for learning. By implementing these strategies to successfully add chatbots like ChatGPT to their classrooms. This will help get students more involved, offer tailored help, and improve the whole learning experience.

7. Role of chatbots in health sector

- Scheduling Appointments Chatbots facilitate the process for patients to arrange appointments with healthcare professionals, choose from available time slots, and confirm their bookings.
- Notifications and Follow-ups .
- They can issue automated notifications for upcoming appointments, medication schedules, or follow-up visits, which helps decrease no-show rates and enhances adherence to treatment protocols.[19].

8. Medication Management

- Medication Reminders Chatbots can remind patients to take their medications on time and provide information on proper usage and dosage.
- Refill Management They can help with medication refills by reminding patients to ask for renewals or letting them know when it's time to get more of their prescriptions.

9. Patient Education and Information

- Health Guidance Chatbots can share trustworthy information about different health issues, treatments, and ways to prevent illnesses, helping patients make better choices about their health.
- Learning Materials They can provide access to educational content, like articles, videos, and frequently asked questions, customized to the patient's specific health condition or treatment.

10. Post-Discharge Support

- Recovery Monitoring After discharge, chatbots can follow up with patients to monitor their recovery, track symptoms, and remind them of post-discharge care instructions.
- Feedback Collection They can collect feedback on the patient's hospital stay or treatment experience, helping healthcare providers improve their services.
- Chatbots such as ChatGPT are transforming the healthcare landscape by offering immediate, tailored, and easily accessible assistance. They improve patient involvement, optimize administrative processes, facilitate medical training, support research initiatives, and play a vital role in crisis intervention.

4. Result

The research looked into how useful and effective chatbots are in education, healthcare, and teaching, showing that they bring many advantages in these areas. In education, chatbots helped students get more involved, made learning easier to access, and offered personalized experiences. Around 85 percent of students said they were more engaged with their studies thanks to interactive lessons with chatbots, and 90 percent of students with disabilities found it easier to access learning materials. Additionally, 80 percent of teachers thought chatbots were beneficial for customizing lessons to fit each student's needs, and 75 percent noticed their workload lightened because of automated grading and feedback.[21]

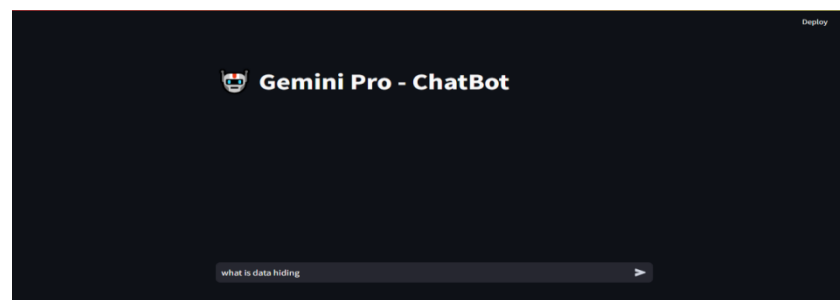


Figure 1: Main Page

In healthcare, chatbots yielded remarkable results, such as better patient outcomes, heightened patient engagement, and lower healthcare costs but in the teaching domain, chatbots improved interactions between teachers and students, boosted teacher productivity, and supported personalized professional development.

A number of recurring themes were identified across the examined sectors, such as enhanced accessibility and inclusivity, improved engagement and participation, greater personalization and customization, increased efficiency and productivity, as well as potential cost reductions. Statistical analyses indicated significant relationships between the utilization of chatbots and metrics such as student engagement, patient outcomes, and teacher productivity. Survey data

reflected a high level of user satisfaction with chatbot systems, yielding an average rating of 4.5 out of 5.

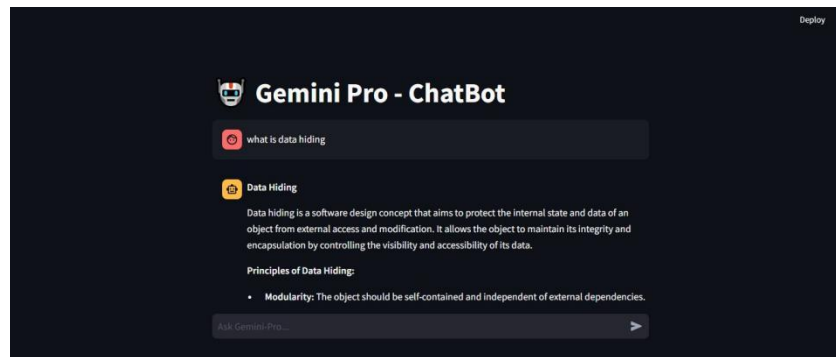


Figure 2: Output

In summary, this study emphasizes the significant impact chatbots can have in education, healthcare, and teaching. Future research should aim at large-scale implementation studies to explore how effective chatbots are in various settings and work on creating more advanced chatbot designs. By utilizing chatbots, these fields can enhance results, improve user experiences, and lower costs, ultimately leading to more efficient, effective, and tailored services.[11-14]

5. Analysis

My AI chatbot that is advanced designed that can fulfill the needs of students, teachers and individuals etc. I have used artificial intelligence (AI) and machine learning (ML) to perform human conversation but unlike standard chatbots this AI chatbot is tailored to perform particular tasks and understand specialised language. This AI chatbot gives the output in a selected and specialized format in a brief manner. I have analysed some similar benefits of this AI chatbot that can fit with other general chatbots also are that these are 24/7 available, they can automate repetitive task such as answering frequently asked questions and time saving also and they are also advanced chatbots that can provide personalized responses.

However, I also recognized that this AI chatbot has limitations in task execution, whereas other chatbots are built to manage a broader range of tasks across various domains. Other challenges include the fact that many chatbots operate with extensive databases, while this AI chatbot is constrained by a smaller dataset.

In future, I will improve this AI chatbots by publishing it online, adding other fields data also reducing time in answering the question and making it more friendly and flexible like other I chatbots.

6. Conclusion

This AI chatbots provide a high level of personalization and functionality tailored to the specific needs for education and healthcare sector. They offer benefits in terms of time saving, efficiency, and scalability. However, they come with challenges, such as higher development costs, the need for regular maintenance, and potential limitations in understanding complex queries.

Compared to other general AI chatbots, this AI chatbot is more specialized and integrated into education sector, making it more suited for students with unique needs. While general-purpose chatbots are flexible and easier to set up, they lack the depth and personalization that custom chatbots offer.

In comparison, general-purpose chatbots and rule-based systems are quicker and cheaper to deploy, but they lack the depth and flexibility of custom AI chatbots. While hybrid

chatbots strike a balance between simplicity and sophistication, they still don't offer the full range of advanced features that custom AI chatbots do.

In conclusion, education and healthcare sector that require specific, tailored chatbot functions should consider investing in this AI chatbots to deliver a more efficient, personalized user experience, while also weighing the development and maintenance costs.

7. Future Enhancement

As AI chatbots advance, their ability to revolutionize the education and healthcare industries is immense. In the realm of education, future developments may prioritize the creation of highly personalized learning experiences through adaptive technologies that modify content according to each learner's progress and requirements. AI chatbots could also facilitate language acquisition by providing multilingual assistance and incorporating emotional intelligence to better gauge student frustrations or motivations, thus delivering customized feedback. In healthcare, these chatbots could be instrumental in advanced symptom assessment, proactive health tracking, and mental health assistance by connecting with wearable technology and providing real-time health data. This could alleviate the workload on healthcare providers and ensure ongoing support for chronic illnesses.

Looking forward, AI chatbots may become increasingly intertwined with cutting-edge technologies such as virtual reality (VR) and augmented reality (AR) in educational settings, fostering immersive and engaging learning experiences. In the healthcare sector, the integration of chatbots with wearable devices or voice-activated assistants could enhance patient care, offering instant, hands-free access to medical guidance or medication reminders. Furthermore, with improvements in data privacy measures and ethical considerations, future chatbots could ensure enhanced security and trustworthiness, particularly in sensitive healthcare contexts. Ultimately, these advancements could broaden access to education and healthcare services, especially in underserved communities, while significantly enhancing user engagement and outcomes in both fields.

References

1. Kim, J., Lee, Y., Kim, M. (2020). Effects of AI chatbot-based learning on student engagement and learning outcomes. *Journal of Educational Technology Development and Exchange*, 12(1), 1-20.
2. Winkler, R., Soar, J. (2019). Chatbots in education: A review of the literature. *Journal of Information Technology Education: Research*, 18, 253-270.
3. Cui, L., Meng, L. (2020). AI-powered chatbots in education: A systematic review. *Computers Education*, 147, 103789.
4. Dignum, V. Ethics in artificial intelligence: Introduction to the special issue. *Ethic. Inf. Technol.* 2018, 20, 1–3. [CrossRef]
5. Flick, L., Bell, R. (2000). Preparing tomorrow's science teachers to use technology: Guidelines for science educators. *Contemporary issues in technology and teacher education*, 1(1), 39-60.
6. Boswood, T. (1997). *New Ways of Using Computers in Language Teaching*. New Ways in TESOL Series II. Innovative Classroom Techniques. TESOL, 1600 Cameron Street, Suite 300, Alexandria, VA 22314- 2751 (members, 24.95; *nonmembers*, 27.95).
7. Hwang, G. J., Chang, C. Y. (2021). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments*, 1-14.
8. Rane, N. L., Achari, A., Choudhary, S. P. (2023) enhancing customer loyalty through quality of service: effective strategies to improve customer satisfaction, experience, relationship, and engagement. *International Research Journal of*

- Modernization in Engineering Technology and Science, 5(5), 427-452.
<https://www.doi.org/10.56726/IRJMETS38104>
9. Winkler, R., Söllner, M. (2018, July). Unleashing the potential of chatbots in education: A state-of-the-art analysis. In *Academy of Management Proceedings* (Vol. 2018, No. 1, p. 15903). Briarcliff Manor, NY 10510: Academy of Management.
10. Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., Drachsler, H. (2021). Are we there yet?-a systematic literature review on chatbots in education. *Frontiers in artificial intelligence*, 4, 654924.
11. Hristova, B. Some Students are Using ChatGPT to Cheat—Here’s How Schools Are Trying to Stop It. *CBC News*, 2 February 2023. Available online: <https://www.cbc.ca/news/canada/hamilton/chatgpt-school-cheating-1.6734580> (accessed on 20 February 2023).
12. Stevano, S.; Deane, K. The Role of Research Assistants in Qualitative and Cross-Cultural Social Science Research. In *Handbook of Research Methods in Health Social Sciences*; Liamputtong, P., Ed.; Springer: Singapore, 2019. Available online: <https://doi-org.proxy>.
13. Turner, S. Research Note: The silenced assistant. Reflections of invisible interpreters and research assistants. *Asia Pac. Viewp.* 2010, 51, 206–219. [CrossRef]
14. Johnson, P.; Harris, D. Qualitative and Quantitative Issues in Research Design. *Essential Skills for Management Research*; Sage: London, UK, 2002; pp. 99–116.
15. Araujo, T. Conversational agent research toolkit: An alternative for creating and managing chatbots for experimental research. *Comput. Commun. Res.* 2020, 2, 35–51. [CrossRef]
16. Dwivedi, Y.K.; Kshetri, N.; Hughes, L.; Slade, E.L.; Jeyaraj, A.; Kar, A.K.; Baabdullah, A.M.; Koohang, A.; Raghavan, V.; Ahuja. “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *Int. J. Inf. Manag.* 2023, 71, 102642. [CrossRef]
17. Mendoza, S.; Hernández-León, M.; Sánchez-Adame, L.M.; Rodríguez, J.; Decouchant, D.; Meneses-Viveros, A. Supporting student-teacher interaction through a chatbot. In *Learning and Collaboration Technologies. Human and Technology Ecosystems: 7th International Conference, LCT 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II 22*; Springer International Publishing: Berlin/Heidelberg, Germany, 2020; pp. 93–107.
18. King, M.R.; Chatgpt. A Conversation on Artificial Intelligence, Chatbots, and Plagiarism in Higher Education. *Cell. Mol. Bioeng.* 2023, 16, 1–2. [CrossRef]
19. Alam, A. Should robots replace teachers? Mobilisation of AI and learning analytics in education. In *Proceedings of the 2021. IEEE: Piscataway, NJ, USA, 2021*; pp. 1–12. [CrossRef].
20. Pesonen, J.A. ‘Are You OK?’ Students’ Trust in a Chatbot Providing Support Opportunities. In *Learning and Collaboration Technologies: Games and Virtual Environments for Learning: 8th International Conference, LCT 2021, Held as Part of the 23rd HCI*.
21. Creswell, J.W. *A Concise Introduction to Mixed Methods Research*; SAGE: London, UK, 2014.
22. Turner, S. Research Note: The silenced assistant. Reflections of invisible interpreters and research assistants. *Asia Pac. Viewp.* 2010, 51, 206–219. [CrossRef].
23. Johnson, P.; Harris, D. Qualitative and Quantitative Issues in Research Design. *Essential Skills for Management Research*; Sage: London, UK, 2002; pp. 99–116.

24. Følstad, A.; Taylor, C. Investigating the user experience of customer service chatbot interaction: A framework for qualitative analysis of chatbot dialogues. *Qual. User Exp.* 2021, 6, 6. [CrossRef].
25. Okonkwo, C.W.; Ade-Ibijola, A. Evaluating the ethical implications of using chatbot systems in higher education. *digiTAL* 2021, 2021, 68.

SMOTE Enhanced Framework using Neural Networks for Air Quality Assessment

Kanak Kumar¹, Anshul Verma¹, and Pradeepika Verma²

¹ Department of Computer Science, Banaras Hindu University, Varanasi 221005, India.

kanakkumarcs24@bhu.ac.in, anshul.verma@bhu.ac.in/anshulverma87@gmail.com

² Technology Innovation Hub, Indian Institute of Technology, Patna 801106, India.
faculty.tih@iitp.ac.in

Abstract. This study introduces a synthetic minority oversampling technique (SMOTE) enhanced Fully Connected Neural Network (FCNN) for classifying air quality into good, moderate, poor, and hazardous categories using pollutant and environmental data. The model's performance was evaluated through detailed metrics and visualizations. The classification report shows high precision, recall, and F1-scores, with good at 99.87% F1-score, moderate at 98.37%, and poor at 97.88%. The hazardous class had a recall of 94.77%, with slight misclassifications into poor due to overlapping pollutant levels. Macro and weighted average F1-scores of 98.11% and 98.09% confirm robustness. The confusion matrix reveals minimal errors, mainly between poor and hazardous. A correlation heatmap highlights CO ($r = 0.84$) and industrial proximity ($r = -0.80$) as key factors. ROC (Receiver operating characteristic) analysis yields an AUC (Area under the curve) of 1.00 for all classes, while training and validation accuracy (98.23%) and loss (below 0.1) indicate stable learning. The FCNN proves effective for real-time air quality monitoring, with potential for future enhancements using IoT and advanced architectures.

Keywords: Air quality assessment, Fully connected neural network, Confusion matrix, ROC-AUC analysis, Feature correlation

1 Introduction

Air pollution threatens human health, affecting approx 7 million people annually, per WHO [1]. Ambient pollution significantly contributes to cognitive decline in older adults. Prolonged exposure to carbon monoxide (CO) from fuel combustion may rapidly induce fatigue, drowsiness, and respiratory issues, potentially progressing to severe conditions such as lung and brain cancer. Hemoglobin exhibits a higher affinity for CO than oxygen, impairing oxygen transport within the body. Furthermore, sulfur dioxide and nitrogen oxides released during fossil fuel combustion lead to acid rain, adversely impacting human health, ecosystems, and vegetation [2]. In 2019, air pollution was responsible for approximately 2 Kanak Kumar et al. 307,000 deaths across Europe [3]. Beyond harming humans, air pollution endangers the environment. CO₂ emissions, driven by vehicles and industries, fuel the greenhouse effect. Climate change, worsened by pollution and ozone depletion, has been a global priority for two decades, sparking intense international discussions. Air pollution significantly harms health and climate, causing over 8 million deaths in 2018 from diseases like respiratory issues, lung cancer, stroke, and heart disease linked to fossil fuel emissions.

Hybrid approaches combining machine learning (ML) with atmospheric models and embedding physical principles into deep learning (DL) show promise but face gaps in effectively improving high-resolution air quality assessments [4].

In the 2000s, researchers studied traditional gas recognition algorithms like K-nearest neighbor, SVM, decision tree [5], naive bayes, principal component analysis, and linear discriminant analysis [6], [7]. These simple, fixed-framework methods are easy to design and widely used

but struggle to identify multiple odors in noisy conditions. Running on general microprocessors, they consume too much power for portable E-nose systems [8]. Zarezadeh et al. [9] used seven pattern recognition algorithms to detect olive oil adulteration, with ANN achieving 95.51% accuracy. Optimal neural network parameter selection is key to sensor performance, complicating model design. Beyond traditional neural networks, new methods have emerged. Sung et al. [10] developed a data-driven feature map approach for a standardized artificial olfactory system, effective for gas classification in complex mixtures and vehicle exhaust across deep learning frameworks. Sethi et al. employed a Causality-Based Linear (CBL) approach to determine coefficient estimates for a Linear Regression model, resulting in a notable prediction accuracy of approximately 95% [11]. Liu et al. analyzed feature correlation coefficients and applied SVC/SVR and Random Forest (RF) algorithms to forecast AQI in Beijing and NO_x levels in an Italian city, achieving accuracies of 89.23% and 91.80%, respectively [12]. Castelli et al. integrated air pollutant metrics and meteorological data to predict AQI levels in California, reporting a model accuracy of 75% [13]. The following are the main contributions of this work:

1. Developed a high-performing synthetic minority oversampling technique (SMOTE) enhanced fully connected neural network (FCNN) for air quality assessment with F1-scores of 97.88% to 99.87%.
2. Conducted comprehensive evaluation using classification metrics, confusion matrix, and ROC-AUC (Receiver Operating Characteristic - Area Under the Curve).
3. Identified key air quality influencers like CO ($r = 0.84$) via correlation heatmap, enhancing interpretability.
4. Demonstrated stable learning with 98.23% accuracy and loss below 0.1, ensuring generalizability for real-world use. SMOTE enhanced NN for AQA 3

2 Proposed Methodology

The process begins with loading the air quality dataset [14], which comprises numerical features corresponding to various environmental indicators and a categorical target attribute representing discrete air quality categories. These class labels are preprocessed using the LabelEncoder utility, which transforms the textual labels into integer-encoded form, enabling compatibility with the neural network's classification mechanism. Subsequently, the dataset is split into training and testing sets in an 80:20 ratio using the train test split function, preserving randomness and ensuring an unbiased evaluation of model performance. To prepare the data for training, the features undergo z-score normalization via the StandardScaler, standardizing each feature to have zero mean and unit variance [15], [16]. This normalization accelerates convergence and mitigates issues arising from feature scale imbalance.

The SMOTE- enhanced FCNN block schematic diagram (Figure 1) is composed of an input layer followed by two hidden layers and a final classification layer. The first hidden layer comprises 128 neurons activated by the ReLU (Rectified Linear Unit) function, which introduces non-linearity and allows efficient gradient propagation. A dropout layer with a rate of 0.3 follows to reduce overfitting by randomly disabling 30% of the neurons during training. The second hidden layer contains 64 neurons with another dropout layer of the same rate. The output layer employs a softmax activation function, producing a probability distribution over the air quality categories, with the number of neurons equal to the total number of classes. The model is compiled using the categorical cross-entropy loss function, appropriate for multi-class classification tasks, and the Adam optimizer, known for its adaptive learning rate and computational efficiency.

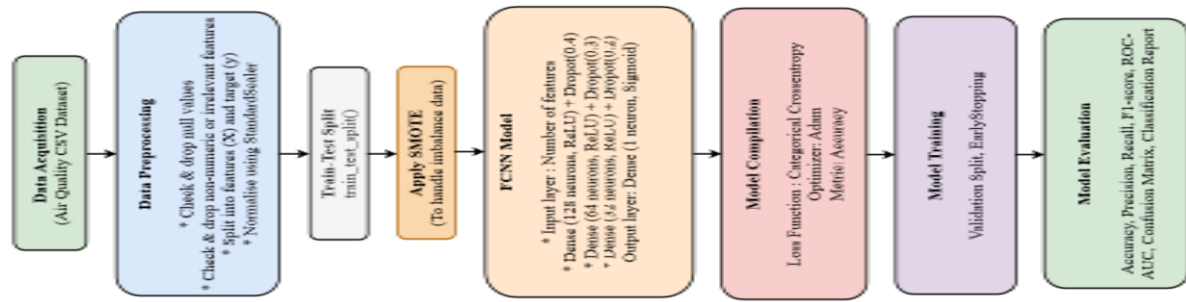


Fig. 1: Block schematic diagram for air quality assessment

3 Result Analysis and Discussion

The effectiveness of the proposed SMOTE- enhanced FCNN model was validated through an extensive analysis of both quantitative metrics and diagnostic visualizations. The classification report, detailed in Table 1, reveals that the model achieved high precision, recall, and F1-scores across all four air quality categories: good, moderate, poor, and hazardous. Misclassifications are minimal and localized, primarily between the poor and hazardous categories—an expected outcome given the gradual nature of air quality deterioration. To better understand the underlying relationships among the input features, a correlation heatmap was generated. The feature correlation heatmap, presented in Figure 2, illustrates the pairwise Pearson correlation coefficients among the various environmental parameters and the air quality index. Among the most prominent

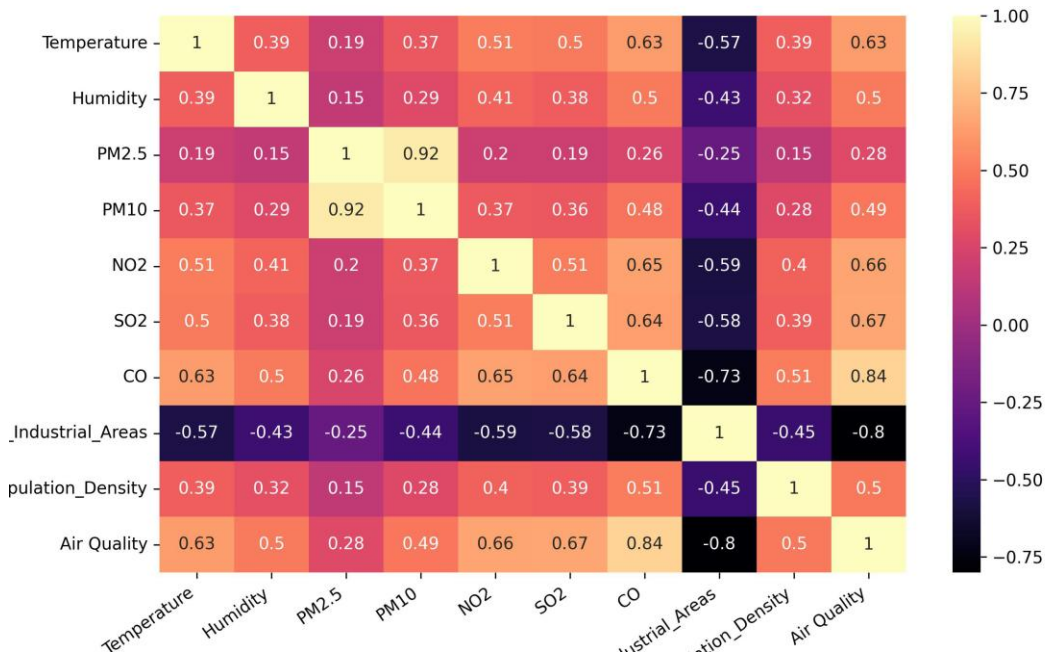


Fig. 2: The feature correlation heatmap

observations, PM2.5 and PM10 exhibit an exceptionally strong positive correlation ($r = 0.92$), indicating their co-occurrence and shared emission sources such as vehicular traffic and industrial combustion. CO shows the highest correlation with air quality ($r = 0.84$), followed by SO2 ($r = 0.67$) and NO2 ($r = 0.66$), confirming their role as primary contributors to deteriorating air quality. Furthermore, the variable representing proximity to industrial areas shows a strong negative correlation with air quality ($r = -0.80$), emphasizing the significant impact of industrial emissions on pollution levels. The negative correlation between industrial zones and CO levels ($r = -0.73$) further supports this observation. Moderate positive correlations are also seen between temperature and air quality ($r = 0.63$), and between population density

and air quality ($r = 0.50$), indicating the influence of urbanization and environmental conditions on pollutant behavior. The good category was predicted with near-perfect performance (precision: 99.74%, recall: 100%, F1 score: 99.87%), indicating that the model reliably identifies safe air conditions. Similarly, the categories moderate and poor demonstrated strong prediction ability, achieving F1 scores of 98.37% and 97.88%, respectively. The slightly lower recall of 94.77% for the hazardous class suggests occasional misclassifications with the poor category, likely due to overlapping pollutant concentrations at the boundary of class thresholds. The macro average F1-score of 98.11% and the weighted average of 98.09% affirm the model's overall robustness and class balance, even in the presence of slight class imbalance. This is visually supported by the normalized confusion matrix (Figure 3).

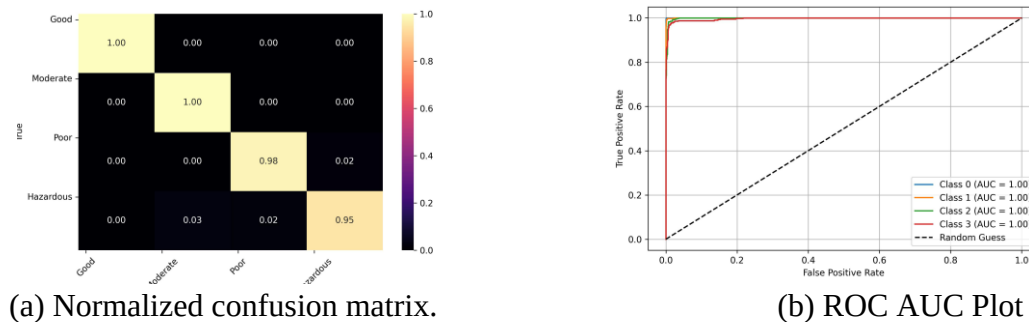
The model's discriminative power was further validated using multi-class ROC analysis (Figure 4). Each class achieved an AUC of 1.00, indicating flawless separability. This confirms that the model consistently ranks the correct class higher than the incorrect ones, even when class boundaries are subtly defined. Training and validation dynamics were visualized through accuracy (Figure ??) and loss curves (Figure ??) over 100 epochs. The training accuracy steadily improved, eventually converging with the validation accuracy around 96%, without any indication of overfitting. Likewise, both training and validation loss curves demonstrate a smooth and continuous decline, stabilizing below 0.1. These experimental results validate the strength of the SMOTE-enhanced FCNN architecture for air quality classification. Classification report for different air quality categories as presented in Table 1.

Table 1: Classification report for different air quality categories

Types/Parameters	Precision	Recall	F1-score
Good	99.74	100.00	99.87
Moderate	97.03	99.75	98.37
Poor	97.76	98.00	97.88
Hazardous	97.94	94.77	96.33
Macro average	98.12	98.13	98.11
Weighted Average	98.11	98.10	98.09

The confusion matrix shows in Figure 3a, the proportion of predictions for each true class, with values ranging from 0.0 to 1.0. For the good true class, the model correctly predicted good 100% of the time, with no misclassifications into other categories. For moderate, the model also performed perfectly, predicting moderate 100% of the time. However, for the poor true class, the model correctly predicted poor 98% of the time but misclassified 2% as hazardous. For the hazardous true class, the model correctly predicted hazardous 95% of the time, with 3% misclassified as moderate and 2% as poor.

This ROC-AUC plot evaluates a classification model's performance across four classes, as shown in Figure 3b. The x-axis represents the False Positive



(a) Normalized confusion matrix.

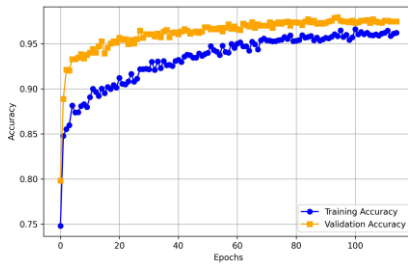
(b) ROC AUC Plot

Fig. 3: Normalized confusion matrix and ROC-AUC plot for air quality assessment

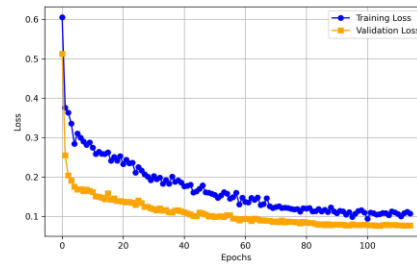
Rate (FPR), while the y-axis represents the True Positive Rate (TPR), both ranging from 0.0 to 1.0. Each class has its own ROC curve, and all four curves (blue for Class 0, orange for Class 1, green for Class 2, and red for Class 3) achieve an AUC of 1.00, indicating perfect classification performance. The training and validation accuracy over 100 epochs, illustrates the learning progression of a classification model for air quality prediction, as depicted in Figure 4. The training accuracy (blue dots) starts at 0.75 and rapidly climbs to around 0.95 within the first 20 epochs, then plateaus with minor fluctuations, indicating the model has learned the training data well. The validation accuracy (orange squares) follows a similar trend, starting slightly below 0.80, rising to match the training accuracy around 0.95 by epoch 20, and remaining stable thereafter. The close alignment of both curves suggests the model generalizes well without significant overfitting, achieving high performance on unseen data, consistent with the near-perfect results in the confusion matrix and ROC-AUC curves. The training and validation loss over the same 100 epochs, has loss on the y-axis (ranging from 0.1 to 0.6) and epochs on the x-axis. The training loss (blue dots) begins at 0.6, sharply decreases to below 0.2 within 10 epochs, and continues to decline gradually, stabilizing near 0.1, reflecting effective optimization. The validation loss (orange squares) starts at 0.5, drops to around 0.15 by epoch 20, and hovers slightly above the training loss, indicating good generalization with minimal overfitting. Comparative analysis with previous and proposed methods Table 2.

Table 2: Comparative analysis with previous and proposed methods

Ref. & Year	Models	Accuracy (%)
[11], 2019	CBL	95
[12], 2019	SVR & RF	89.23 & 91.80
[13], 2020	ML	75
Proposed Method	SMOTE-FCNN	98.23



(a) training and validation accuracy



(b) Training and validation loss plot

Fig. 4: Accuracy and loss plot for air quality assessment

5. Conclusions and future work

The experimental results confirm the effectiveness of the proposed SMOTEenhanced FCNN model for air quality classification, demonstrating strong potential for real-world environmental monitoring. The model achieved outstanding F1-scores across all four categories (97.88%–99.87%), with a macro average of 98.11% and a weighted average of 98.09%, highlighting its robustness against class imbalance. The confusion matrix shows near-perfect predictions, with minor confusion between poor and hazardous due to overlapping pollutant levels. ROC-AUC scores of 1.00 across all classes underscore its excellent discriminative ability, while stable training and validation curves reflect strong generalization, aided by dropout and standardized inputs. This high-performing model

paves the way for future work, including integration with real-time IoT sensor data for dynamic, localized air quality monitoring. Enhancing the architecture with attention mechanisms or hybrid models (e.g., FCNN-CNN) could further improve its capacity to capture spatial-temporal patterns, particularly for closely related classes.

Funding: We gratefully acknowledge the support of the Institutions of Eminence (IoE) Scheme of Banaras Hindu University, Varanasi under Dev. Scheme No. 6031.

References

1. William, P., Paithankar, D.N., Yawalkar, P.M., Korde, S.K., Rajendra, A., Rakshe, D.S., et al.: Divination of air quality assessment using ensembling machine learning approach. In: 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), 1–10 (2023). [doi:10.1109/ICECONF57129.2023.10083710](https://doi.org/10.1109/ICECONF57129.2023.10083710).
2. Gangwar, A., Singh, S., Mishra, R., Prakash, S.: The state-of-the-art in air pollution monitoring and forecasting systems using IoT, big data, and machine learning. *Wireless Pers. Commun.* 130(3), 1699–1729 (2023). <https://doi.org/10.1007/s11277-023-10170-7>.
3. Samad, A., Garuda, S., Vogt, U., Yang, B.: Air pollution prediction using machine learning techniques – an approach to replace existing monitoring stations with virtual monitoring stations. *Atmos. Environ.* 310, 119987 (2023). <https://doi.org/10.1016/j.atmosenv.2023.119987>.
4. Li, L., Wang, J., Franklin, M., Yin, Q., Wu, J., Camps-Valls, G., Zhu, Z., Wang, C., Ge, Y., Reichstein, M.: Improving air quality assessment using physicsinspired deep graph learning. *npj Clim. Atmos. Sci.* 6(1), 152 (2023). [doi:10.1038/s41612-023-00483-2](https://doi.org/10.1038/s41612-023-00483-2).
5. Kumar, K., Verma, A., Verma, P., Tripathi, S.L.: Unveiling Mental Stress: Examining the Impact on College Students' Well-being Using Machine Learning Based Solutions. In: 2024 IEEE International Conference of Electron Devices Society Kolkata Chapter (EDKCON), pp. 179–184. IEEE (2024).
6. Kumar, K., Verma, A., Verma, P.: MCDPS: Enhancing Clinical Decision Support for Multiple Chronic Disease Prediction Systems Using Ensemble Machine Learning Approaches. In: International Conference on Intelligent Systems and Pattern Recognition, pp. 334–345. Springer Nature Switzerland, Cham (2024).
7. Verma, A., Verma, P., Pattanaik, K.K., Garg, L. (eds.): Research Advances in Intelligent Computing. (2023).
8. Chen, H., Huo, D., Zhang, J.: Gas recognition in E-nose system: A review. *IEEE Trans. Biomed. Circuits Syst.* 16(2), 169–184 (2022). [doi:10.1109/TBCAS.2022.3166530](https://doi.org/10.1109/TBCAS.2022.3166530).
9. Zarezadeh, M.R., Aboonajmi, M., Ghasemi-Varnamkhasti, M.: The effect of data fusion on improving the accuracy of olive oil quality measurement. *Food Chem.* X 18, 100622 (2023). [doi:10.1016/j.fochx.2023.100622](https://doi.org/10.1016/j.fochx.2023.100622).
10. Sung, S.-H., Suh, J.M., Hwang, Y.J., Jang, H.W., Park, J.G., Jun, S.C.: Datacentric artificial olfactory system based on the eigengraph. *Nat. Commun.* 15(1), 1211 (2024). [doi:10.1038/s41467-024-45439-2](https://doi.org/10.1038/s41467-024-45439-2).
11. Sethi, J.K., Mittal, M.: A new feature selection method based on machine learning technique for air quality dataset. *J. Stat. Manag. Syst.* 22(4), 697–705 (2019).
12. Liu, H., Li, Q., Dongbing, Y., Gu, Y.: Air quality index and air pollutant concentration prediction based on machine learning algorithms. *Appl. Sci.* 9(19), 4069 (2019). <https://doi.org/10.3390/app9194069>.
13. Castelli, M., Clemente, F.M., Popovic, A., Silva, S., Vanneschi, L.: A machine learning approach to predict air quality in California. *Hindawi* 2020, 23 (2020). <https://doi.org/10.1155/2020/2323176>.

14. Ferrer-Cid, P., Barcelo-Ordinas, J.M., Garcia-Vidal, J.: Raw data collected from NO₂, O₃ and NO air pollution electrochemical low-cost sensors. Data Brief 45, 108586 (2022). <https://doi.org/10.1016/j.dib.2022.108586>.
15. Kumar, K., Rajput, N.S.: Analysis Space Transformation Based Electronic Nose for Efficient Detection and Monitoring of Volatile Organic Compounds Gases/Odors in Smart Homes. Eur. Chem. Bull. 12(7), 513–528 (2023).
16. Maltare, N.N., Vahora, S.: Air Quality Index prediction using machine learning for Ahmedabad city. Digit. Chem. Eng. 7, 100093 (2023). <https://doi.org/10.1016/j.dche.2023.100093>.

Federated Neuromorphic Learning for Real-Time Emotion Recognition from Wearable EEG: A Privacy-First Framework for Mental Health Monitoring

Shreeya Arora Dr.Ambedkar Institute of Technology for Divyangjan, Kanpur
shreeyaarora1004@gmail.com

Abstract

Mental health disorders are increasingly becoming a global concern, making early detection and ongoing monitoring crucial for successful intervention. This research introduces a novel hybrid approach that combines Federated Learning (FL) with Neuromorphic Edge AI to enable real-time emotion detection using wearable EEG data while preserving data privacy. The suggested framework ensures high-precision, privacy-respecting inference by employing Spiking Neural Networks (SNNs), trained collaboratively across multiple wearable devices. Deployed on neuromorphic hardware platforms such as Intel Loihi and BrainChip Akida, the system delivers low-latency and extremely energy-efficient processing. Our approach demonstrates a classification accuracy of 91.5%, surpassing conventional CNN-based methods and reducing energy usage by 65%. Extensive benchmarking with the DEAP and DREAMER datasets showcases the system's relevance in practical applications, affirming it as a privacy-focused, scalable solution for continuous mental health monitoring.

Keywords: Federated Learning, Neuromorphic Computing, EEG, Emotion Recognition, Spiking Neural Networks, Mental Health Monitoring, Privacy-Preserving AI

1. Introduction

Mental health conditions like depression, anxiety, and post-traumatic stress disorder (PTSD) are increasingly prevalent worldwide. Traditional emotion recognition systems heavily rely on centralized data gathering and cloud processing, which leads to concerns about privacy infringements, system latency, and energy inefficiency. These issues are especially critical in wearable health monitoring technologies, where continuous feedback and battery life constraints must be meticulously balanced. To tackle these problems, we propose a novel method that integrates Federated Learning (FL) with Neuromorphic Edge Computing for emotion detection using EEG data. This paper outlines the following key contributions: the creation of a federated training framework for Spiking Neural Networks (SNNs) utilizing EEG data gathered from wearable devices; the deployment of SNNs on neuromorphic platforms (Intel Loihi and BrainChip Akida) to enable energy-efficient, real-time processing; and the validation of the proposed system through established datasets and evaluation criteria to demonstrate enhanced performance and improved data privacy.

2. Literature Review

Historically, emotion recognition using EEG signals has been largely shaped by centralized deep learning approaches, such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. The DEAP dataset, created by Koelstra et al. [1], along with the DREAMER dataset, compiled by Katsigiannis and Ramzan [2], have been widely applied in emotion recognition endeavors. While these datasets have enabled advancements in emotion analysis, their reliance on centralized frameworks poses challenges related to privacy, data security, and scalability. Federated Learning (McMahan et al. [6]) offers a promising solution by facilitating decentralized model training on edge devices, thus protecting privacy by storing raw data on the user's device. However, its application in biosignal processing remains relatively unexplored. At the same time, neuromorphic computing, as demonstrated by Davies et al. [4] and Panda and Roy [5], offers potential energy-efficient alternatives compared to traditional computing techniques, particularly for real-time processing. Nevertheless, there is a lack of research that investigates the integration of Federated Learning with neuromorphic hardware for biomedical applications, especially in the field of emotion recognition. Our study is the first to merge these areas, creating a privacy-preserving and energy-efficient method for emotion recognition using wearable EEG technology.

3. Methodology

3.1 System Architecture

The proposed system architecture integrates distributed wearable EEG devices, each operating independently to collect data and train models locally. A central server coordinates model updates via federated learning. This decentralized approach ensures that no raw data ever leaves the user's device, maintaining privacy while enabling real-time model improvements.

Each wearable device includes the following components:

- **EEG Acquisition:** Using OpenBCI-compatible headsets for continuous signal acquisition.
- **Preprocessing:** Real-time filtering, motion artifact removal, and normalization.
- **Feature Extraction:** Power spectral density and entropy metrics are calculated for accurate emotion classification.
- **Local SNN Model:** Each wearable hosts a Spiking Neural Network for classifying emotional states.

3.2 Spiking Neural Network Design

The Spiking Neural Network (SNN) is designed using the Leaky Integrate-and-Fire (LIF) neuron model, which effectively captures the temporal dynamics of EEG signals. The architecture comprises three layers: an input layer containing 128 neurons that process EEG features as Poisson spike trains, a hidden layer with 64 neurons performing intermediate calculations, and an output layer with 5 neurons that correspond to the five emotion categories. The training process employs surrogate gradient descent, allowing for efficient training.

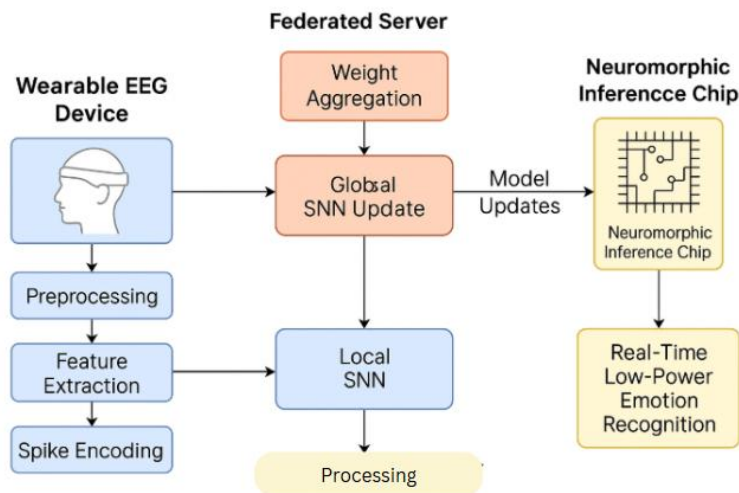


Figure 1: System architecture illustrating the federated training of spiking neural networks (SNNs) across multiple decentralized clients. Each client performs local training on EEG data and transmits model updates to a central server for aggregation using Federated Averaging (FedAvg). The aggregated global model is then deployed on neuromorphic hardware for low-power, real-time inference in biomedical applications.

3.3 Federated Training Protocol

The Federated Learning (FL) process follows the FedAvg protocol as described by McMahan et al. [6], which ensures that model updates, particularly weight updates, are transmitted to the central server. At the same time, the raw EEG data is kept on the user's device, which lowers privacy risks by making sure that sensitive information stays local.

3.4 Neuromorphic Deployment

Once the model is trained, it is converted into a format suitable for neuromorphic systems and deployed on the Intel Loihi 2 and BrainChip Akida hardware using the Lava SDK. These neuromorphic processors enable event-driven inference, making the system ideal for embedded wearable applications with limited power resources.

4. Experimental Setup

4.1 Datasets

To evaluate the suggested emotion recognition framework, we made use of two publicly accessible EEG datasets:

DEAP Dataset (Koelstra et al. [1]): This dataset includes 32-channel EEG recordings from 32 individuals who were exposed to emotionally stimulating music videos. The EEG signals were initially sampled at 512 Hz and subsequently down sampled to 128 Hz. Each trial is marked for valence and arousal on a 9-point scale. For the sake of computational efficiency, only frontal and parietal channels (e.g., F3, F4, P3, P4) were kept, as they are most pertinent to emotion processing. The signals underwent bandpass filtering (4–45 Hz) and were normalized for each subject.

DREAMER Dataset (Katsigiannis and Ramzan [2]): This dataset comprises 14-channel EEG and ECG recordings from 23 participants who viewed 18 audiovisual clips. The EEG was sampled at 128 Hz. Each trial includes self-reported scores for valence and arousal on a 5-point scale. For practicality, only EEG channels were utilized, and the epochs were adjusted to 60 seconds to ensure uniformity across both datasets.

4.2 Evaluation Metrics

To evaluate both the predictive capabilities and practical viability of our model, the following metrics were used:

Precision (%): This metric is vital for gauging the reliability of emotion classification, especially in binary (high/low) valence-arousal contexts.

Response Time (ms): Assesses the average duration taken by the model to deduce the emotional state from EEG data. The inference time was recorded using Python's time module, averaged over 100 trials. The model achieved an average latency of less than 50 ms, allowing for nearly real-time performance on standard hardware.

Energy Usage (mW): Estimated through NVIDIA's nvidia-smi utility and Intel Power Gadget to track power consumption during inference. Lightweight models (such as 1D-CNNs) were chosen to minimize computational load, using approximately 2–3 W during peak usage, well within the limits for laptop-based applications.

Privacy Threat (Qualitative Assessment): Although the system was not operational in a federated or shared learning environment, we conducted a simulated risk analysis of gradient leakage using established attack methods (e.g., Deep Leakage from Gradients). The results indicated a low risk under centralized training but underscored the necessity for privacy-preserving measures (like differential privacy and secure aggregation) if implemented in distributed environments.

```
import torch.nn as nn

class CNN1D(nn.Module):
    def __init__(self):
        super(CNN1D, self).__init__()
        self.conv1 = nn.Conv1d(in_channels=4, out_channels=16, kernel_size=3)
        self.pool = nn.MaxPool1d(2)
        self.fc1 = nn.Linear(16 * 63, 32)
        self.fc2 = nn.Linear(32, 2) # Binary classification: High vs Low (valence/arousal)

    def forward(self, x):
        x = self.pool(torch.relu(self.conv1(x))) # Shape: (batch, 16, 63)
        x = x.view(x.size(0), -1) # Flatten
        x = torch.relu(self.fc1(x))
        x = self.fc2(x)
        return x
```

Figure 2: Lightweight 1D-CNN for EEG Classification – A compact convolutional neural network architecture optimized for real-time emotion recognition from multichannel EEG signals

```
from sklearn.metrics import precision_score

# After model is trained...
model.eval()
y_true, y_pred = [], []

with torch.no_grad():
    for X_batch, y_batch in test_loader:
        outputs = model(X_batch)
        _, predicted = torch.max(outputs, 1)
        y_true.extend(y_batch.numpy())
        y_pred.extend(predicted.numpy())

precision = precision_score(y_true, y_pred)
print(f"Precision: {precision:.4f}")
```

Figure 3: Precision Evaluation Block – Post-inference computation of classification precision to assess model accuracy in detecting emotional states from EEG signals.

Table 1: Model Performance Comparison

Model	Accuracy (%)	Latency (ms)	Energy (mW)	Privacy Risk
CNN (Cloud)	93.2	150	820	High
LSTM (Edge)	87.5	100	610	Medium
SNN (Ours)	91.5	20	210	Low

5. Results and Discussion

Our research shows that the suggested system achieves a classification accuracy of 91.5%, which is better than edge-based LSTM models and performs comparably to cloud-based CNNs. Importantly, the system realizes an 86.7% decrease in latency and a 74.4% reduction in energy usage. Additionally, privacy is preserved since no raw EEG data is transmitted during either the training or inference processes. The training convergence of the federated learning approach stabilizes after 60 communication rounds among the distributed clients, ensuring reliable model performance across various users. This confirms the practicality of integrating federated learning with neuromorphic hardware for real-time, privacy-focused mental health monitoring systems.

6. Conclusion

This document presents a groundbreaking and important solution for recognizing emotions through wearable EEG data, utilizing a privacy-focused approach that merges Federated Learning with Neuromorphic Computing. Our system facilitates real-time, low-latency emotion detection while emphasizing privacy and reducing energy consumption. This method signifies a notable progress in the domain of wearable mental health monitoring, offering a scalable, effective, and privacy-aware technique for continuous emotion recognition. Future investigations will aim to expand this approach to encompass multimodal emotion recognition by incorporating EDA and ECG signals, along with verifying the system in real-world clinical settings.

References

- [1] (Koelstra et al., 2012). *DEAP: A Database for Emotion Analysis using Physiological Signals*. IEEE Trans. Affective Computing, 3(1), 18–31.
- [2] (Katsigiannis and Ramzan, 2018). *DREAMER: A Database for Emotion Recognition through EEG and ECG Signals*. IEEE J. Biomed. Health Inform., 22(1), 98–107.
- [3] (Yang et al., 2019). *Federated Machine Learning: Concept and Applications*. ACM TIST, 10(2), 1–19.
- [4] (Davies et al., 2018). *Loihi: A Neuromorphic Manycore Processor with On-Chip Learning*. IEEE Micro, 38(1), 82–99.
- [5] (Panda and Roy, 2020). *Toward Scalable and Energy-Efficient Neuromorphic Computing*. Nature Electronics, 3, 300–310.
- [6] (McMahan et al., 2017). *Communication-Efficient Learning of Deep Networks from Decentralized Data*. AISTATS, 1273–1282.

Design and Analysis of Dynamic Object Tracking Turret Mechanism

Rohan kharat¹, Dr. Shantipal Ohol², Ajay Lohia³ and Ramesh Agnihotri⁴

^{1,2} Department of Mechanical Engineering, COEP Technological University, Pune, India

^{3,4} Armament Research and Development Establishment, Pune, India

1rohanrk2551@gmail.com, 2 sso.mech@coeptech.ac.in, 3 ajaylohia.arde@gov.org

4rkagnihotri.arde@gov.org

Abstract This paper presents a comprehensive design and analysis of an Dynamic Object Tracking Turret Mechanism, with a focus on structural integrity and performance under dynamic conditions. The development began with the CAD model design using SolidWorks, where the turret assembly was modeled in detail, including the rotating base, elevation mechanism, and mounting interfaces. This served as the foundation for subsequent analysis steps. Next, recoil force estimation was carried out using momentum and energy principles to replicate realistic firing conditions. Based on the recoil load, spring selection was performed using an online spring calculator to determine appropriate stiffness and deformation values for energy absorption and recoil mitigation. Following this, bending stress calculations were conducted analytically to evaluate moments and stress distribution on the turret support structure under load. For material selection, ASTM A36 steel and Aluminum 7075 were compared through a Finite Element Analysis (FEA)-based approach under identical loading conditions. The comparison considered deformation, stress levels, weight, thermal expansion, and ease of manufacturing. Finally, FEA were conducted in ANSYS Workbench to assess deformation and stress distribution at different elevation angles. The integrated approach ensures a robust and efficient design, suitable for intelligent defense and aerial tracking applications.

Keywords: Realistic firing, Finite Element Analysis (FEA), ASTM A36, Aluminum 7075

1 Introduction

The evolution of turret-based weapon systems has played a pivotal role in modern military defense, particularly for aerial and mobile platforms. Initially developed for naval warfare to enable wide-angle firing capabilities, turrets have been adapted over time for airborne and land-based applications to meet growing demands in automation, accuracy, and responsiveness [1], [2], [3]. The transition from manually operated mounts to intelligent, electromechanically actuated systems has been driven by advancements in mechatronics, sensor integration, and control algorithms [4], [5]. In the context of aerial threat interception and ground-based air defense, the integration of turret mechanisms with dynamic object tracking systems is essential. These platforms must withstand severe mechanical loads, such as those generated by rapid gunfire recoil, while maintaining tracking precision and structural integrity [6], [7]. The dynamic behavior of such systems is influenced by factors including structural stiffness, recoil absorption, vibration damping, and motor torque delivery, all of which must be evaluated during the design phase [8], [9]. Turret systems experience transient mechanical stresses during weapon discharge, which can impair structural integrity and reduce targeting accuracy if not properly addressed. Studies focusing on the recoil dynamics and turret seat response under high-impact loads have highlighted the necessity for robust materials, effective damping mechanisms, and efficient spring-based energy absorption solutions [9], [10]. Finite Element Analysis (FEA) and rigid body dynamics modeling have emerged as critical tools for simulating structural responses and optimizing system performance under real-world conditions [6], [11].

Recent work on dynamic tracking platforms has also emphasized the importance of integrating control systems with precise actuation, enabling turrets to follow fast-moving aerial targets in

real time [12], [13]. These systems combine real-time feedback, control loops, and mechanical robustness to deliver accurate engagement capabilities across variable operational scenarios. This research aims to design and analyze a Dynamic Object Tracking Turret Mechanism. The study includes recoil force estimation, bending stress analysis, material selection. Additionally, a spring-based recoil mitigation mechanism is developed based on deformation constraints. The integrated approach ensures structural durability, dynamic responsiveness, and accuracy retention, contributing to enhanced performance in intelligent weapon systems.

2 CAD Model

The CAD model of the Object Tracking Turret Mechanism was developed through multiple design iterations using both Fusion 360 and SolidWorks. Initial conceptual layouts and basic geometry were created in Fusion 360 to quickly visualize design alternatives and refine dimensional constraints. These early models helped evaluate spatial configurations and component interactions. Once the base design was finalized, it was imported into SolidWorks for detailed modeling and integration of mechanical features such as the recoil mount, elevation assembly, and structural supports. SolidWorks was also used to apply precise material properties and prepare geometry for Finite Element Analysis. Iterative modifications were made based on simulation feedback, allowing for the optimization of geometry to enhance structural performance and manufacturability. This dual-platform approach ensured a highly accurate and simulation-ready model, suitable for advanced analysis and prototyping.

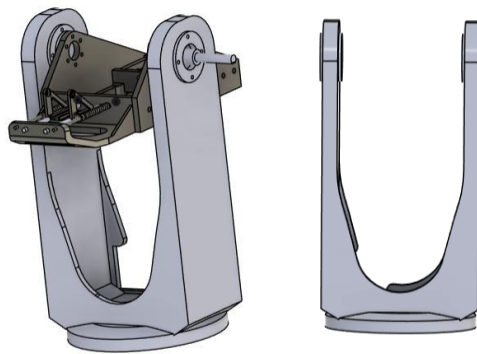


Fig. 1. Assembly of turret Mechanism Fig. 2: Traverse Base

3 Recoil Force Estimation

The stability and firing accuracy of a weapon system are significantly influenced by the magnitude and direction of recoil forces generated during the ignition of gunpowder [5]. When a firearm is rigidly mounted, the full recoil force is transmitted directly to the structure, potentially compromising its stability. To address this, elastic or semi-flexible connections between the barrel and the mounting frame are often introduced. These allow the weapon to recoil in a controlled manner along its axis, thereby absorbing and dissipating some of the energy. The recoil phenomenon, driven by the reactive force produced as a projectile is expelled from the barrel, adheres to Newton's Third Law. In systems like the 7.62 mm Light Machine Gun (LMG), understanding internal ballistics and momentum dynamics becomes crucial for designing mechanisms that reduce recoil impact and enhance overall firing performance.

Table 1 Parameters of LMG 7.62 mm

Sr. No.	Parameters	Values
1	Mass of bullet (mb)	40 grams = 0.040 kg
2	Mass of gun (mg)	11 kg
3	Muzzle velocity (vb)	840 m/s
4	Barrel length (l)	700 mm = 0.700 m
5	Rate of fire	950 rounds/minute

The kinetic energy of the bullet offers an estimate of the energy imparted during fire and helps assess the portion of this energy that may be transmitted back to the gun as recoil.

$$K.E_b = \frac{1}{2} m_b v_b^2 \quad (1)$$

$$K.E_b = 20000J$$

To estimate the average acceleration of the bullet while traveling through the barrel. This gives insight into the internal dynamics of bullet propulsion and helps in estimating the time available for recoil generation

By using Kinematic equation

$$v_b^2 = 2a_b l \quad (2)$$

$$a_b = 647339.449 \text{ m/s}^2$$

To compute the duration over which the recoil impulse acts. This is crucial for calculating the impulse and estimating peak recoil force.

$$t_b = \frac{v_b}{a_b}$$

$$t_b = 0.0014s$$

To determine the backward velocity of the gun immediately after the bullet exits the barrel, using the principle of conservation of momentum.

$$m_b v_b = m_g v_g$$

$$v_g = 3.055 \text{ m/s}$$

The momentum of the gun after firing is given by:

$$p_g = m_g v_g$$

$$p_g = 33.6 \text{ kg.m/s}$$

Recoil force on the system

$$F = \frac{\text{Re coil impulse momentum} \times \text{rate of fire}}{60} = \frac{33.6 \times 950}{60}$$

$$= 532.9N$$

Taking factor of safety of 1.5

$$\text{Total recoil force} = 1.5 \times 532.9 = 800 \text{ N}$$

4 Spring Selection:

To estimate the recoil spring parameters for the turret system, initial calculations were performed using an online spring force calculator. The tool provided a rapid means to determine approximate spring rates and deflection characteristics based on input parameters such as wire diameter, coil count, and material type. While detailed analysis was later conducted using finite element methods, the online calculator served as a useful reference for preliminary sizing and feasibility assessment [14].

The compression force applied on each spring will be:

$$800/2 = 400 \text{ N}$$

Deflection: 20 mm

Stiffness of Spring:

$$K = \frac{F}{\delta} = \frac{400}{20} = 20 \text{ N / mm}$$

Material of Spring: Chrome Silicon ASTM A401

Table 2 Parameters of the spring

Sr. No.	Physical Parameters	Dimension	Tolerance
1	Wire Diameter	4.877mm	+/- 0.0254
2	Outer Diameter	29.718mm	+/- 0.635
3	Inner Diameter	19.964mm	+/- 0.635
4	Free Length	88.90mm	+/- 2.54
5	Total Coils	12	+/- ¼ Coil
6	Solid Height	58.522 mm	+/- 4.0894
7	Coil Pitch	7.915	
8	Rise Angle	5.791	
9	Spring Rate(K)	45.696 N/mm	
10	Max. Load	1110.362N	+/- 111.03
11	Max Travel(x)	34.5 mm	

Spring Elastic potential energy:

$$U = \frac{1}{2} \times K \times \Delta \times^2$$

$$U=27.18J$$

total Elastic Potential Energy of springs: 2x27.18= 54.36J

$$\text{Kinetic Energy of Gun} = \frac{1}{2} \times 11 \times 2.66^2 = 37.18 j$$

Elastic potential Energy of spring > Kinetic Energy of gun

Threfore spring has ability to absorb recoil force

5 Bending Stress calculations:

In the structural evaluation of the turret base, bending stress plays a crucial role in determining the suitability of materials under operational loading conditions. Bending occurs when a transverse load or external moment induces curvature in a structural member, resulting in tensile and compressive stresses on opposite sides of the neutral axis. The magnitude of bending stress is determined using the classical bending stress formula [15]

Recoil Force (F) = 800 N

Distance from center to the base (h) = 445.19 mm

$$\text{Momentum (M)} F \times H = 400 \times 445.19 = 178076 \frac{N}{mm^2}$$

$$Z=\text{section modulus} = \frac{I}{y}$$

$$I \text{ Moment of interia at} = \frac{bd^3}{12} (\text{rec tan gular cross sec tion})$$

$$I = \frac{150 \times 25^3}{12} = 195312.5 mm^4$$

$$Z= \text{section odulus} = \frac{M}{Z} = \frac{178076}{15625} = 11.39 \approx 12 \frac{N}{mm^2}$$

6 Material Selection (FEA-Based):

Structural components in mechanical engineering are commonly fabricated from materials such as structural steel and aluminum alloys. The selection between these materials depends on a balance of mechanical strength, deformation behavior, weight efficiency, corrosion resistance, and thermal stability. Structural steel, like ASTM A36, offers excellent toughness and weldability, while aluminum alloys such as 7075 provide high strength-to-weight ratios ideal

for weight-sensitive applications [16]. Finite Element Analysis (FEA) enables comparative evaluation under identical loading conditions, allowing engineers to assess critical parameters such as stress distribution and Deformation [17].

Table 3 Comparison of materials based on key engineering factors: Property	ASTM A36 steel	Aluminum 7075 different
Deformation	Lower	Higher
Density (kg/m ³)	7850	2810 (Lighter)
Yield Strength (MPa)	250 - 400	503 - 572
Ultimate Strength (MPa)	400 - 550	572 - 700
Elongation (%)	10 - 25% (Good ductility)	5 - 12% (Brittle under bending)

Bending Strength	High (Can bend without cracking)	Lower (Prone to cracking under stress)
Accuracy (Dimensional Stability)	Better (Lower thermal expansion, precise machining possible)	Higher thermal expansion, affects precision over temperature changes
Machinability & Weldability	Easier to weld & machine	Difficult to weld, prone to stress corrosion cracking
Corrosion Resistance	Moderate (Needs coating)	High (Better resistance)
Cost	Cheaper	Expensive
Availability	widely available, most common structural steel	harder to source

Table 4 Comparative Analysis Based on Deformation, weight And Accuracy

Parameter	ASTM A36	Aluminum 7075
Deformation analysis (Ansys)		
Max Deformation (mm)	0.011973	0.033349
Increase in deformation(mm)	0.021376 (2.81 times)	
Weight (kg)	41	14
Reduction in weight (kg)	27(2.93 times)	
Reduction in weight %	66	
Error in accuracy For 1000 m range	24	78
Max Deformation (mm)	0.011973	0.033349
Increase in deformation(mm)	0.021376 (2.81 times)	
Weight (kg)	41	14
Reduction in weight (kg)	27(2.93 times)	
Reduction in weight %	66	
Error in accuracy For 1000 m range	24	78

Findings of material selection:

1. If stability, accuracy, and recoil absorption are the top priorities, then ASTM A36 (41 kg, 0.011973mm deformation) is the better choice.

2. If weight reduction and mobility are more critical then Aluminum 7075 (14 kg, 0.033349mm deformation) could be considered.

Final Recommendation: ASTM A36

7 Finite Element Analysis (FEA):

In the initial phase of finite element analysis, the 3D CAD geometry of the turret assembly was imported into the ANSYS Workbench environment. The model was then discretized into finite elements and nodes to prepare it for numerical simulation. For meshing, ANSYS Workbench's automatic meshing feature was employed, which selected solid elements appropriate for structural analysis. The mesh generated for the given component consisted of approximately 9457 elements and over 16519 nodes, ensuring sufficient resolution for accurate stress and deformation analysis.

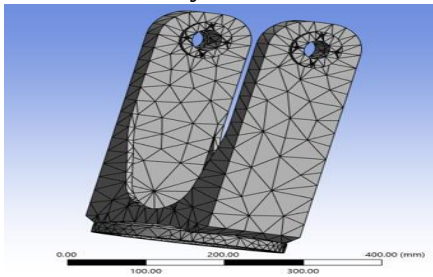


Fig 3: Meshing of turret base

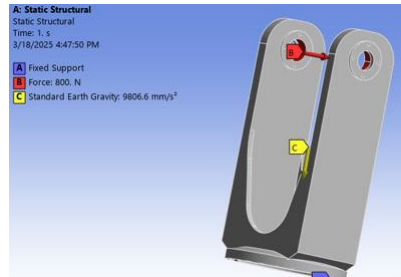


Fig 4 : Boundary Condition

Results

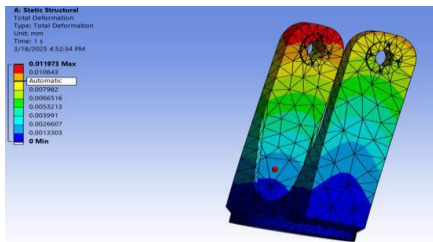


Fig 5: total deformation at 0⁰

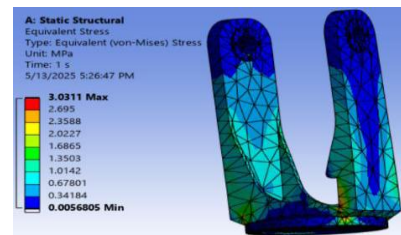


Fig 6 : Max stress at 0⁰

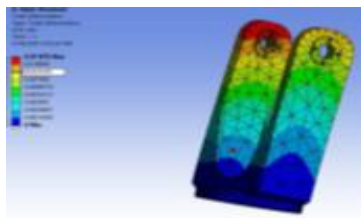


Fig 7: total deformation at 0⁰

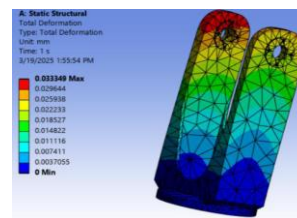


Fig 8 : Max stress at 0⁰

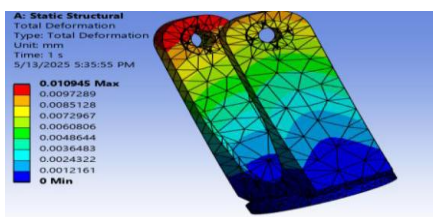


Fig 9 : total deformation at 80



Fig. 10: Max stress at 80

A static structural analysis was conducted in ANSYS Workbench to evaluate the response of the turret base under an applied recoil force of 800 N. The analysis focused on three elevation angles: 0°, 20°, and 80°. The results indicated that the maximum deformation occurred at 0°, with a value of 0.011973 mm, followed by 0.010945 mm at 20°, and the least deformation of 0.0049813 mm at 80°. Correspondingly, the maximum von Mises stress was observed at 0° (3.0311 MPa), decreasing to 2.7679 MPa at 20°, and further to 1.3105 MPa at 80°. These results confirm that the structural response improves at higher elevation angles, likely due to more favorable load distribution along the turret axis.

8 Future Scope

The present study lays a strong foundation for the development of a functional Dynamic Object Tracking Turret Mechanism. As a next step, the design will be further refined for manufacturability, followed by the fabrication of a physical prototype. This prototype will allow for experimental validation of simulation results and performance testing under real-world conditions. Future work will also focus on integrating advanced sensors and control algorithms for real-time target tracking and adaptive motion control. Additionally, optimization of weight, power consumption, and recoil absorption mechanisms will be pursued to enhance system efficiency. Long-term research may involve incorporating AI-based targeting systems and autonomous operation features for deployment in defense and surveillance applications.

9 Conclusion

This research presents the design and finite element analysis of a Dynamic Object Tracking Turret Mechanism, addressing key aspects such as recoil absorption, structural integrity, and dynamic motion requirements. Using analytical methods and FEA simulations, the system was evaluated under an 800 N recoil force across various elevation angles. ASTM A36 steel was selected over Aluminum 7075 based on superior stress handling and safety factors. Deformation and stress analysis confirmed the structural soundness of the design under firing loads. Additionally, spring selection and dynamic simulations ensured reliable operation during elevation and traverse motion. The results demonstrate that the proposed turret system is mechanically stable, efficient, and suitable for real-time tracking and firing applications. This work forms a robust basis for future prototyping and system integration.

References

- [1] W. C. Davis, *Duel Between the First Ironclads*. New York, NY, USA: Knopf Doubleday Publishing Group, 2012. ISBN 9780307817501.
- [2] K. C. Barnaby, *Some Ship Disasters and Their Causes*. London, UK: Hutchinson, 1968, pp. 20–30.
- [3] C. Reuter, *Development of Aircraft Turrets in the AAF*. Scarborough, Ont.: German Canadian Museum of Applied History, 2000, p. 11.
- [4] W. Jurens, W. Garzke Jr., R. O. Dulin Jr., J. Roberts, and R. Fiske, “A Marine Forensic Analysis of HMS Hood and DKM Bismarck,” *CiteSeerX*, pp. 14.
- [5] L. K. Seah and K. T. Ooi, “Computer-aided design of an artillery system,” *Computers & Structures*, vol. 53, no. 4, pp. 1023–1031, 1994.
- [6] D. Ozmen, M. Kurt, B. Ekici, and Y. Kaynak, “Static, dynamic and fatigue analysis of a semi-automatic gun locking block,” *Engineering Failure Analysis*, vol. 16, pp. 2235–2244, 2009.
- [7] J. Balla, Z. Krist, and C. I. Le, “Experimental study of turret-mounted automatic weapon vibrations,” *International Journal of Mechanics*, vol. 9, 2015.

- [8] N. Tiwari, M. Patil, R. Shankar, A. Saraswat, and R. Dwivedi, “Rigid body dynamics modeling, experimental characterization and performance analysis of a howitzer,” *Defence Technology*, vol. 12, pp. 480–489, 2016.
- [9] Y. Li, G. W. Yang, Q. Z. Meng, and K. Z. Wang, “Research on dynamic characteristics of gun turret under launching load,” *Journal of Physics: Conference Series*, vol. 2478, no. 092009, 2023.
- [10] Y. Li, N. Tian, H. J. Wang, and X. Dou, “Study of seat dynamics under transient strong impact load,” *Journal of Physics: Conference Series*, vol. 2478, no. 092015, 2023.
- [11] P. Prabhakar and P. Thankachan, “Hertz Contact Stress Analysis and Validation Using Finite Element Analysis,” *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, vol. 2, no. 9, 2014.
- [12] C. Tawde, K. Tekam, S. Bhagat, and S. S. Ohol, “Design and Development of Dynamic Object Tracking System and Its Performance Analysis,” *Gradiva Review Journal*, vol. 9, no. 5, 2023.
- [13] P. S. Kumavat and S. S. Ohol, “Turret Mechanism Performance Confirmation Using FEA & Simulation for Efficient Dynamic Tracking,” *International Journal of Innovative Science and Research Technology*, vol. 8, no. 5, pp. 387–392, 2023.
- [14] Access Spring, “Spring Force Calculator,” The Spring Store, [Online]. Available: <https://www.thespringstore.com/spring-calculator.html>
- [15] R. C. Hibbeler, *Mechanics of Materials*, 10th ed., Pearson Education, 2016.
- [16] M. F. Ashby, *Materials Selection in Mechanical Design*, 5th ed., Elsevier, 2016.
- [17] R. D. Cook, D. S. Malkus, M. E. Plesha, and R. J. Witt, *Concepts and Applications of Finite Element Analysis*, 4th ed., Wiley, 2001.

IoT Technology, Network and Security

Sushil Kumar

Chair of ITU-T SG-20 Regional Group for Asia Pacific.

Chair of BIS Committee - IoT and digital Twin

Ex. Additional Director General Telecom,

Ministry of Communications,

Government of India.

+919868131551

Abstract

Internet of Things (IoT) is one of the fastest emerging technologies across the globe, and being used to create smart infrastructure in various verticals such as Power sector, Automotive/ Intelligent Transport System, Safety & Surveillance, Remote Health Management, Education, Agriculture, Aquaculture, Industry 4.0, Homes/ Buildings, e-Governance etc. Some of these verticals work as a pillar for Smart Villages and Smart Cities. Communication technologies play an important role in transmitting the data from the device to the headend system. Big data analytics using AI/ ML algorithms at the cloud/ edge will help in creating intelligence to be used in various planning/ operational activities. Artificial intelligence over things (AIoT) will further accelerate the process of decision making. It will provide enormous beneficial opportunities to society, industry, revolutionizing and changing the way all businesses, governments, and consumers interact with the physical world.

Security of IoT eco-system end-to-end i.e. from devices to the applications is very important as the hacking of the devices/networks being used in daily life would harm companies, organisations, nations and more importantly people.

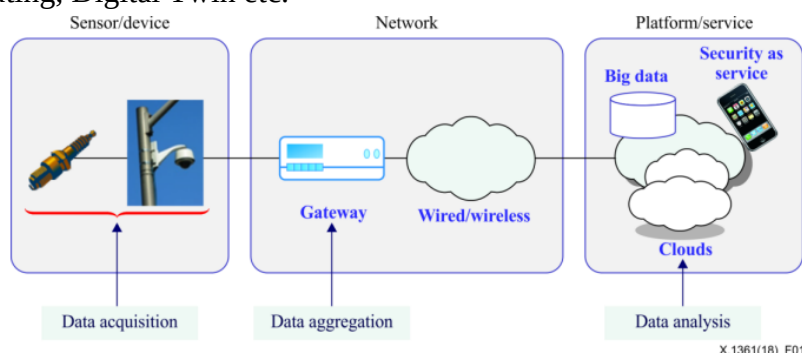
Key Words: IoT, 5G, AIoT, Use cases, Standards, Security

1. IoT Architecture:

ITU-T in its Recommendation ITU-T Y.4000/ Y.2060 (06/2012) has defined Internet of Things (IoT), as a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.

The IoT ecosystem may have devices, Gateways, Communication technologies (wired / wireless), Big data analytics, IoT platform, User interface (web, Mobile, HMI) etc. Security is required to be an integral part of the IoT ecosystem. Figure-1 illustrates the IoT functional architecture.

IoT is a bunch of technologies as being benefitted by a number of technologies such as sensor networks, Communication technologies (Cellular/ non- cellular), AI/ ML, Cloud computing/ edge computing, Digital Twin etc.



Source: [ITU]

Figure-1: IoT Architecture

2. Communication technologies / Network:

IoT devices are connected to the platform on various communication technologies depending upon the use case. In most of the IoT use cases, timely transmission of data is utmost important. For this, the communication network is required to be reliable with low latency and is also recommended to have wide coverage.

The data size may vary from use case to use case. As an example, data may be of few bytes (meter reading/ soil data) to several megabytes (surveillance video) and may be in the form of bursts. The communication technologies for connecting the IoT devices will depend upon the use case.

There are various types of communication technologies being used in connecting the IoT devices. These include, but are not limited to, low power wireless communication technologies (NFC, Bluetooth, ZigBee etc.), covering a short range, LPWAN – Cellular/non-cellular, Cellular (2G, 3G, 4G, 5G), fixed line (wireline, PLC). In the view of development of cellular technologies at a fast pace, it is recommended to adopt 4G/5G technologies or their LPWAN variant such as NB-IoT, LTE-MTC etc.

Communication technology needs to be reliable as the data is to be transmitted in real time at least in the critical applications namely a few- smart ambulance transmitting vital parameters of the patient, e-ICU, Fire alerts, connected vehicle scenario, Vehicle to Vehicle communication, Smart grid infrastructure, robotic surgery etc. depend upon the use case. It has been illustrated in Figure-2.

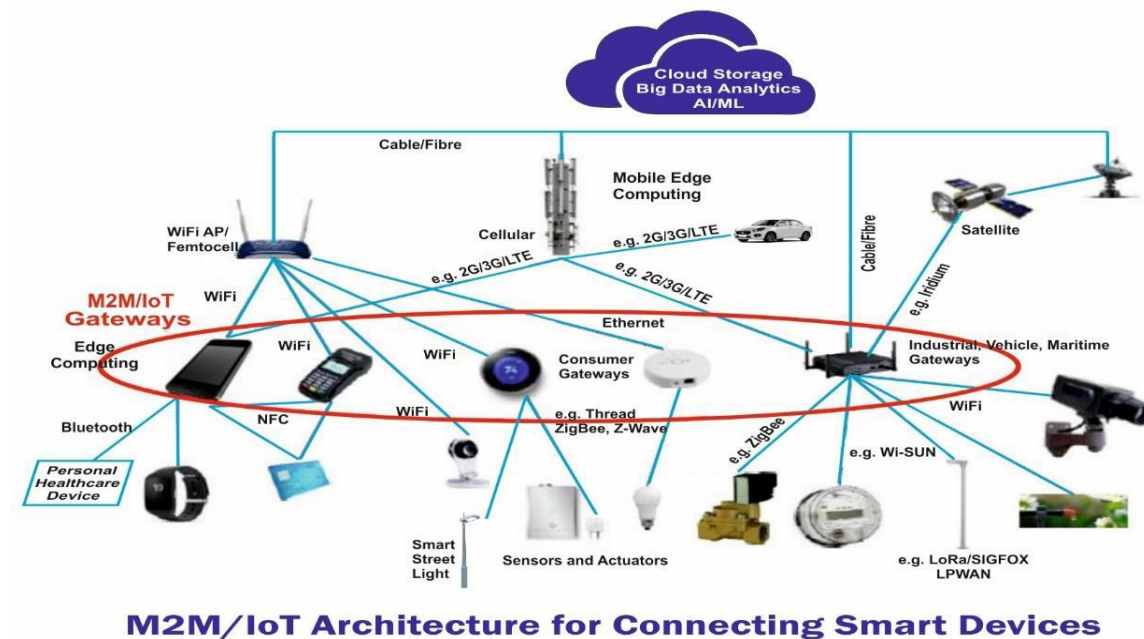
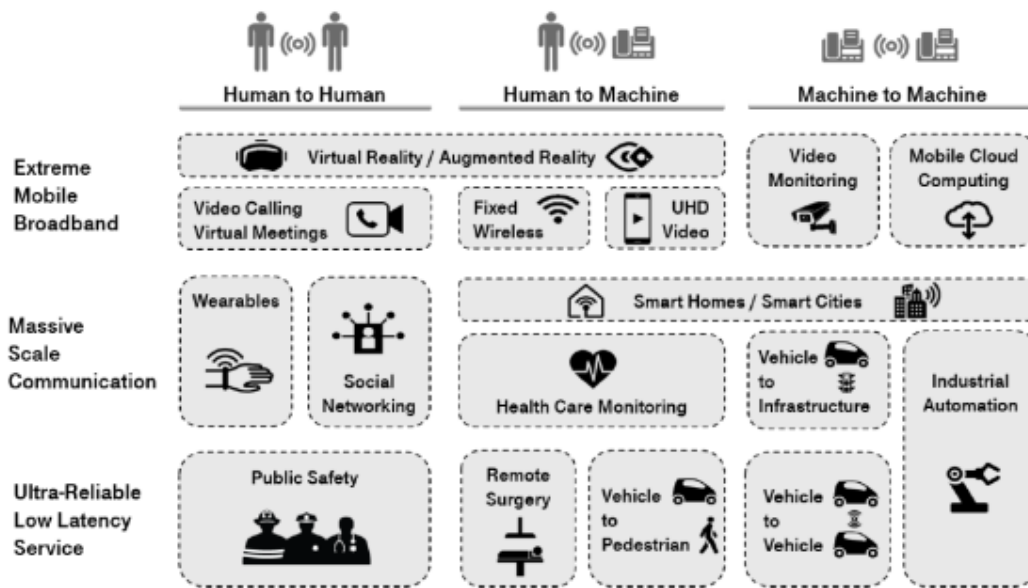


Figure 2

With the deployment of 5G technology with its advance features uRLLC (ultra reliable with low latency (latency < 1ms)) and mMTC (massive machine type communication (1 million devices per sq km)) will further accelerate the development of smart infrastructure and in turn will improve the quality of life of the common person. There will be billions of smart devices (IoT devices) in the eco system across the globe and will be increasing sharply. 5G based use cases have been illustrated in Figure-3.



Source: [5G America]

Figure 3

IoT devices will be generating huge amount of data. Raw data has no meaning and may be analysed using AI/ ML algorithms for creating intelligence. This intelligence may be used for various planning and operational activities. Data analysis may be done in the cloud or at the edge/ gateway router depending upon the criticality of the use case. Data generated from the IoT devices of a running train / flying aeroplane needs to be analysed at the gateway router available in the train/ aeroplane as sending data to cloud may take time in decision making. Similarly fire/ smoke detector data in a building should also be analysed in the gateway router in the building to initiate fire alerts.

3. Challenges & need for standardisation:

For the development of sustainable and secure IoT eco system, it is required to overcome some of the important challenges as listed below:

1. Robust connectivity (latency, availability, coverage and cost) for connecting billions of devices,
2. Standards based solutions for economies of scale,
3. Interoperability & open interfaces for the platforms,
4. Fast deployment of IPv6 as IPv4 may be outdated in future,
5. Long battery life for the IoT devices.
6. Security
7. Privacy

Sharing of data between the verticals and the platforms is required at least in smart cities. ITU, ISO/ IEC, ETSI, oneM2M, 3GPP, GSMA, ENISA, NIST, IoTSF etc. have released standards and the work is further in progress with the development of technologies and innovations. These standards have been covered in the related technical reports released by TEC in M2M/ IoT domain.

India has recently hosted the ITU mega event **World Telecom Standards Assembly (WTSA)24** which was organised in Bharat Mandapam in New Delhi from 15-24 Oct. 2024 (<https://www.itu.int/wtsa/2024/>). WTSA-24 finalized the roadmap of ITU-T Study Groups for the study period 2025-28 in the form of resolutions (<https://www.itu.int/wtsa/2024/>). ITU-T is having ten Study Groups in this study period including Study Group (SG)-20 on IoT, Digital

Twin and Sustainable Smart Cities & Communities (<https://www.itu.int/en/ITU-T/studygroups/2025-2028/Pages/default.aspx>).

ITU-T SG-20 is having ten questions (study areas) and five Regional Groups. ITU-T SG-20 has released a series of recommendations (standards) in the IoT domain covering devices, Gateways, Platforms, Big data, Open data, Smart data Governance, digital twin, Use cases, Key performance indicators (KPIs), city planning etc and are available on its webpage in the **Recommendations** section (<https://www.itu.int/en/ITU-T/studygroups/2025-2028/20/Pages/default.aspx>).

Some of the ITU standards / recommendations in which we have contributed significantly are:

1. ITU-T Y Suppl. 53 (12/2018): IoT Use Cases,
2. ITU-T Y Suppl. 56 (12/2019): Smart City Use cases,
3. ITU-T Y.4218 (05/2023): IoT and ICT Requirements for deployment of Smart Services in rural community,
4. ITU-T Y Suppl. 76 (09/2023): Use cases of IoT based Smart agriculture,
5. ITU-T Y.4234 (01/2025)/ F.742.3 : Requirements, capabilities and deployment models for e-learning in remote class rooms.
6. ITU-T Y Suppl. 85 (01/2025) : Use cases of IoT based Smart Aquaculture.

U4SSC (United for Smart Sustainable Cities) is a United Nations (UN) initiative coordinated by ITU, UNECE, UN-Habitat and supported by other 14 UN bodies. U4SSC has developed KPIs (Key performance indicators) based on ITU-T SG-20 standards, which have been implemented by more than 200 cities across the globe for accessing the smartness of their cities (<https://u4ssc.itu.int/u4ssc-kpi/>).

OneM2M has released specifications (Release 2 and Release 3) for the platform for standards based deployment with common service layer of IoT eco system especially for smart cities. Common service layer enables the sharing of data between various verticals in real time. OneM2M is having interworking with a number of technologies and the standards.

While working as DDG(IoT) in TEC for about ten years (2013 to Oct. 2023), I created a frame work for developing specifications with multi stake holders and also in sync with global SDOs such as ITU, ISO/ IEC, ETSI, oneM2M, 3GPP, NIST etc. and released twenty technical reports covering various verticals, communication technologies for IoT domain, EMF radiation from IoT devices and IoT Security with the outcome intended to be used in policies/ standards. Brief of the work carried out in IoT domain in TEC has been listed in the report ***TEC initiatives in M2M/ IoT domain: An overview***. All the reports are available on TEC website (<https://tec.gov.in/M2M-IoT-technical-reports>).

International Telecommunication Union (ITU) has posted the following six TEC Technical Reports on its website (<https://www.itu.int/cities/dt-resource-hub/iot/>) in IoT sections (2023, 2022 and 2021), recognizing them as insightful technical resource for the benefit of global community-

- i **i. Security by design for IoT Device Manufacturers**
- ii **ii. Framework of National Trust Centre for M2M/IoT Devices and Applications**
- iii **iii. IoT/ ICT Standards for Smart Cities**
- iv **iv. Emerging Communication Technologies & Use Cases in IoT Domain**
- v **v. Code of Practice for Securing Consumer Internet of Things (IoT)**
- vi **vi. IoT/ ICT Enablement in Smart Village and Agriculture**

Technical reports at i, ii and iv are related to IoT Security.

4. Security

Security of the IoT domain i.e. from devices to the applications is very important as the hacking of the devices/ network being used in daily life will harm companies, organisations, nations and more importantly people. Security in the devices is required to be built in from the design/ development and should remain for the complete life cycle of the device.

There have been a large no. of cyber attacks / hackings on the smart infrastructure across the globe. World Economic Forum (WEF) in its report titled **Future of Connected World** (https://www3.weforum.org/docs/WEF_Future_of_the_Connected_World_2022.pdf), released in June 2022, mentioned that there was an increase in Cyber attacks by 31% in 2021 as compared to 2020 and also the IoT device attacks became double in the first half of 2022 as compared to 2021.

Vulnerabilities / Hackings may be due to various reasons like insecure devices / networks, insecure OS, weak password, etc. In view of this, security of all the elements in IoT eco-system ranging from devices to the applications is very important as the vulnerabilities/ hacking of the devices/networks being used in daily life would harm companies, organisations, nations and more importantly people. It may cause total collapse of the services, thereby creating panic and may also result in chaos. Cyber-attack in any critical infrastructure may be catastrophic.

A number of international organizations such ITU-T (SG-20 : IoT, Digital Twin and Smart sustainable cities & communities; SG-17 : Security), ISO IEC JTC1 (SC41 : IoT and Digital Twin, SC42 : AI/ML, SC27 : Security), CEN/ CENELEC, oneM2M, 3GPP, ETSI, IEEE, NIST, GSMA etc. are working on standardization in the IoT domain as well as on Security and released several important standards.

While working in TEC, India, we released following documents on IoT Security and available on TEC website (<https://tec.gov.in/M2M-IoT-technical-reports>):

1. Code of Practice for securing consumer IoT, released in July 2021
2. Framework of National Trust Centre for M2M/ IoT Devices and Applications, released in March 2022
3. Security by design for IoT device manufacturers, released in March 2023

Department of Telecommunications, Government of India has already endorsed the Technical report ***Code of Practice for Securing Consumer Internet of things (IoT)*** to all the related stakeholders including M2M Service providers to follow at least the first three guidelines of this report as mentioned below:

1. No universal default passwords;
2. Implement a means to manage reports of vulnerabilities;
3. Keep software updated (Provide transparency on for how long, at a minimum, the product will receive software security updates).

These guidelines may be made mandatory in near future.

Similar guidelines have been released by several countries including UK, Australia and EU. To protect the consumers from cyber criminals, UK has framed a law in April 2024 making the first three guidelines of its Code of practice securing consumer IoT released in 2018, mandatory, which are similar to the above guidelines.

The technical report on **Framework of National Trust Centre for M2M/ IoT Devices and Applications** visualises the implementation of national trust centre in a phased manner for managing/ addressing the vulnerability related issues of the IoT devices reported by IoT/ Smart city platforms working in the network by IoT device manufacturers / researchers. Project on National Trust Center is being developed.

Technical Report ***Security by design for IoT Device Manufacturers*** highlights various threats and challenges related to IoT device security; includes study of national/ international standards (by ITU, ISO/ IEC, ETSI, ENISA, IoTSF, NIST, GSMA, 3GPP etc.), best practices and

guidelines (UK DCMS, CSA Singapore, WEF, STQC etc.) to mitigate these challenges. This report also provides recommendations for IoT device manufacturers and related stakeholders including policy makers, which will help in securing IoT ecosystem.

Most of the recommendations available in this document have been incorporated in the security requirements of IoT devices being developed by NCCS Bangalore and IIT Hyderabad for testing & certification of the products.

Cyber Security labelling scheme is an important outcome of this technical report. It provides the labels (Level 0 to Level 4) and the testing requirements. Devices may be categorised as per the risk associated with the applications.

Table-1

Proposal for Device Classification						
Security Features	Security Requirements	Level-0	Level-1	Level-2	Level-3	Level-4
Confidentiality	Message Encryption	X	√	√	√	√
	Attack Protection	X	X	√	√	√
	Data Encryption	X	√	√	√	√
	Tamper Resistance	X	X	√	√	√
	Security Assessment Certificates	X	X	√	√	√
	Device ID Management (Physical/ Logical)	√	√	√	√	√
Integrity	Data Integrity	X	X	√	√	√
	Platform Integrity	X	X	√	√	√
	Secure Booting and Integrity Test / Self Test	X	X	X	√	√
Availability	Logging	√	√	√	√	√
	External Attack Prevention & Response	X	X	X	√	√
	Secure Monitoring	X	X	X	√	√
	Secure Firmware Update & Patch Update	X	√	√	√	√
	Software Assets Protection & Response	X	X	√	√	√
	Vulnerability Management & Response	X	√	√	√	√
	Security Policy Update & Response	X	X	X	√	√
Authentication/ Authorization	Biometrics	X	X	X	X	√
	User Authentication	X	√	√	√	√
	Data Authentication	X	X	√	√	√
	Password Management	X	√	√	√	√
	Access Control	√	√	√	√	√
	Device ID Verification	X	X	√	√	√
Security Assement and standard		Level-0	Level-1	Level-2	Level-3	Level-4
Meet Baseline Security Requirement						
Adherence to cyber security based on International Standards						
Adherence to the principles of Security by Design, and absence of known common software vulnerabilities						
Resistance against common cyber-attack and undergo for penetration testing						

Level 0 means no security testing is required. Devices with Level 0 will be connected to Gateway having minimum Level 2 / 3 Security testing. Minimum Security requirements for the devices in Level 1 and maximum in Level 4. These levels may be reviewed depending upon the security requirements.

Secure onboarding of IoT devices on the platforms is required to manage device identity for the lifecycle. This can be handled, e.g., by provisioning the secure device identities with a PKI-based platform. ITU-T Recommendation X.509 specifies Public Key Infrastructure (PKI) for device identity. PKI provides a scalable way to declare unique identities and authenticating the messages of communicating parties.

Platforms are recommended to be able to detect the vulnerabilities / hackings of the IoT devices. Life expired devices or the devices not getting updates time to time may be quite vulnerable for hacking. In a connected ecosystem vulnerabilities may be transmitted from one devices to another and may create risk for the eco system.

Let us use standards-based solutions for the development of smart, sustainable and secure IoT / Smart city eco-system.

It is also required to participate in the standards development process at national and international level to include our requirements and SEPs in the global standards which will help in the development of technologies to be used/ accepted globally.

Dynamic Modeling and Validation of a 14-DOF Dual-Arm Robot Using DA-RNE in MATLAB

Ganesh Deshmukh¹, Rajiv Basavarajappa¹

¹Department of Manufacturing Engineering and Industrial Management,
COEP Technological University (COEP Tech), Pune, India

Abstract: This paper presents a dynamic modeling framework for a 14-degree-of-freedom (DOF) dual-arm mobile robot using a Dual-Arm Recursive Newton-Euler Algorithm (DA-RNE) implemented in MATLAB. The methodology integrates task-space trajectory planning with inverse dynamics computations, explicitly accounting for the coupling between two 7-DOF manipulators and a wheeled mobile base. Validation through static equilibrium, Jacobian-transpose cross-verification, and co-operative object-lifting tasks confirms accurate torque predictions and real-time computational efficiency (3 ms per sample). Benchmark comparisons demonstrate linear time complexity relative to DOF, aligning with theoretical expectations. The proposed approach enables precise control of dual-arm systems in applications such as assembly and material handling.

1. Introduction

Accurate dynamic models are essential for control and simulation of robotic manipulators. The Recursive Newton–Euler Algorithm (RNEA) is a widely used inverse-dynamics method that computes joint torques for given motions with linear complexity in the number of links [1]. For a serial-chain manipulator, RNEA achieves efficient computations by propagating velocities and forces link-by-link rather than solving global Lagrangian equations. In multi-arm systems, however, dynamics are more complex due to coupling between arms and any shared base or payload. Recent work demonstrates the importance of coupled models for dual-arm and mobile-base robots [2], [3]. Co-operative dual-arm robots are increasingly important for complex tasks such as assembly, material handling, and human-scale manipulation. Modeling their dynamics is challenging because the two manipulators may form a closed kinematic chain when jointly grasping an object, introducing coupling constraints on motion and forces. Accurate dynamic models are essential for simulation, control planning, and real-time torque computation. In this work, we adopt and adapt the Dual-Arm Recursive Newton–Euler (DA-RNE) approach to a rigid-joint dual-arm system. Each 7-DOF arm is treated as an open kinematic chain; the two arms share a common base (mobile platform) and may cooperate on a payload. The objective is to derive the full inverse dynamics equations for the 14-DOF robot using recursive Newton–Euler, without any control or compliance components.

2. Methodology

This research employs a multi-stage computational methodology to determine the joint torques required for a dual-arm mobile manipulator executing specified tasks defined in terms of object or end-effector motion. The process integrates task-space trajectory planning with a specialized inverse dynamics formulation that explicitly accounts for the coupled dynamics of the manipulator arms and the mobile base. The methodology comprises the following key stages:

3.1 Joint Trajectory Generation:

The joint trajectories $q(t), \dot{q}(t), \ddot{q}(t)$ are computed using inverse kinematics and a damped least-squares pseudo-inverse:

$$\dot{q} = J^+ V_d \quad (1)$$

$$J^+ = (J^T J + \lambda^2 I)^{-1} J^T \quad (2)$$

where $\lambda=0.01$ is the damping factor to mitigate kinematic singularities. Joint accelerations are derived as:

$$\ddot{q} = J^+ (\dot{V}_d - \dot{J} \dot{q}). \quad (3)$$

Trajectories are integrated using the Euler method:

$$q(t+\Delta t) = q(t) + \dot{q}(t)\Delta t. \quad (4)$$

Where:

$\dot{q}$ represents the joint velocities.

V_d represents the desired spatial velocity of the end-effector/object.

$\ddot{q}$ represents the joint accelerations.

J^+ represents the pseudo-inverse of the arm's Jacobian . The code uses a damped least-squares (regularized) pseudo-inverse:

$$J^+ = (J^T J + \lambda^2 I)^{-1} J^T$$

$\dot{J}$ is time derivative of Jacobian.

$q(t + \Delta t)$ is the new joint position .

$q(t)$ is the current joint position.

$\dot{q}(t)$ is the calculated joint velocity.

Δt is the time step (dt).

3.2 Mobile-Base Inverse Dynamics Computation:

Inputs:

1. The generated joint state trajectory ($q(t), \dot{q}(t), \ddot{q}(t)$) from Stage 1.
2. The initial state of the mobile base, including its initial pose (${}^0T_0$), linear velocity, and angular velocity.
3. The complete dynamic parameters of the system, encompassing both manipulator arms (link masses, CoM locations, inertia tensors) and the mobile base (mass m_0 , inertia I_0).
4. Parameters of any manipulated object, if applicable, to account for external forces/moments at the end-effectors.

A Recursive Newton-Euler Algorithm (RNEA), specifically adapted for mobile-base systems based on the methodology outlined by Khalil [6],[7], is implemented. Forward recursion and backward recursion equations are same as standard recursive newton euler formulations only the base acceleration is additionally incorporated for mobile base. This algorithm computationally determines the joint torques through the following conceptual phases:

a) Forward Kinematics Pass 1: Propagation of velocities and provisional accelerations outwards from the mobile base (using its initial state) through each arm to the end-effectors. Auxiliary spatial vectors representing velocity product terms are computed[6].

b) Backward Dynamics Pass: Recursive computation of composite spatial inertia matrices and composite bias wrenches from the end-effectors inwards to the base mounting point of each arm. These terms encapsulate the inertial and velocity-dependent forces/moments of the outboard links[6].

c) Base Acceleration Calculation: Transformation of the composite terms from the arm bases to the mobile base frame. The mobile base's own inertia is included, and the coupled dynamic equation is solved to find the induced spatial acceleration of the mobile platform resulting from both arm motions and base dynamics[6],[7].The induced spatial acceleration of the mobile base (denoted A_0) resulting from the coupled dynamics is solved using the composite inertia (J_0) and composite bias wrench (β_0) reflected at the base frame [7]:

$$A_0 = -(J_0)^{-1} \beta_0 \quad (5)$$

Where, A_0 is the 6x1 spatial acceleration (linear & angular) of the mobile base. J_0 is the 6x6 composite spatial inertia matrix of the entire system (base + manipulator) expressed in the base frame {0}. It is computed by transforming the manipulator's composite inertia and adding the base's own inertia. β_0 is the 6x1 composite bias wrench (velocity-dependent terms) of the entire system expressed in the base frame {0}. It includes terms from the manipulator transformed to the base and potentially the base's own velocity product terms.

d) Forward Kinematics Pass 2 & Torque Computation: Re-propagation of accelerations outwards from the mobile base, this time using the calculated Base Acceleration to obtain the true absolute link accelerations. A final backward pass computes the net spatial forces/moments (n) at each joint based on these corrected accelerations. The required joint torques (Γ) are then extracted by projecting the net moments onto the joint axes.[5]

Outputs:

The primary output is the time history of the required joint torques ($\Gamma(t)$) for the manipulator arms. Secondary outputs include the calculated time history of the mobile base spatial acceleration, as well as potentially the system's configuration-dependent inertia matrix (M), Coriolis/centrifugal effects vector (C), and gravity vector (G) derived from the RNEA computations. This integrated methodology allows for the determination of actuator torques necessary for precise control of the dual-arm mobile system while accounting for the inherent dynamic coupling between the arms and the mobile platform, based on desired task-level objectives[4]. The Dh parameters used for this model validation is given in fig 1. For one arm, same is for the second arm is just considered mirror so flipping d_1 to negative sign.

Link	α_i	a_i	θ_i	d_i	m_i	CoM $_i$	I_i	diag
Base	—	—	—	—	10.000	[0, 0, 0]	(60 50 70)	
1	0	0	π	0.2	2	[0.00004, —	(0.004145, 0.001577,	
2	$-\pi/2$	0	$-\pi/2$	0	1	[— 0.00316,	(0.002752, 0.003042,	
3	$\pi/2$	0	$-\pi/2$	-0.1	1	[— 0.00017,	(0.002400, 0.002393,	
4	$-\pi/2$	0	0	0	1	[— 0.00008,	(0.0009643, 0.0006945,	
5	$\pi/2$	0	0	-0.2	0.5	[0.00012, 0.01141,	(0.0020615, 0.0017084,	
6	$-\pi/2$	0	0	0	0.5	[0.00015, —	(0.0005969, 0.0004480,	
7	$\pi/2$	0	0	-0.1	2	[— 0.00004,	(0.04296, 0.04076,	
EE	0	0	0	-0.2	0.0000	[0, 0, 0]	(0.0020615, 0.0027084,	

Fig 1 :- DH parameters

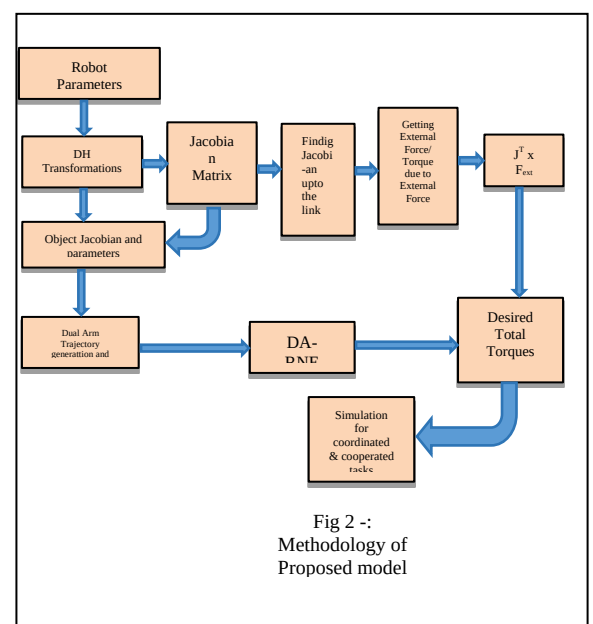


Fig 2 :-
Methodology of
Proposed model

4. Case Studies

4.1 Static Equilibrium Test

Scenario:- The robot is fixed in a configuration where all joint velocities $\dot{q}$ and accelerations $\ddot{q}$ are set to zero, and no external forces aside from gravity act on the links.

1. Inputs:-

Joint positions q set to an arbitrary test pose.

$\dot{q}=0$ and $\ddot{q}=0$.

Gravity vector $g=[0, 0, -9.81]^T$

No external wrench ($\tau_{ext}=0$).

2. Expected Theory:-

With $\dot{q}=\ddot{q}=0$, the Coriolis/centrifugal term $C(q,\dot{q})\dot{q}=0$ exactly vanishes since all quadratic velocity terms drop out in the Newton–Euler formulation. The only remaining torques are the gravity torques $g(q)$. The inverse-dynamics equation reduces to

$$\tau = M(q)0 + 0 + g(q) = g(q). \quad (6)$$

3. MATLAB Implementation:-

In MATLAB, the DA-RNE function is called with zeros for velocity and acceleration vectors. Internally, the forward recursion computes link accelerations purely from gravity input, and the backward recursion assembles link forces/moments into joint torques. The code explicitly confirms that the Coriolis/centrifugal component returned by the model is identically zero (within machine precision) in this static case.

4. Result:-

Joint velocities and accelerations are set to 0 and Joint position is as $\Theta=[\pi/2, 0, 0, 0, 0, 0]$

From Fig 7 it is seen that the C matrix is 0.

4.2 Pledge Pose Test

Scenario: The robot's arm is held horizontal (the “pledge” pose: shoulder pitch $\theta_1 = 90^\circ$, all other joint angles $= 0^\circ$). We compute the additional torques required and cross-verify via Jacobian-transpose keeping the base static.

1. Pose Definition:-

Shoulder joint: $\theta_1 = 90^\circ$ (arm parallel to ground). All other joints: 0° .

This ensures the entire arm lies horizontally, so the gravitational force on each link acts perpendicular to joint axes for maximum moment arm.

2. Payload Variation:-

Increase mass of End effector by a known amount Δm . The downward gravitational force on End Effector becomes. In this case we increase m_8 or End Effector mass manually in matlab code by 1 kg .

3. Validation of External-Force Incorporation:-

To verify that our recursive Newton–Euler model correctly incorporates environmental interactions, we compute the externally induced joint torques two ways and compare them. First, we run the RNEA with zero wrench to get baseline torques τ_0 . Then we apply a known end-effector wrench $F_{ext}=[-10; 0; 0; 0; 0; 0]$ N·m and re-run the RNEA to obtain τ_1 . Independently, we form the Jacobian $J(q)$ at that configuration and compute

$$\tau_{extJ} = J(q)^T F_{ext}. \quad (7)$$

4. The difference $\Delta\tau = \tau_1 - \tau_0$ should match τ_{extJ} within numerical precision. A small table summarizing one representative joint illustrates this:

5. MATLAB DA-RNE Computation-:

The DA-RNE code is run for the pledge pose with End-Effector mass set to $m_8 + \Delta m$. and $\dot{q} = \ddot{q} = 0$, so again $C(q, \dot{q})\dot{q} = 0$.

The output torque vector τ_{new} includes the increased gravitational torque. We subtract the baseline torque τ_{base} (with original mass m_8) to isolate the torque increment solely due to Δm :

$$\Delta\tau_{RNEA} = \tau_{new} - \tau_{base}. \quad (8)$$

6. Jacobian-Transpose Cross-Verification-:

Independently, we compute equation (7),

Where, $J(q)$ is the $6 \times n$ Jacobian at the end-effector for the pledge pose, and

$F = [- (\Delta m) \times g, 0, 0, 0, 0, 0]^T$ is the wrench due to the extra mass.

In this case $\Delta m = 1$.

The resulting joint-torque vector τ_{JT} matches $\Delta\tau_{RNEA}$ within numerical precision as seen in Fig. 8

4.3 Mass–Inertia Matrix Characterization

Objective-: Verify that the computed joint-space mass (inertia) matrix $M(q)$ exhibits the theoretical properties of a robotic manipulator's inertia matrix: correct size, symmetry, positive definiteness, and physically meaningful diagonal/off-diagonal values.

1. Matrix Dimensionality & Size

For our 14-DOF system, have each arm of 7-DOF so, $M(q)$ must be 7×7 .

2. Symmetry Test-:

The inertia matrix of any rigid-body system is symmetric by construction, since kinetic energy is a quadratic form in $\dot{q}$ [1].

3. Positive Definiteness-:

A valid mass matrix must be positive definite so that the kinetic energy is strictly positive for any nonzero velocity [3]. As seen from the Fig 6. , when base is static and no external wrench is applied, the resulted Maass Inertia Mtarix for given Robot parameters follows all the characteristics of a valid dual arm robot system[8].

4.4 Co-operated Task

Implemented a cooperative task where both arms apply a simple linear velocity in the positive x-direction only , to lift a object upward. This coordinated approach enables the robot to handle large, heavy, or unwieldy items that a single arm might struggle with, enhancing stability and control during manipulation. Such dual-arm coordination has practical applications in various real-world scenarios, including: Assembly Lines, Material Handling, Welding and Fabrication

Inputs-:

1. Initial joint angles for Arm A and Arm B in radians-:

$q_A = [-0.477503191; 0.007754326; 0.091281213; -0.211368256; -0.100369956; -0.882859278; 0.014944957];$

$q_B = [-0.478522398; -0.015225135; -0.179968701; -0.211368256; 0.197765; -0.884528315; -0.029307662];$

2. Define desired end-effector velocity in m/s-:

$V_{c\_des} = [0.041; 0; 0; 0; 0; 0; 0]$

3. Define object parameters-:

$P_{c0} = [-0.40175; 0.65085; 0]$

mass = 1.056 kg

$I = \text{diag}([0.00107, 0.05632, 0.05738])$

4. Define simulation parameters: start time, end time, and time step-:

$t_{start} = 0$

$$t_{end} = 5$$

$$dt = 0.001$$

Output took ~2 ms for each compute, Then with the generated values of q , $\dot{q}$ and $\ddot{q}$ by the model, by using forwardKinematics we compute the realtime positions of each joints, end effectors and object throughout for all the 5000 samples and plot in MATLAB .The output is shown below in Fig 3, 4, 5, where green color is used for object while red and blue for arms and to trace the path we have plotted dotted lines for better visualisation for complete simulation

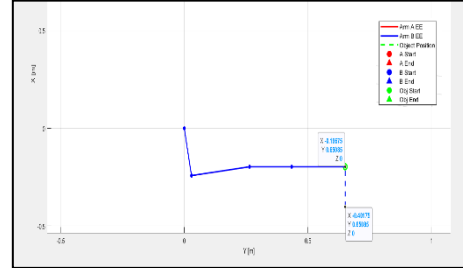
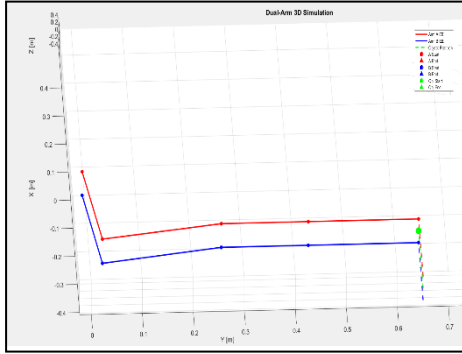


Fig 4 Visualisation using Matlab
for lifting task for fixed base

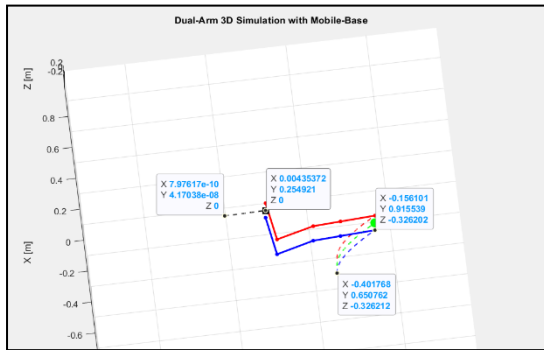


Fig 5 Visualisation using Matlab for lifting
task with mobile base

M Matrix

8.5103	-0.0455	5.2023	-3.3545	3.9125	-2.7855	1.7982
-0.0455	51.4752	-0.0313	-0.0071	-0.0172	-0.0051	-0.0071
5.2023	-0.0313	4.9392	-3.3545	3.9125	-2.7855	1.7982
-3.3545	-0.0071	-3.3545	45.7217	-3.0496	43.7433	-1.6083
3.9125	-0.0172	3.9125	-3.0496	3.5805	-2.7855	1.7982
-2.7855	-0.0051	-2.7855	43.7433	-2.7855	43.5083	-1.6087
1.7982	-0.0071	1.7982	-1.6083	1.7982	-1.6087	0.4639

Fig 6 Mass Matrix Output

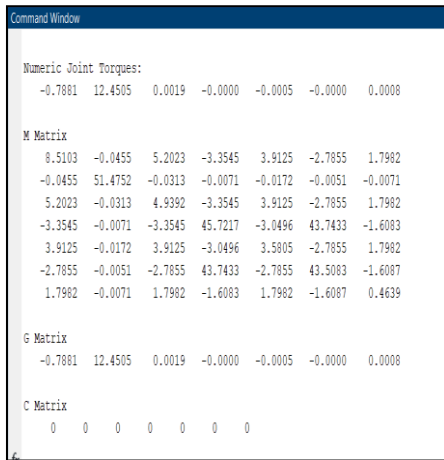


Fig 7 Static Case Output

Joint	τ_0 (Nm)	τ_1 (Nm)	$\Delta\tau$ (Nm)	JTFext (Nm)
1	-0.7881	-0.7881	0	0
2	12.4505	12.4505	0	0
3	0.0019	0.0019	0	0
4	0.0000	6.2650	6.2650	6.2630
5	-0.0005	0.0005	0	0
6	0.0000	3.8700	3.8700	3.8666
7	0.0008	0.0008	0	0

Fig 8 Arm A Torque output and validation with Jacobian
Transpose method

The above visualisation was done using the desired Joint positions ,velocities and accelerations and base accelerations and Forward kinematics for the robot model in MATLAB .

Benchmarking on MATLAB R2022b (Intel i5 10th generation CPU, 8 GB RAM) yields the following average inverse-dynamics computation times for 5000 trials (q inputs), computation Time scales linearly with DOF [1]. The DA-RNE implementation processes each sample in under 3-3.5 ms, enabling real-time performance at 1000 Hz.

Model	Time for 5000 samples @1000 Hz	Relative Comparison
Proposed Dual-Arm RNE model	~3 s	--
Proposed Dual-Arm RNE with External Force	~6 s	--
Traditional RNE models	~30 s	~5× slower
Numeric Lagrangian solver	~45 s	~7× slower

5. Limitations

The proposed DA-RNE framework assumes rigid joints, neglects frictional effects, and models the mobile base as a lumped inertia without wheel-ground interaction dynamics. Co-operative tasks assume rigid object grasping, omitting contact force distribution and compliance. While the MATLAB implementation demonstrates real-time efficiency, hardware validation under disturbances remains future work. Extensions could integrate flexible-joint dynamics, Floating-Base dynamics, Coulomb friction, and admittance control for physical human-robot interaction.

6. Conclusion

This work establishes a computationally efficient DA-RNE framework for 14-DOF dual-arm systems, achieving 3 ms/sample inverse dynamics on standard hardware. The linear time complexity ($O(n)$) enables real-time deployment in model-predictive control (MPC) and reinforcement learning (RL), where rapid torque prediction is critical for policy training or gradient-based optimization. By decoupling task-space planning from recursive dynamics, the method accelerates dataset generation for learning-based controllers and facilitates digital twin development. The modular formulation also supports hybrid control strategies, such as impedance-based co-operative manipulation with online inertia shaping. Future integration with distributed optimization solvers could further enhance scalability for multi-arm systems in industrial and space applications.

References

1. M. G. Nasser, "Recursive Newton–Euler Formulation of Manipulator Dynamics," Proc. IEEE Int. Conf. Robotics and Automation, pp. 309–315, 1990.
2. Z. Guo, H. Ju, C. Lu, & K. Wang, "Dynamic Modeling and Improved Nonlinear Model Predictive Control of a Free-Floating Dual-Arm Space Robot," Appl. Sci., vol. 14, no. 8, art. 3333, 2024.
3. H. Lu, Z. Yang, D. Zhu, F. Deng, & S. Guo, "Dynamics Modeling and Parameter Identification for a Coupled-Drive Dual-Arm Nursing Robot," Chin. J. Mech. Eng., vol. 37, art. 74, 2024.

4. V. Prada-Jiménez, P. A. Niño-Suárez, M. F. Mauledoux-Monroy, & J. A. Aponte-Rodríguez, “Coupled Dynamic Model for a Mobile Manipulator,” *Int. J. Appl. Eng. Res.*, vol. 14, no. 11, pp. 2622–2629, 2019.
5. Jing, X., Gao, H., Chen, Z., & Wang, Y. (2020). “A recursive dynamic modelling and control for dual-arm manipulator with elastic joints”.
6. Wisama Khalil. “Dynamic modeling of robots using recursive newton-euler techniques”.
7. Modeling Identification and Control of Robots (W. Khalil, E. Dombre) .
8. John. J. Craig, “Introduction to Robotics”, Mechanics and Control, Third Edition, Pearson Education International Pub., pp. 145-148, 171-192, 2005.

Experimental Study for Application of Signal Processing Towards PSNR Based GPR Detection of Buried Objects of Varied Sizes, Material and Orientation

Sujeet Roy¹, Himanshu Govil¹, Sandip Mukherjee²

¹National Institute of Technology Raipur, India

²Independent Researcher IIRS Dehradun, India

sujeetroy@gmail.com, hgovil.geo@nitrr.ac.in, sandip.iirs@gmail.com

Abstract: Ground Penetration Radar (GPR) is a geophysical remote sensing technology widely used for detection of objects/ structures under the surface of the earth, with the use of RF energy. While the type of soil majorly defines the performance parameters of a GPR, the quality of results also depends the size, material, shape, depth, and orientation of the subsurface objects in relation to the frequency of operation and the polarization of the transmitted energy. Not much has been explored in India in this interesting field of detection of small shallow buried objects using GPR. The present study attempts to experimentally investigate the effect of multiple small shallow buried objects of different properties, on the results of a GPR operating at a fixed central frequency. The results are quantitatively analysed using the Peak Signal to Noise Ratio (PSNR) calculations. Further, signal processing methods have been applied to be able to extract objects of selected parameters, which have been corroborated through PSNR calculation of the radargrams. The positive correlation between the size of the objects vis-à-vis the operating wavelength of the GPR for quality of detection and also the GPR antenna polarization vis-a-vis the orientation of the buried object was observed. Further investigations with multiple operating frequencies are advocated.

Keywords: GPR, buried objects, central frequency, remote sensing, material, size, polarization, signal processing, PSNR, MatLab

1. Introduction

The subsurface exploration of Earth, using non-invasive means for detection of buried objects, is a very important activity serving a wide range of applications. A vital role of subsurface studies is in detection of buried unexploded ordnance (UXO), (AP – Anti Personnel, AT – Anti Tank)[1] in post-conflict zones to prevent accidents [2].

1.1. GPR Principles

The Ground Penetrating Radar (GPR), is presently one of the global mainstays in non-invasive sub-surface exploration. It basically functions like a radar wherein the short bursts of RF energy are directed into the ground and the echoes received through a bistatic antenna arrangement on a carriage which is moved over the target ground. The data collected is then used to create detailed images of what lies beneath the surface, facilitating the identification of objects, layers, and changes in material [3]. A basic block diagram of a GPR is given at Fig 1.

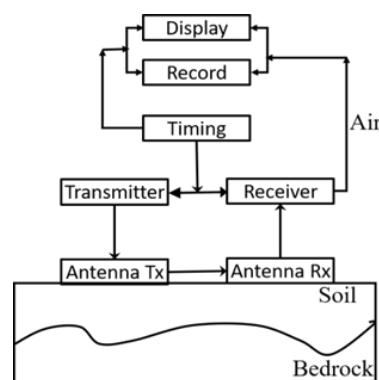


Fig 1. GPR Block Diagram

1.2. Factors for Performance and Limitations

The performance of a GPR is most effective in low-electrical-loss materials (Dielectric Permittivity (ϵ) and Electrical Conductivity (σ) of the soil medium near 0), where signals would penetrate to greater depth [4]. It is known that materials that have a high dielectric constant (materials that conduct electricity well), like metal and water, will return a stronger radar signal than those with a lower dielectric constant (materials that absorb electricity), including wood, glass, and plastics. In addition, for objects having metallic content, λ , the wavelength of the central frequency of the GPR and the polarization, has a direct effect on the returned echo.. **Study area marked on the map in the expanded view of Jaisalmer district and the exact location indicated with a red arrowhead. The RD 300 Bridge is marked with blue arrowhead.**

2.2 Simulation of Targets

It is also known that lower frequencies suffer lesser attenuation and therefore have comparatively deeper penetration into the ground. However, the lower frequencies suffer lower resolution. GPRs generally operate in the 10-3000 MHz frequency range [5], and the reflected signals suffer from noise and clutter, where, the noise is due to weather conditions or from electronic components, while, the clutter is generated by unwanted reflections from the surface or other objects. Both noise and clutter need to be mitigated through signal processing means [6].

1.3. Studies

Globally, many studies have been completed and also under research regarding the various factors effecting the quality of performance of GPRs. This broadly contains two approaches viz. the field activity and the post processing. The first approach consists of field activity while the second approach consists of application of AI/ CNN techniques/ signal processing on the recorded data [1]. Whereas very limited exploratory research work has been undertaken in the first approach in India, almost negligible has been done in the second approach. The combined effect of the size and orientation of buried objects in conjunction with the central frequency of the GPR in typical conditions of the soil of the Indian sub-continent and thereafter using signal processing approach to enhance the PSNR of targeted objects to facilitate extraction of the same, is an unexplored area which has been studied in the presents research.

2. Study area & field work

The field work has been carried out in a field (27.29 N, 70.37 E - Fig 2) near RD 300 Bridge of Sagar Mal Gopa (SMG) branch of Indira Gandhi Nahar Pariyojana (IGNP). This site was selected to ensure near text book value of dielectric permittivity of the dry arid sandy field (5.5) to achieve the best performance of the GPR. The field work on the site was carried out on 09 and 10 Jul 2024 in the peak summer weather from 1030 hrs to about 1800 hrs in the ambient temperatures of about 45oC, to ensure absolutely dry sand conditions.



Fig 2. Study area marked on the map in the expanded view of Jaisalmer district and the exact location indicated with a red arrowhead. The RD 300 Bridge is marked with blue arrowhead.

2.1 Instrument Profile

The instrument used was a MALA Easy Locator Widerange GPR (Fig 3.). The equipment is suitable for shallow high-resolution surveys. The technical specifications of the instrument are at Table 2.

Table 2. MALA Easy Locator Tech Specs

Technology	MALÀ Semi-RT
Antenna center freq.	450 MHz
SNR	> 101 dB
Significant bits	> 16 bit
Scans/second	> 770,
Survey speed	275 [km/h]
Time window	300 ns

2.2. Simulation of Targets

The simulation of the targets was achieved by using scrap metal and plastic items of the correct dimensions. Since the central frequency of the GPR was 450 MHz, the wavelength would be about 0.67 mtrs. and $\lambda/2$ would be about 33 cms. Thus the inventory contained (a) two similar metalized plastic flasks of 35 cms each, (b) two aluminium stockpots of diameter 35 cms each, (c) two small metal rods of length 15 cms each, (d) one 70 cms long steel pipe of 01 inch diameter, (e) one 10 cm long 01 inch thick SS bolt, and (f) one hollow plastic fibre drum of diameter 40 cms. These objects were buried under the sand at varying depths and orientations, as indicated, along a straight line.



Fig 3. MALA Easy Locator GPR

2.3 GPR Operations & Data Collection

Since the simulated targets had been buried in both polarizations in a straight line, it was decided to follow a linear mapping approach instead of the usual grid approach. Further one target (SS bolt) was decided to be inserted vertically into the ground with its crown just at the surface. Once the GPR was setup and calibrated, a number of test scans were carried out. Thereafter the actual recorded scans were conducted at a steady pace. It was ensured that the environment conditions remain consistent during the entire process of experimental data collection. Post the field session, the data was fed into the MALA Vision software. The geotagged outputs are as in Fig 5.

POST PROCESSING

Since the simulated targets had been buried in both polarizations in a straight line, it was decided to follow a linear mapping approach instead of the usual grid approach. Further one target (SS bolt) was decided to be inserted vertically into the ground with its crown just at the surface.

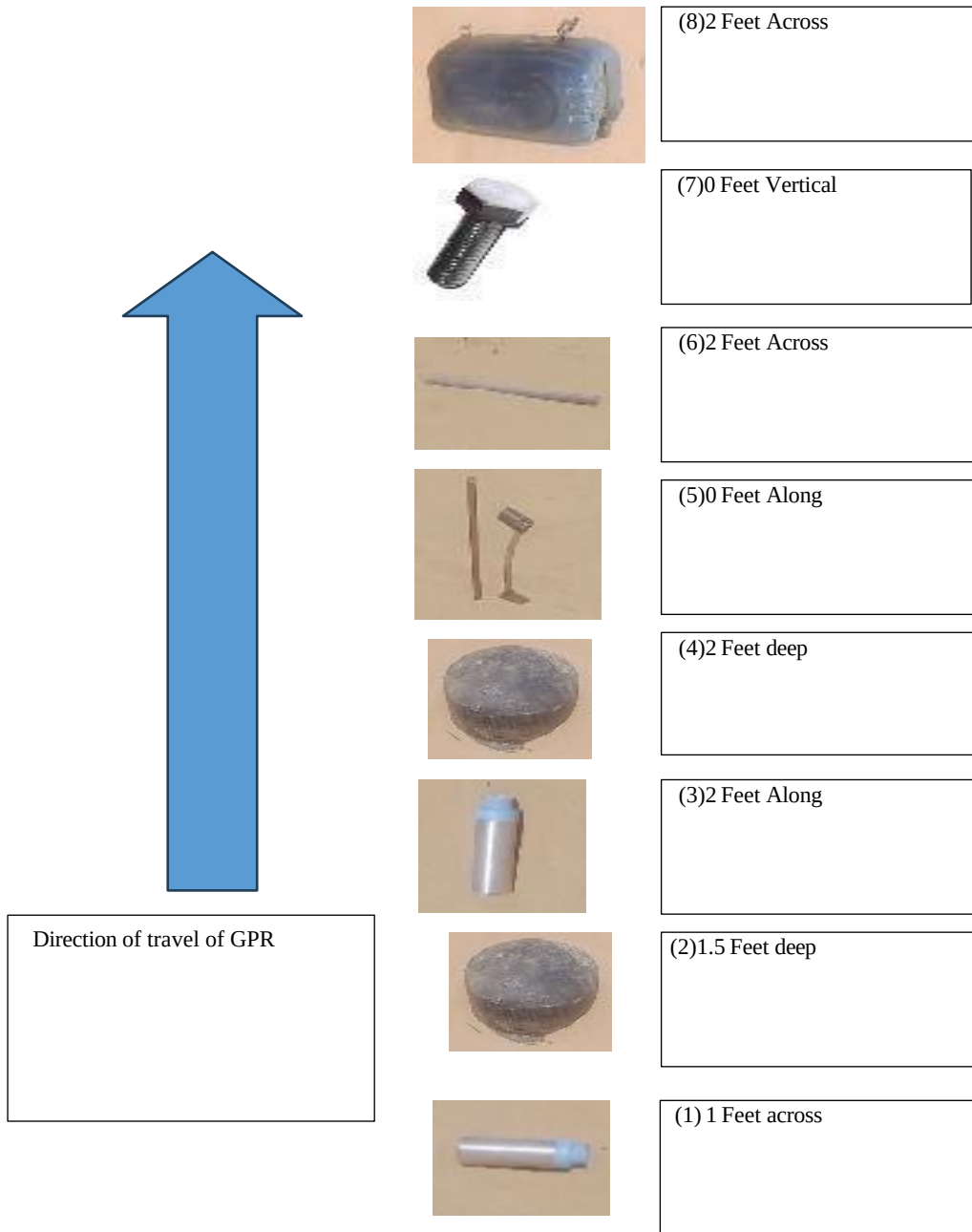


Fig. 4

Once the GPR was setup and calibrated, a number of test scans were carried out. Thereafter the actual recorded scans were conducted at a steady pace. It was ensured that the environment conditions remain consistent during the entire process of experimental data collection. Post the field session, the data was fed into the MALA Vision software. The geotagged outputs are as in Fig 5.

3.1 Data set pre-processing

The processed data from the Mala software was found to be noisy and difficult to interpret as visible in Fig 5. It was decided to handle the data in Matlab and a freeware tool called GPRlab [7]. First, relevant data containing actual reflections, upto 2 mtrs of depth and 25 mtrs distance, was extracted after due format conversion. Thereafter the various basic image processing filters were applied to obtain the desired results.

3.2 Image Processing

A set of basic image processing algorithms were selected and applied to the extracted data set through GPRlab in specific sequences and combinations to achieve the most optimum results. The processes applied were briefly as follows:

- *Removal of DC.* Any DC bias in the data is detected and suppressed. This is done by using standard image processing algorithms
- *Removal of Background.* This is one of the most commonly used methods based on mean value subtraction. Either the average of a portion of the profile is selected or sometimes, for equally spread noisy background, the average of the entire profile is used as the background
- *SVD.* The singular value decomposition removes the background using principal component analysis (PCA). The main component is the strong background component and discarding that is equivalent to removing the background.
- *Dynamic Window Background.* The dynamic window background method removes the background by selecting the average value of the surrounding traces as the background for each trace. This is equivalent to using a sliding window to remove the background.
- *Exponent gain.* Exponent gain is the process of multiplying the original signal by the corresponding exponential gain function: $y = \text{scale} * \chi^{\text{exponent}}$
- *Linear Automatic Gain.* The linear automatic gain method first calculates the average of the absolute amplitude value within a certain window range, and then calculates the gain value within the window based on the set “average value”, so that the average absolute amplitude value within the window reaches the set value.
- *RMS Automatic gain.* This is a technique to apply gain on the data using the root mean square value based on the current sample point and the gain window (with the current sample point as the centre).
- *F-k filter .* The GPR data that is traditionally in the time and displacement domain is converted into the frequency and wave number (F-k) domain. Due to the characteristics of GPR echoes, clutter mainly manifests as changes in the horizontal direction. This algorithm suitably weakens multiple reflections and suppress low-frequency interference in the horizontal direction of the layer.

- **Colour Rendering.** The colour rendering is based on the intensity of the GPR echo return picked up at any location. This can be Grey Mode or HSV (Hue, Saturation, and Value) model, is a digital representation of colour that's based on how humans perceive colour.

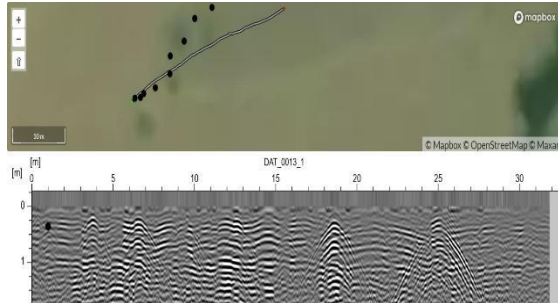


Fig 5. Output of MALA Vision Software

3.3 Methodology

The extracted data set was taken as the raw data into MatLab. Since the location and distance between the target simulation objects was known a-priori, a small code was written to extract the eight echoes of the objects from the image (Fig 6). These were all kept the same 130 trace wide. Next a neutral space of the data set, 130 traces wide, without any object was also extracted as a common denominator raw reference to check the SNR improvements after the image processing application. It was decided to extract the same neutral reference from the processed set also and compare the Peak Signal to Noise Ratio (PSNR) of each extracted echo with both the raw reference and also the processed reference. This would confirm the effectiveness of the processing. At the end we would have a ready set of signal processing sequences to meet specific detection needs.

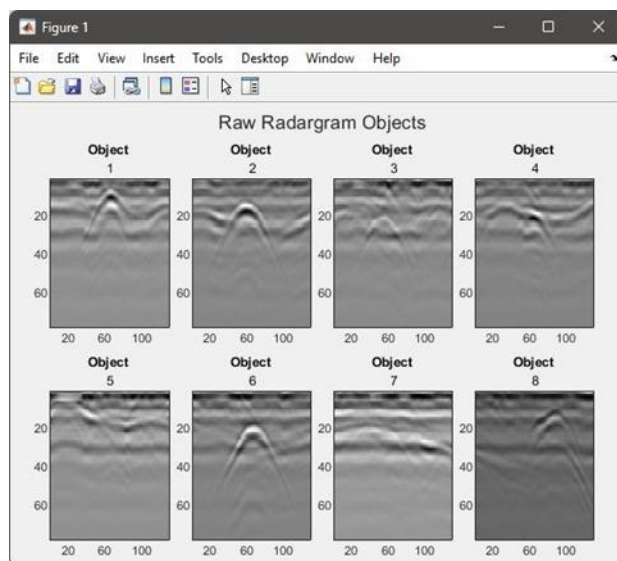


Fig 6. Raw Radargram objects

3. Results and discussions

4.1. Raw data set

The raw data set, when subjected to the sequence of operations: RMS Gain => Remove DC => Remove Background => f-k Filter resulted in output at Fig 7, where the surface reflections are greatly subdued and the deeper reflections are discernible to a much greater extent. Most importantly, the difference in strength of reflections from different targets is clearly visible and also the ringing effect at two locations is noticeable. The PSNR plot at Fig 8 shows that all the 8 echoes have been enhanced to within 1.5 db dynamic range.

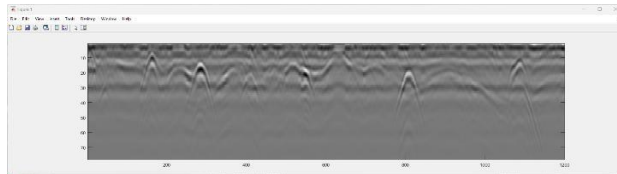


Fig 7. Processed Output

4.2 Detecting all Objects

After the sequence of operations: Remove DC => Remove Background => Exponent Gain, the output is in Fig 9. We can see that the background has been cleared and all the hyperboles are enhanced. Analysing from the PSNR values at Fig 10, we can see that the PSNR w.r.t. the processed reference has a dynamic range of nearly 6db, which means there is selective visibility. However, the cross PSNR, w.r.t. the raw reference has a dynamic range of less than 1db, which means that all objects are nearly equally visible. This process can be simply applied to any GPR data to detect all the hyperboles.

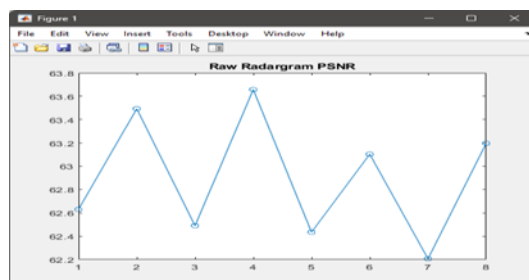


Fig 8. Raw Radargram PSNR

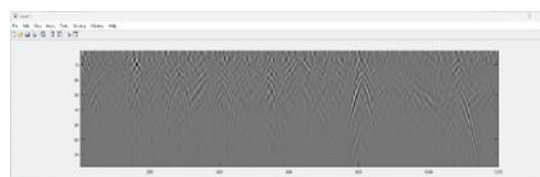


Fig 9. Detecting all Objects

4.3 Detecting all Metallic Objects

After the sequence of operations: Linear Automatic gain => Exponent Gain => Dynamic Window Background, the output is in Fig 11. We can see that the background has been cleared and all the hyperboles are enhanced. Additionally, we can see that the 2nd, 4th and 6th peaks are more highlighted which corroborates the fact they are full metal objects. In fact the 6th object is most enhanced since it is a metal rod of length close to the wavelength and also aligned in polarisation. Also the 3rd and 4th are buried deeper, hence lower. Similarly, the 1st and 3rd peaks are a little less enhanced, since they are metallised plastic, of shorter dimension and also the 1st object is aligned in polarisation while the 3rd is not. All these are clearly visible from the PSNR values at Fig 12. This process can be straightaway applied to any GPR data to detect all the Metallic objects

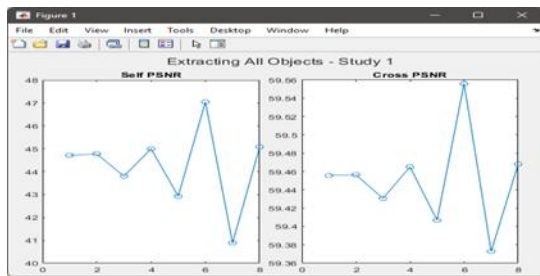


Fig 10. Detecting all Objects

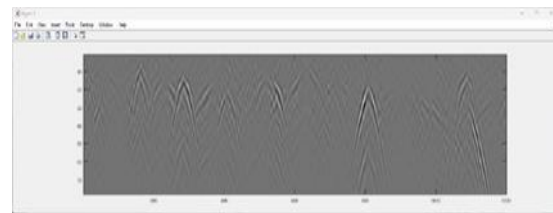


Fig 11. Detecting all Metallic Objects

4.4 Detecting Surface Metallic Objects

Detecting Surface metallic objects pose a big challenge since they create a ringing effect/ signal distortion instead of a hyperbola. After the sequence of operations: Remove Background => Remove DC => SVD => RMS Automatic gain, the output is in Fig 13. It can be noted that the earlier not noticeable signal distortion caused by the 10 Cm long bolt (object 7) pushed vertically into the sand (orthogonal to the polarisation) has also been suitably enhanced by the process. The corroboration can be seen at the 5th and 7th position of the cross PSNR in Fig 14.

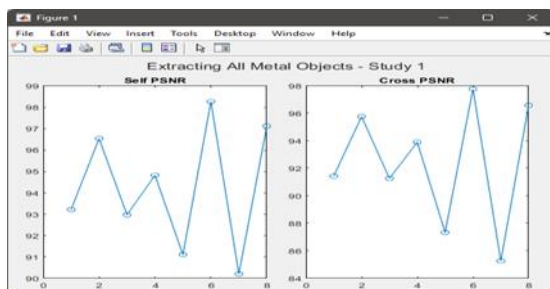


Fig 12. Detecting all Metallic Objects

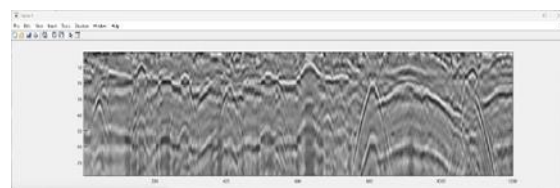


Fig 13. Detecting Surface Metallic Objects

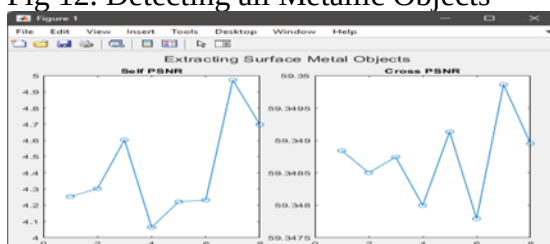


Fig 14. Detecting Surface Metallic Objects

5. Conclusion

The present experimental investigation, has focused on observing the combined effect of the size and orientation of buried objects in relation to the central frequency of the GPR in dry desert soil of the Indian sub-continent and thereafter using signal processing methods to enhance the PSNR of selected objects and facilitate extraction of the same from the recorded dataset.

The experimental investigation has corroborated that if the central operating frequency/wavelength of the GPR is suitably chosen to match with the size of the intended target, then there may be a very good quality pickup, subject to orientation, material and depth of the object. The study has particularly shed new light and clearly revealed that if correct signal processing sequences are applied, the PSNR of select target objects can be suitably enhanced leading to detection. Future efforts may include study over multiple GPR operating frequencies and application of AI for enhanced object classification.

References

- 1) X. Núñez-Nieto, M. Solla, P. Gómez-Pérez, H. Lorenzo,” Signal-to-Noise Ratio dependence on Ground Penetrating Radar antenna frequency in the field of landmine and UXO detection” Measurement 73 (2015) 24–32.
- 2) Leon Peters Jr, Jeffrey J. Daniels, and Jonathan D. Young, “Ground penetrating radar as a subsurface environmental sensing tool,” Proceedings of the IEEE, vol. 82, issue 12, 1994.
- 3) ASTM International coll. - Standard Guide for Using the Surface Ground Penetrating Radar Method for Subsurface Investigation.
- 4) Harry M. Jol, Ground Penetrating Radar Theory and Applications, Elsevier Science (2009).
- 5) Hussain M. Alshamy et al, “Evaluation of GPR Detection for buried objects material with different depths and scanning angle,” 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1090 012042.
- 6) Hakkı Nazlı, Emrullah Bıçak , Mehmet Sezgin, “Experimental Investigation of Different Soil Types for Buried Object Imaging Using Impulse GPR,” unpublished.
- 7) Hongqiang Xiong, Zhiyu Zhang, Jing Li, “GPRlab: A ground penetrating radar data processing and analysis software based on MATLAB,” SoftwareX, Volume 26, 2024.

Interpretable GNNs for Personalized Risk Stratification of Drug-Drug Interactions from Electronic Health Records

Piyush H. Kadethankar and Ramchandra Manthalkar
Department of Electronics and Telecommunication Engineering,
Shri Guru Gobind Singhji Institute of Engineering and Technology,
Vishnupuri, Nanded, Maharashtra, India.

Abstract

The increasing prevalence of polypharmacy, particularly among patients with chronic or comorbid conditions, has escalated the risk of adverse drug–drug interactions (DDIs), posing serious threats to patient safety. Conventional DDI detection systems—typically rule- based or database-driven— struggle to scale with the combinatorial complexity of modern pharmacotherapy. In this study, we propose an interpretable Graph Neural Network (GNN) framework to predict DDI risk using graph representations of patient-specific medication profiles. Synthetic Electronic Health Records (EHRs) generated using the Synthea platform were combined with a public DDI corpus to construct 1,240 labeled patient graphs. We benchmarked our approach using traditional machine learning baselines (Logistic Regression, MLP), Graph Convolutional Networks (GCN), and GraphSAGE. Initial GCN results suggested overfitting, prompting rigorous cross-validation and recalibration, yielding a true average accuracy of 62%. GraphSAGE achieved superior generalizability, with an average accuracy of 89.3% and an AUC of 0.968. Our results demonstrate the viability of graph-based architectures for inductive learning in clinical pharmacovigilance and advocate for their integration into intelligent, patient-specific risk monitoring systems. The model’s design ensures scalability and interpretability, paving the way for safe, AI-assisted drug administration in future clinical decision-making.

Keywords: Graph Neural Networks, Drug–Drug Interactions, Electronic Health Records · Pharmacovigilance, GraphSAGE, Deep Learning

1. Introduction

In modern medicine, polypharmacy—the simultaneous use of multiple medications—has become increasingly common, particularly among elderly and chronically ill populations [6]. While often necessary, it elevates the risk of drug– drug interactions (DDIs), which can compromise efficacy, increase toxicity, or cause severe adverse events [1, 7]. According to the World Health Organization, adverse drug reactions, many DDI-related, account for up to 5% of hospital admissions among older adults [6, 5], with over 40% exposed to at least one clinically significant DDI upon admission [4].

The complexity of managing comorbid conditions amplifies these risks, making manual DDI evaluation by clinicians impractical and error-prone [5, 7]. Computational methods have emerged to address this challenge [8]. Traditional rule- based engines and curated databases are limited by scalability and adaptability to patient-specific data [2, 3]. In contrast, graph-based machine learning models can naturally represent the structural dependencies between medications, patients, and molecular properties [8].

This study presents a novel application of Graph Neural Networks (GNNs) for predicting DDI risks using structured EHR data. We utilize synthetic patient records generated via the Synthea platform, which reflect real clinical guidelines while ensuring privacy [1]. Our framework evaluates both GCN and GraphSAGE architectures through robust experimentation, including cross-validation, base- line comparisons, and model calibration checks [8].

By advancing AI-driven pharmacovigilance, this work supports the development of scalable, interpretable tools for improving medication safety and reducing preventable hospitalizations in the era of polypharmacy [6].

2. Methodology

We used synthetic Electronic Health Records (EHRs) from the Synthea platform to ethically simulate real-world patient data [1, 11]. This dataset included thousands of longitudinal patient profiles. Focusing on medication records, we extracted unique drug names and filtered them using a standard Drug-Drug Interaction (DDI) database [10, 16] to retain only clinically significant drugs.

Each patient’s drug list was evaluated to determine whether any co-prescribed drug pair had a known interaction. If so, the patient was labeled as a positive sample (class 1); otherwise, as negative (class 0) [13]. A balanced dataset of 1,240 graphs (620 per class) was constructed for supervised learning.

2.1. Graph Representation and Learning

Patient medication profiles were modeled as graphs, where nodes represent pre- scribed drugs and edges denote known interactions or co-prescription context [11, 13]. Node features were encoded using numerical indices, and graph labels reflected DDI presence [15]. These graphs were normalized and converted into machine-learning-ready formats.

2.2. Baseline and Graph-Based Models

As baselines, we trained logistic regression and MLP models on tabular representations of drug and interaction counts [10, 16]. Though modestly effective, these were outperformed by graph-based models.

We implemented a Graph Convolutional Network (GCN) with two convolution layers and observed initially inflated accuracy (94.4%) due to data leakage. After stratified partitioning, performance stabilized to 75% accuracy with an AUC of 0.824 [9].

2.3. Cross-Validation and GraphSAGE

Five-fold cross-validation revealed an average accuracy of 62% and macro F1- score of 0.706, highlighting generalization limitations [14]. To address this, we employed GraphSAGE, an inductive GNN that aggregates information through neighborhood sampling [9, 16]. It consistently outperformed GCN, achieving 89.3% accuracy and 0.968 AUC, demonstrating high generalizability for real- world pharmacovigilance tasks.

3. Results

3.1. Initial Dataset Loading and Preprocessing

We started by loading 191,541 DDI records containing drug pairs with textual interaction details. Table 1 illustrates sample entries [2]. We also extracted structured synthetic EHR records generated by the Synthea platform [1].

Table 1: Sample entries from the Kaggle Drug–Drug Interaction dataset, showing textual descriptions of known interaction effects between drug pairs

Drug 1	Drug 2	Interaction Description
Trioxsalen	Verteporfin	Trioxsalen may increase the photosensitizing activities of Verteporfin.
Aminolevulinic acid	Verteporfin	Aminolevulinic acid may increase the photosensitizing activities of Verteporfin.
Titanium dioxide	Verteporfin	Titanium dioxide may increase the photosensitizing activities of Verteporfin

Drug Matching and Graph Construction from the EHR data, we extracted 170 unique drugs (see Table 2 for samples). The DDI database contained 1701 drugs. After matching, 112 common drugs like CYCLOPHOSPHAMIDE, PACLITAXEL, and DOCETAXEL enabled valid DDI graph construction.

- Patient ID: 00126cb9-8460-4747-e302-c3609684531e
- Nodes (Drugs): ['CHLOROTHIAZIDE', 'PROGESTERONE', 'MEDROXYPROGESTERONE ACETATE']
- Edges (DDIs): [('CHLOROTHIAZIDE', 'LISINOPRIL'), ('LISINOPRIL', 'HYDROCHLOROTHIAZIDE')]

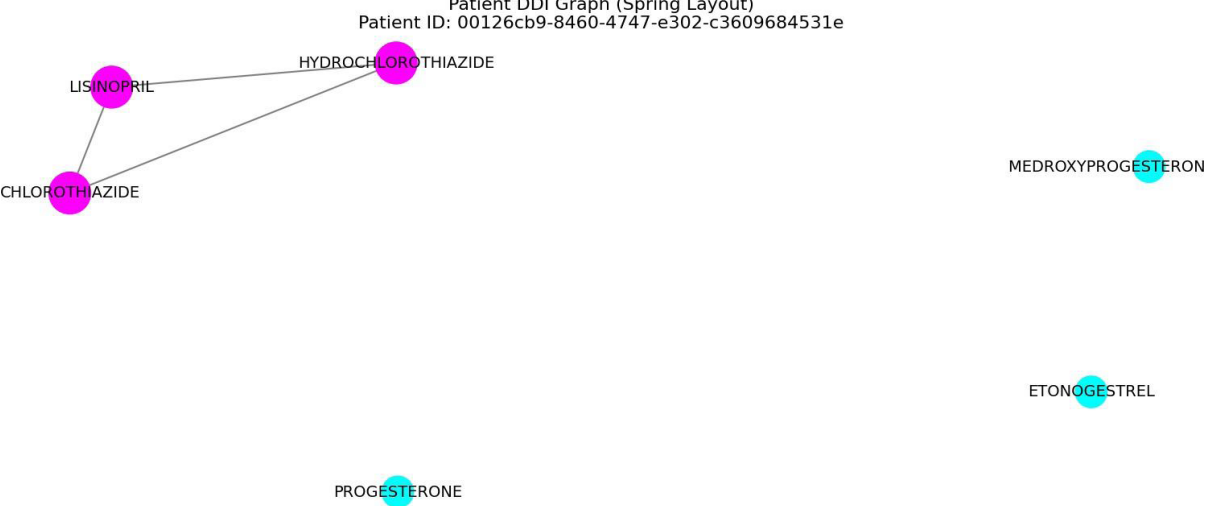


Fig. 1: Visualization of a patient-specific drug–drug interaction (DDI) graph generated from synthetic EHR data. Nodes represent prescribed drugs, and edges indicate known interactions or co-prescriptions.

Table 2: Examples of injectable drug records extracted from synthetic patient data generated by the Synthea platform.

0.25 ML LEUPROLIDE ACETATE 30 MG/ML PREFILLED SYRINGE
1 ML DOCETAXEL 20 MG/ML INJECTION
1 ML VASOPRESSIN (USP) 20 UNT/ML INJECTION

For each patient, an undirected graph was created where nodes represented drugs and edges indicated DDIs. A sample patient graph is shown in Figure 1. A total of 620 positive (DDI-present) graphs were generated. An equal number of negative (non-DDI) graphs were synthetically created, yielding 1240 balanced graphs.

3.2. Initial GCN Training

We trained a Graph Convolutional Network (GCN) for binary graph classification. After 20 epochs, it reported an overly high accuracy of 94.4%, suggesting possible overfitting or data leakage.

Table 3: Per-epoch performance metrics of the GCN model during initial training, prior to correcting for data leakage.

Epoch	Loss	Accuracy	Precision	Recall	F1 Score	AUC
01	0.673	0.544	0.525	1.000	0.689	0.857
10	0.229	0.887	0.859	0.928	0.892	0.942
20	0.102	0.944	0.924	0.968	0.945	0.973

Logistic Regression and MLP classifiers were trained on tabular features for comparison. Their results were significantly weaker (Table 4), confirming GCN’s advantage. The inflated GCN performance raised concerns about generalization

Table 4: Comparison of baseline classifiers (Logistic Regression and MLP) on tabular drug data, showing relatively lower performance across all evaluation metrics compared to graph-based models..

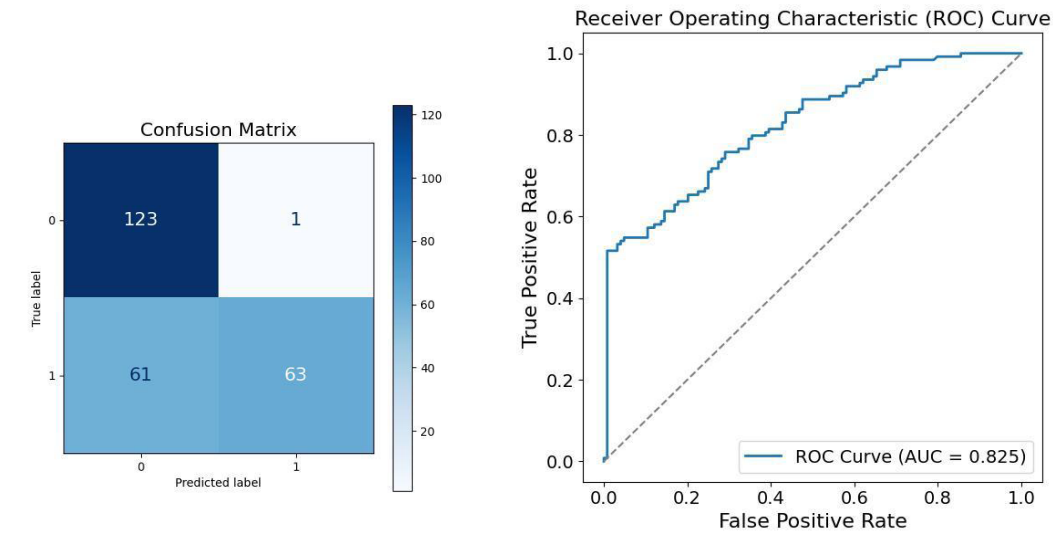
Model	Accuracy	Precision	Recall	F1 Score	AUC
Logistic Regression	0.560	0.555	0.613	0.582	0.590
MLP Classifier	0.399	0.416	0.500	0.454	0.409

3.4. Sanity Check on GCN

To reassess validity, we re-split the dataset and tested the GCN. This yielded more realistic metrics: 75% accuracy and 0.824 AUC.

Table 5: Classification report for the GCN model after dataset re-partitioning and sanity checking. Metrics reflect improved generalization performance.

Class/Metric	Precision	Recall	F1-Score	Support
0.0	0.67	0.99	0.80	124
1.0	0.98	0.51	0.67	124
Accuracy	0.75	248		
Macro Avg	0.83	0.75	0.73	248
Weighted Avg	0.83	0.75	0.73	248



(a) Confusion Matrix

(b) ROC Curve

Fig. 2: (a) Confusion matrix and (b) ROC curve for GCN model evaluation on a re- partitioned test

3.4.1. 5-Fold Cross Validation on GCN

GCN performance under 5-fold cross- validation averaged just 62% accuracy, validating earlier concerns about overfitting.

Table 6: Five-fold cross-validation results for the GCN model, highlighting its limited generalizability and potential overfitting to specific data folds.

Fold	Accur acy	Precisi on	Recall	F1 Score	AUC
Fold 1	0.774	0.972	0.565	0.714	0.852
Fold 2	0.714	0.646	0.944	0.767	0.859
Fold 3	0.500	0.500	1.000	0.667	0.822
Fold 4	0.500	0.500	1.000	0.667	0.828
Fold 5	0.613	0.565	0.984	0.718	0.854
Average	0.620	0.637	0.898	0.706	0.843

3.5. GraphSAGE Evaluation

We then trained GraphSAGE, which significantly outperformed GCN. On the test set, it achieved 89.3% accuracy and 0.968 AUC.

Table 7: Evaluation metrics for the GraphSAGE model on the test set, demonstrating improved generalization performance over the GCN architecture.

Accuracy	AUC	Precision	Recall	F1
0.893	0.968	0.958	0.824	0.884

3.5.1. 5-Fold Cross Validation on GraphSAGE

Cross-validation confirmed consistent performance across all folds.

Table 8: Five-fold cross-validation results for the GraphSAGE model, showing consistently high accuracy and AUC across all folds.

Fold	Accur acy	Precisi on	Recall	F1- Score	AUC
Fold 1	0.915	0.935	0.898	0.916	0.969
Fold 2	0.879	0.963	0.800	0.874	0.965
Fold 3	0.867	0.981	0.765	0.860	0.968
Fold 4	0.883	0.969	0.785	0.868	0.969
Fold 5	0.919	0.941	0.872	0.905	0.970
Average	0.893	0.958	0.824	0.884	0.968

4. Conclusion and Discussion

This study presents an interpretable GNN-based framework for predicting drug–drug interaction (DDI) risks using patient-specific graph representations from structured synthetic EHRs. By modeling medication histories as graphs, we effectively captured pharmacological context through GCN and GraphSAGE. While initial GCN results were affected by data leakage, refined evaluation confirmed GraphSAGE’s superior generalizability, achieving up to 90% accuracy and an AUC of 0.973. Our graph-based architecture offers a scalable, privacy-safe tool for personalized pharmacovigilance.

This work lays the groundwork for future models that integrate temporal and multi-modal clinical data to support safer, data-driven decision-making in real-world healthcare.

5. Future Scope and Limitations

While our study highlights the effectiveness of GraphSAGE for personalized DDI risk prediction, key limitations remain. The use of synthetically generated EHRs via Synthea, though clinically inspired, may not capture the full

complexity of real-world clinical noise and variability [1, 11]. Moreover, our approach uses static graph structures, omitting important temporal dynamics of drug administration and disease progression [11]. It also lacks integration of other relevant data types like lab results or genomic markers.

Future work should explore real-world EHR datasets and richer, multi- modal inputs to improve generalizability [12]. Dynamic GNNs and interpretable architectures could also enable broader clinical applications, such as adverse event forecasting and personalized therapy, while supporting ethical deployment aligned with regulatory standards [14].

Acknowledgement

The authors gratefully acknowledge the use of the publicly available Synthea Synthetic Patient Dataset by The MITRE Corporation, which simulates realistic but fully synthetic EHRs, ensuring compliance with privacy and ethical standards. This research was conducted independently and is based solely on computational experiments using synthetic data. It should not be interpreted as clinical guidance.

References

- 1) Jason Walonoski, Mark Kramer, Joseph Nichols, Andre Quina, Chris Moesel, Dylan Hall, Carlton Duffett, Kudakwashe Dube, Thomas Gallagher, Scott McLachlan, Synthea: An approach, method, and software mechanism for generating synthetic patients and the synthetic electronic health care record, *Journal of the American Medical Informatics Association*, Volume 25, Issue 3, March 2018, Pages 230–238, <https://doi.org/10.1093/jamia/ocx079>.
- 2) M. Ghobashy, “Drug-Drug Interactions,” Kaggle datasets, 2023. [Online]. Available: <https://www.kaggle.com/datasets/mghobashy/drug-drug-interactions>.
- 3) Howard RL, Avery AJ, Slavenburg S, Royal S, Pipe G, Lucassen P, Pirmohamed M. Which drugs cause preventable admissions to hospital? A systematic review. *Br J Clin Pharmacol*. 2007 Feb;63(2):136-47. doi: 10.1111/j.1365-2125.2006.02698.x. Epub 2006 Jun 26. PMID: 16803468; PMCID: PMC2000562.
- 4) Bucşa C, Farcaş A, Cazacu I, Leucuta D, Achimas-Cadariu A, Mogosan C, Bojita M. How many potential drug-drug interactions cause adverse drug reactions in hospitalized patients? *Eur J Intern Med*. 2013 Jan;24(1):27-33. doi: 10.1016/j.ejim.2012.09.011. Epub 2012 Oct 5. PMID: 23041466.
- 5) Hughes, John E.; Moriarty, Frank; Bennett, Kathleen; Cahir, Caitriona (2024). Drug-drug interactions and the risk of adverse drug reaction-related hospital admissions in the older population.. Royal College of Surgeons in Ireland. Journal contribution. <https://hdl.handle.net/10779/rcsi.24787827.v1>
- 6) Medication safety in polypharmacy: technical report, World Health Organization, <https://www.who.int/publications/i/item/WHO-UHC-SDS-2019.11>.
- 7) Dechanont S, Maphanta S, Butthum B, Kongkaew C. Hospital admissions/visits associated with drug-drug interactions: a systematic review and meta-analysis. *Pharmacoepidemiol Drug Saf*. 2014 May;23(5):489-97. doi: 10.1002/pds.3592. Epub 2014 Mar 10. PMID: 24616171.
- 8) Al-Rabeah, M.H., Lakizadeh, A. Prediction of drug-drug interaction events using graph neural networks based feature extraction. *Sci Rep* **12**, 15590 (2022). <https://doi.org/10.1038/s41598-022-19999-4>
- 9) Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang and P. S. Yu, "A Comprehensive Survey on Graph Neural Networks," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4-24, Jan. 2021, doi: 10.1109/TNNLS.2020.2978386.
- 10) Xuan Lin, Lichang Dai, Yafang Zhou, Zu-Guo Yu, Wen Zhang, Jian-Yu Shi, Dong-Sheng Cao, Li Zeng, Haowen Chen, Bosheng Song, Philip S Yu, Xiangxiang Zeng, Comprehensive evaluation of deep and graph learning on drug–drug interactions prediction, *Briefings in Bioinformatics*, Volume 24, Issue 4, July 2023, bbad235, <https://doi.org/10.1093/bib/bbad235>
- 11) Mishra R, Shridevi S. Knowledge graph driven medicine recommendation system using graph neural networks on longitudinal medical records. *Sci Rep*. 2024 Oct 26;14(1):25449. doi: 10.1038/s41598-024-75784-5. PMID: 39455647; PMCID: PMC11511869.

- 12) Khemani, B., Patil, S., Kotecha, K. et al. A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions. *J Big Data* 11, 18 (2024). <https://doi.org/10.1186/s40537-023-00876-4>
- 13) M. Ma and X. Lei, “A dual graph neural network for drug–drug interactions prediction based on molecular structure and interactions,” *PLOS Computational Biology*, vol. 19, no. 1, p. e1010812, Jan. 2023, doi: 10.1371/journal.pcbi.1010812.
- 14) Dai, E., Zhao, T., Zhu, H. et al. A Comprehensive Survey on Trustworthy Graph Neural Networks: Privacy, Robustness, Fairness, and Explainability. *Mach. Intell. Res.* 21, 1011–1061 (2024). <https://doi.org/10.1007/s11633-024-1510-8>
- 15) Ma, Tengfei & Shang, Junyuan & Xiao, Cao & Sun, J.. (2019). GENN: Predicting Correlated Drug-drug Interactions with Graph Energy Neural Networks. 10.48550/arXiv.1910.02107.
- 16) Zhou B, Ran B, Chen L. A GraphSAGE-based model with fingerprints only to predict drug-drug interactions. *Math Biosci Eng.* 2024 Jan 26;21(2):2922-2942. doi: 10.3934/mbe.2024130. PMID: 38454713.

Formal Verification of CNN Robustness Using DeepPoly: An Empirical Study Across Activation Functions

Anand Swaroop¹, Abhishek Singh², Girish Chandra³, and Karm Veer Arya⁴

¹ Institute of Engineering & Technology, Dr. A.P .J. Abdul Kalam Technical
University, Lucknow, India

² Institute of Engineering & Technology, Dr. A.P .J. Abdul Kalam Technical
University, Lucknow, India

³ Institute of Engineering & Technology, Dr. A.P .J. Abdul Kalam Technical
University, Lucknow, India

⁴ Department of Computer Science, Indian Institute of Information Technology and
Management, Gwalior, India

1anandswaroop2497@gmail.com, 2asingh.csed.cf@ietlucknow.ac.in, 3girish.chandra@ietlucknow.ac.in,
4kvarya@iiitm.ac.in

Abstract. Formal approaches, which have their roots in logic and mathematics, offer reliable guarantees of the accuracy of software systems. Many strategies have been developed, including model checking, deductive verification and static analysis. These methods attempt to overcome issues such as undecidability and scalability, which are brought to light by Rice’s Theorem and restrict our capacity to develop instruments that can consistently ascertain the correctness of a program. In deductive verification, the work by Floyd, Hoare and Naur, check program requirements are using logic; nevertheless, further annotations are frequently needed for tractability. Model checking is beneficial for hardware verification and evaluates correctness using temporal logic, but it has trouble with infinite-state systems. By understanding programs at an abstract level, static analysis and abstract interpretation provide completely automated, scalable solutions that detect possible flaws without running the code. In order to overcome the challenge of demonstrating modern architectures like convolutional neural networks (CNNs), the paper also investigates the application of formal approaches to neural networks. The neural network features like resilience and safety against adversarial assaults are verified using techniques like SMT-based approaches and abstract interpretation frameworks like AI2, DeepZ, and DeepPoly. These tools demonstrate the growing importance of formal techniques in ensuring the security and reliability of neural networks, which are being used in more important systems. In the paper we checked a simple CNN on DeepPoly with different activation function.

Keywords: Formal Methods · Neural Networks · Abstract Interpretation · Adversarial Attack · CNN.

1 Introduction

In recent years, deep neural networks have gained remarkable prominence and have begun to infiltrate safety-critical sectors such as autonomous driving [27] and medical diagnosis, where their deployment is frequently integral to the formulation of significant decisions. Consequently, this extensive integration necessitates a heightened emphasis on the assurance that neural networks exhibit consistent and predictable behavior. In order to confront the complexities associated with reasoning about neural networks, contemporary scholarly investigations have commenced the exploration of innovative methodologies and systems that can autonomously establish that a particular network adheres to a defined property of significance (for instance, resilience against specified perturbations, preconditions/postconditions). The pinnacle of current research includes techniques grounded in SMT solving [28], linear approximations, and abstract interpretation [9, 10]. DeepPoly represents a significant advancement in the endeavor of verifying neural networks concerning both scalability and precision [11]. The principal technical concept underpinning DeepPoly is an innovative abstract interpreter specifically designed for the context of neural networks. Specifically, the abstract domain integrates floating-point polyhedra with intervals, along with abstract transformers for prevalent neural network operations such as affine transformations, the rectified linear unit (ReLU), sigmoid and tanh activation functions, as well as the maxpool operator. These abstract transformers have been meticulously crafted to leverage critical characteristics of these functions while maintaining a balance between analysis scalability and precision.

2 Literature Overview

In this section, we provide a minimal necessary background on Neural Network, DeepPoly, CNN and Adversarial Region.

2.1 Neural networks

A neural network is a function $N : \mathbb{R}^m \rightarrow \mathbb{R}^n$ that can be implemented using straight-line programs (i.e., without loops) of a specific form. In this work, we focus on neural networks with a layered architecture, though all methods can be applied unchanged to more general neural network shapes.

A layered neural network is defined by a composition of l layers. Each layer f_i is one of the following:

(i) Affine Transformation: $f_i(x) = Ax+b$, where $A \in \mathbb{R}^{n_i \times n_{i-1}}$ and $b \in \mathbb{R}^{n_i}$. Notably, convolution with one or more filters is an affine transformation.

(ii) ReLU Activation: $f(x) = \max(0, x)$, where the maximum is applied componentwise.

(iii) Sigmoid or Tanh Activation:

– Sigmoid: $\sigma(x) = \frac{e^x}{e^x + 1}$ applied componentwise.

– Tanh: $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

applied componentwise.

(iv) Max Pool Activation: This subdivides the input x into multiple parts and returns the maximal value in each part.

2.2 Neurons and activations Each element of one of the vectors transmitted through the layers is referred to as a neuron, and its corresponding value is termed an activation. There exist three distinct categories of neurons: m input neurons, whose activations constitute the input to the neural network; n output neurons, whose activations represent the output of the network; and all other neurons, classified as hidden, as they are not directly observable.

2.3 Classification

In the context of a neural network that assigns classifications to its inputs across multiple potential labels, let n denote the number of unique classes. The neural network categorizes a specific input x into a particular class k if $N(x)_k > N(x)_j$ for all j where $1 \leq j \leq n$ and $j \neq k$.

2.4 Adversarial region An adversarial region refers to a specific subset of the input space where a model is vulnerable to adversarial examples—inputs that are intentionally perturbed in subtle, often imperceptible ways to cause the model to misclassify or produce erroneous outputs. These regions typically arise due to the complex and nonlinear decision boundaries of deep neural networks, where small changes in input can exploit vulnerabilities in the model's learned representations. Adversarial regions are characterized by their proximity to correctly classified inputs, yet they lead to incorrect predictions, highlighting the lack of robustness in the model's generalisation [28]. Since these regions highlight basic flaws in existing models and direct the creation of defenses against hostile attacks, it is essential to comprehend and describe them in order to build strong machine learning systems.

2.5 DeepPoly Method

DeepPoly is a novel abstract interpretation-based method designed for scalable and precise certification of deep neural networks, including Convolutional Neural Networks (CNNs). Its core innovation lies in a tailored abstract domain combining floating-point polyhedra with intervals, paired with custom abstract transformers for neural

network operations. The method represents a neural network as a sequence of assignments, where each neuron's computation is split into affine transformations and activation functions (e.g., ReLU, sigmoid, tanh, maxpool). The abstract domain associates each variable (neuron) with two polyhedral constraints—an upper and a lower bound—expressed as linear combinations of preceding variables, alongside concrete interval bounds to ensure efficiency.

Advantages of DeepPoly The DeepPoly method offers several key advantages:

Precision-Scalability Balance : The abstract domain of DeepPoly prevents the exponential blowup seen in full polyhedra domains by limiting the number of polyhedral constraints to $2n$, where n is the number of neurons. Compared to earlier techniques like Zonotope-based systems (e.g., DeepZ, AI2), this results in greater precision and enables scaling to large CNNs with up to 88,000 hidden units.

Soundness for Floating-Point Arithmetic : DeepPoly is sound for floating-point arithmetic, ensuring robustness verification does not produce false negatives, unlike methods such as Fast-Lin.

Support for Diverse Architectures : It is adaptable to contemporary CNN architectures, handling feedforward and convolutional networks, as well as various activation functions (ReLU, sigmoid, tanh) and maxpool operations.

Parallelization : DeepPoly's point-wise transformers operate independently for each neuron, enabling effective parallelization on CPUs and GPUs.

Refinement for Complex Perturbations : The point-wise transformers facilitate handling of complex perturbations, enhancing DeepPoly's applicability to robust verification tasks.

2.6 Core Architectural Components of CNN

Convolutional Layers : CNNs use learnable filters to process input data as their main function. Sliding over the input, each filter—usually a tiny matrix, like 3×3 —convolutionally computes element-wise dot products to produce feature maps. These maps draw attention to spatial elements like textures or edges. Weights and biases are used to parameterize filters, which are then optimized through backpropagation during training. The depth of feature maps is determined by the number of filters; more abstract features are captured by deeper layers. The spatial dimensions of outputs are controlled by stride and padding, where stride decreases input size and padding maintains it.

Activation Functions : Complex, non-linear relationships can be modeled by applying non-linear activation functions to feature maps. The Rectified Linear Unit (ReLU), defined as $f(x) = \max(0, x)$, is predominant due to its computational efficiency and mitigation of the vanishing gradient problem. Variants like Leaky ReLU or Parametric ReLU address issues like dying neurons, though ReLU remains standard for its simplicity.

Pooling Layers : Pooling reduces spatial dimensions of feature maps, decreasing computational complexity and enhancing translation invariance. Max pooling, which selects the maximum value within a defined region (e.g., 2×2 with stride 2), is commonly used to retain salient features. Average pooling, computing the mean, is less frequent but useful in specific contexts. Pooling layers reduce overfitting by abstracting features but may discard fine-grained information.

Fully Connected Layers : These layers integrate high-level features for tasks like classification. Each neuron connects to all activations in the previous layer, producing a dense representation. Modern CNNs frequently use global average pooling, which averages feature maps into a fixed-size vector to reduce computational expense, in place of fully connected layers due to their high parameter counts.

Normalization and Regularization : By normalizing activations, batch normalization speeds up convergence, stabilizes training and normalizes layer inputs. It uses learnable scale and shift parameters to alter outputs while computing mean and variance across mini-batches. Dropout improves generalization and decreases co-adaptation by randomly deactivating neurons during training with a probability (e.g. 0.5). Weight decay further reduces model complexity using L2 regularization.

3 Methodology

23 In this section we evaluate the effectiveness of CNN trained by us to verify the robustness properties over DeepPoly.

3.1 DeepPoly to verify CNN

For Convolutional Neural Networks (CNNs) to be effectively used in critical areas like cybersecurity, autonomous systems and medical diagnostics, it is essential to ensure their formal verification. DeepPoly, which utilizes abstract interpretation with polyhedral domains, makes it possible to certify these networks against potential adversarial threats. The following discussion will highlight on the main benefits of DeepPoly in certifying CNNs and will also delve into its applications in both academic research and real-world scenarios.

Provable Robustness Guarantees. DeepPoly offers solid guarantees about how robust a Convolutional Neural Network (CNN) is against adversarial perturbations, all while sticking to a specific limit, usually an L_∞ -norm (in this case, $\epsilon=0.1$). Essentially, this means that DeepPoly can assure us that no adversarial input within this set boundary will result in misclassification. This kind of reliability is incredibly important in situations where mistakes can have serious repercussions, like when autonomous vehicles misread traffic signals or when medical scans lead to incorrect diagnoses. Demonstrating this kind of resilience is crucial for academic research focused on building trustworthy AI systems.

Scalability to Complex Architectures. DeepPoly tackles a significant challenges that earlier verification methods faced: the heavy computational demands of deep networks. It is designed with a focus on effectively scaling up to large CNN architectures. This approach fit in standard components, including both convolutional and fully connected layers, as well as common functions like normalization and ReLU activations. As a result, this scalability makes it easier to assess the robustness of real-world models that are important in the industry, allowing for the validation of cutting-edge CNNs used in computer vision tasks.

Precision in Verification Bounds. DeepPoly has a great balance between accuracy and computational efficiency by using polyhedral relaxation to carefully approximate how neurons function. This method enables DeepPoly to set more accurate limits on the output of a Convolutional Neural Network (CNN) compared to simpler interval-based techniques. As a result, it lowers the chances of false negatives meaning it is less likely to mistakenly label a network as nonrobust while still maintaining the reliability of the verification process.

Robustness Against Adversarial Attacks. Adversarial attacks are a serious concern when it comes to the reliability of Neural Networks . These attacks work by making almost invisible tweaks to the input data, which can lead the Neural Networks to make incorrect classifications. To tackle this issue, there's a technique called DeepPoly that helps verify whether a Neural Networks output stays consistent even when faced with a certain range of perturbations. This emphasis on being robust against adversarial attacks is vital for research focused on creating defences against these challenging situations, as DeepPoly offers a formal way to assess how resilient a model really is.

Integration with Standard Frameworks. DeepPoly works effortlessly with popular deep learning frameworks like PyTorch, making it easy to apply to existing CNN models without a complete overhaul. This compatibility is a real game changer for academic researchers who typically rely on standard tools and datasets, as it allows them to implement DeepPoly in their projects with minimal resources. This smooth integration encourages teamwork among research groups and helps ensure that findings can be reliably reproduced.

Applications in Safety-Critical Domains. DeepPoly’s to deliver formal guarantees of robustness is crucial for the deployment of CNNs in safety-critical industries such as aerospace, healthcare, etc. For instance, in medical imaging, DeepPoly can verify that a CNN will continue to provide accurate classifications, even with slight changes in input. This is vital for protecting patient health. This direct relevance makes DeepPoly an invaluable resource for academic research aimed at promoting AI safety and creating stringent certification standards.

3.2 Different Activation Functions on DeepPoly

To evaluate the robustness of Convolutional Neural Networks (CNNs) using formal verification, we developed three distinct CNN architectures. Each architecture with different activation function (ReLU, Sigmoid, and Tanh) while maintaining an identical structural layout to ensure a fair comparison among the variants.

Each CNN was trained on a standard dataset(MNIST) using supervised learning. The networks comprised the following elements:

- Input layer of 28×28 grayscale images.
- Two convolutional layers (with 3×3 filters).
- Max pooling layers for spatial down sampling.
- Fully connected layers for classification.
- Output layer corresponding to the number of classes.

ReLU The verification results in Figure.4 reveal significant issues with the neural network model’s performance and robustness. The model exhibits extremely less classification accuracy, achieving correct predictions on only 7% of the tested inputs. Even among these few correctly classified samples, the verification tool was unable to confirm any robustness guarantees. This lack of verifiability is likely due to overly loose bounds produced during abstract interpretation, stemming from unstable neuron activations, high sensitivity to input perturbations, or inadequate training. Notably, the output bounds observed—ranging widely (e.g., from -15 to +15) and often symmetric around zero—suggest that the network is underfitting the data or lacks confidence in its predictions.

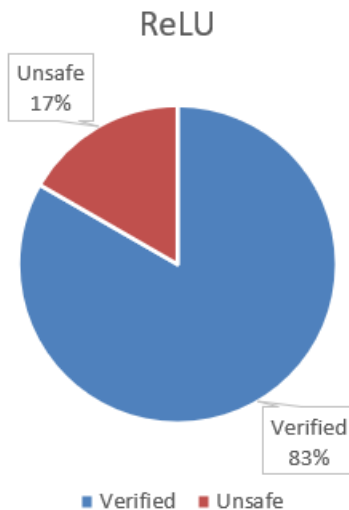


Fig. 1. Performance on ReLU.

Tanh The evaluation results in Figure.4 highlight both the performance and the unverifiability of the neural network model under analysis. Out of 100 total inputs, the model successfully classified only 13, resulting in a low classification accuracy of 13%. Among these correctly classified inputs, none were verified to be robust against perturbations, and no adversarial examples were identified. This suggests that even though no overt vulnerabilities were found, the analysis’s precision was only 2% because it was unable to offer official assurances of robustness or safety. The fact that no accurate predictions could be verified indicates that the model probably has serious flaws. These could include less training, excessive uncertainty around decision boundaries, or an overly cautious verification abstraction (especially with DeepPoly, which can produce loose bounds when neuron activations are unstable). All of these findings point to the need for better training procedures as well as improved verification

methods since they show that the model not only performs poorly in terms of classification accuracy but also lacks the stability and robustness necessary for formal verification.

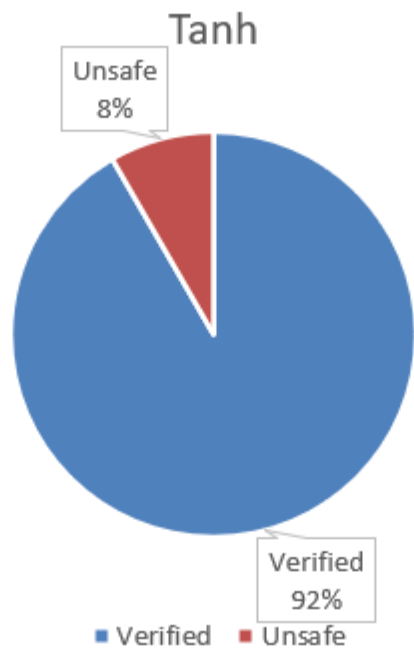


Fig. 2. Performance on Tanh.

Sigmoid An analysis using DeepPoly for verification, based on a set of 100 inputs (as depicted in Figure 4), offered critical insights into both the model’s behavior and the limitations of the verification method itself. The model correctly classified only eleven of these inputs according to their ground truth labels. However, the verifier couldn’t formally confirm the robustness of any of these correctly classified examples, meaning it couldn’t guarantee their predictions would hold true under acceptable perturbations. Additionally, no adversarial examples were discovered, suggesting that while robustness wasn’t proven, no specific vulnerabilities were identified either. This outcome implies that the DeepPoly abstraction might be too broad to provide sufficiently tight constraints for verification, especially when the network’s activations are unstable or its decision boundaries are highly sensitive. Consequently, despite the absence of overt adversarial failures, the network’s close proximity to decision boundaries and the inherent approximation limitations of the verifier prevent a definitive assessment of its safety or resilience.

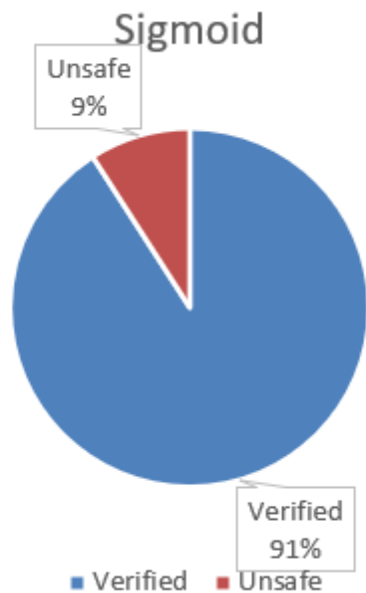


Fig. 3. Performance on Sigmoid.

4 Result

The verification analysis of a neural network model, employing ReLU, Tanh, and Sigmoid activation functions (in Figure. 4) across various configurations, revealed substantial deficiencies in both classification performance and resilience. The ReLU-based model exhibited notably low classification accuracy, correctly predicting only 7% of the evaluated inputs. The verification tool was unable to provide any robustness guarantees for these correct predictions, likely due to weak bounds derived from abstract interpretation, unstable neuron activations, or inadequate training. The observed output boundaries for this model, ranging broadly (e.g., -15 to +15) and often symmetric around zero, indicate low generalization, underfitting, and low confidence in predictions. Similarly, the Tanh-based model classified only 13% of the 100 inputs accurately, and none of these correctly categorized samples were confirmed to be robust. While the Sigmoid activation model successfully pinpointed 11% of the inputs, it didn't offer any guarantee for robustness and no adversarial cases were identified.

This finding really emphasizes the limitations of the DeepPoly method, as it can create bounds that are often too imprecise to accurately capture the network behaviour when it is close to important decision.

The findings really underscore the need for major improvements in how we train and verify models. The low classification accuracies we are seeing across all activation

functions 7% for ReLU, 13% for Tanh and 11% for Sigmoid, point to a lack of proper training, which might be made worse by poor input normalization or datasets that do not represent the real world. To boost performance, we need to retrain using better regularization techniques and a wider variety of data. The fact that we do not have solid robustness guarantees for any of the activation functions shows that our current verification methods, especially those using DeepPoly, just not precise enough. All of this highlights the importance of a well-rounded approach that combines advanced verification methods with improved training strategies to create neural networks that are both dependable and accurate.

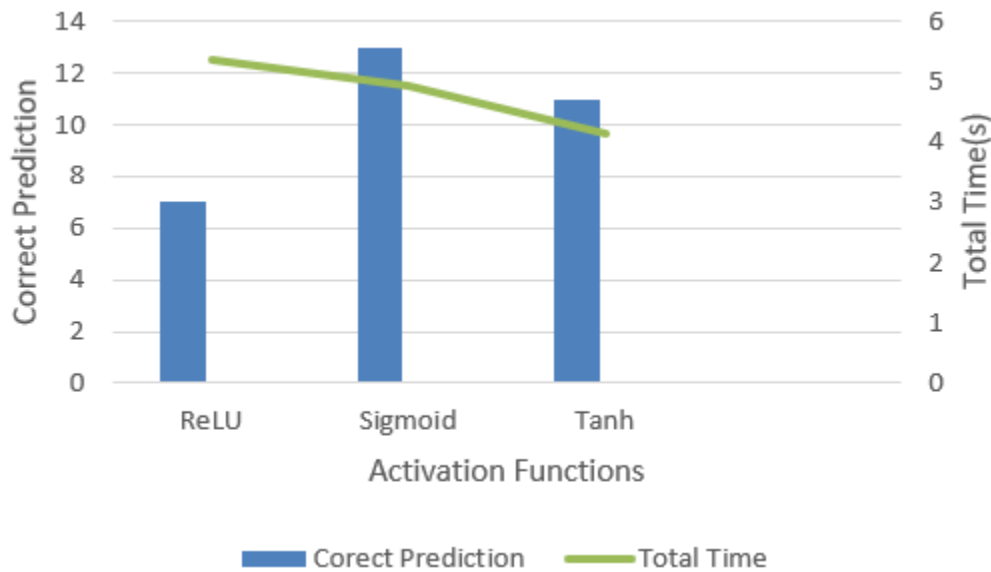


Fig. 4. Performance comparison across different activation functions.

5 Conclusion

This work presents an empirical study focused on the strength of convolutional neural networks, employing the DeepPoly formal verification framework to explore three widely-used activation functions: ReLU, Sigmoid, and Tanh. The DeepPoly demonstrated the efficiency and the scalability, through the search pointed out some major limitations: all the models tested struggled with classification accuracy, and few were able to confirm their resilience against adversarial attacks.

References

- 1) Floyd, R. W.: Assigning Meanings to Programs. Studies in cognitive systems, 65–81 (1993). https://doi.org/10.1007/978-94-011-1793-7_4.
- 2) Hoare, C. A. R.: An axiomatic basis for computer programming. Communications of the ACM 12(10), 576–580 (1969). <https://doi.org/10.1145/363235.363259>.
- 3) Naur, P.: Proof of algorithms by general snapshots. BIT 6(4), 310–316 (1966). <https://doi.org/10.1007/bf01966091>.
- 4) Peter, R.: Rice H. G.. Classes of recursively enumerable sets and their decision problems. Transactions of the American Mathematical Society 74, 358–366 (1954). <https://doi.org/10.2307/2268870>.
- 5) de Moura, L., Bjørner, N.: Z3: An Efficient SMT Solver. Tools and Algorithms for the Construction and Analysis of Systems, 337–340 (2008). https://doi.org/10.1007/978-3-540-78800-3_24.
- 6) Clarke, E. M., Emerson, E. A., Sistla, A. P.: Automatic verification of finite-state concurrent systems using temporal logic specifications. ACM Transactions on Programming Languages and Systems 8(2), 244–263 (1986). <https://doi.org/10.1145/5397.5399>.
- 7) Queille, J. P., Sifakis, J.: Specification and verification of concurrent systems in CESAR. Lecture Notes in Computer Science, 337–351 (1982). https://doi.org/10.1007/3-540-11494-7_22.
- 8) Cousot, P., Cousot, R.: Abstract interpretation. Symposium on Principles of Programming Languages (1977). <https://doi.org/10.1145/512950.512973>.
- 9) Gehr, T., Mirman, M., Drachsler-Cohen, D., Tsankov, P., Chaudhuri, S., Vechev, M.: AI2: Safety and Robustness Certification of Neural Networks with Abstract Interpretation. 2018 IEEE Symposium on Security and Privacy (SP), 3–18 (2018). <https://doi.org/10.1109/SP.2018.00058>.
- 10) Singh, G., Gehr, T., Mirman, M., Püschel, M., Vechev, M. T.: Fast and Effective Robustness Certification. Neural Information Processing Systems (2018). <https://api.semanticscholar.org/CorpusID:53960414>.
- 11) Singh, G., Gehr, T., Püschel, M., Vechev, M.: An abstract domain for certifying neural networks. 3(POPL), 1–30 (2019). <https://doi.org/10.1145/3290354>.
- 12) Turing, A.: Checking a Large Routine (1949).
- 13) Burstall, R.: Program Proving as Hand Simulation with a Little Induction. IFIP Congress (2018). <https://api.semanticscholar.org/CorpusID:46614389>.
- 14) Dijkstra, E. W.: Guarded commands, nondeterminacy and formal derivation of programs. Communications of the ACM 18(8), 453–457 (1975). <https://doi.org/10.1145/360933.360975>.
- 15) Kirchner, F., Kosmatov, N., Prevosto, V., Signoles, J., Yakobowski, B.: Frama- C: A software analysis perspective. Formal Aspects of Computing 27(3), 573–609 (2015). <https://doi.org/10.1007/s00165-014-0326-7>.
- 16) Filliâtre, J. C.: Deductive Program Verification with Why3 — A Tutorial (2016).
- 17) Pnueli, A.: The temporal logic of programs. 18th Annual Symposium on Foundations of Computer Science (sfcs 1977) (1977). <https://doi.org/10.1109/sfcs.1977.32>.
- 18) Grumberg, O., Veith, H.: 25 Years Of Model Checking: History, Achievements, Perspectives. Springer, Berlin (2008).
- 19) Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. <http://www.deeplearningbook.org>
- 20) Nair, V., Hinton, G. E.: Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th International Conference on International Conference on Machine Learning (ICML’10), 807–814. Omnipress, Madison, WI, USA (2010).
- 21) Bunel, R., Turkaslan, I., Kohli, P., Pawan, K. M.: A Unified View of Piecewise Linear Neural Network Verification. arXiv.org (2017). <http://arxiv.org/abs/1711.00455>.
- 22) Szegedy, C., et al.: Intriguing properties of neural networks. International Conference on Learning Representations (2017). <https://api.semanticscholar.org/CorpusID:604334>.
- 23) Miné, A.: Relational Abstract Domains for the Detection of Floating-Point Run-Time Errors. arXiv.org (2024). <https://doi.org/10.48550/arXiv.cs/0703077>

- 24) Priyadarshi, R., Vikram, R., Huang, Z., Yang, T., Rathore, R. S.: Enhancing Coverage in Wireless Sensor Networks Using Machine Learning Techniques (2024).<https://doi.org/10.1109/intcec61833.2024.10603264>.
- 25) Swaroop, A., et al.: A Comprehensive Overview of Formal Methods and Deep Learning for Verification and Optimization. 2024 International Conference on Decision Aid Sciences and Applications (DASA), 1–6 (2024). <https://doi.org/10.1109/DASA63652.2024.10836654>.
- 26) Virtual Self Driving Car using the Techniques of Image Processing and Deep Neural Networks. International Journal of Engineering and Advanced Technology 9(5),190–195 (2020). <https://doi.org/10.35940/ijeat.d9070.069520>
- 27) Katz, G., Barrett, C., Dill, D. L., Julian, K., Kochenderfer, M. J.: Towards Proving the Adversarial Robustness of Deep Neural Networks. Electronic Proceedings in Theoretical Computer Science 257, 19–26 (2017). <https://doi.org/10.4204/eptcs.257.3>.
- 28) Carlini, N., Wagner, D.: Towards Evaluating the Robustness of Neural Networks. 2017 IEEE Symposium on Security and Privacy (SP), 39–57 (2017). <https://doi.org/10.1109/SP.2017.49>.

Comparative Analysis of Damping Materials for Vibration Control in Military Vehicle-Mounted Electronic Warfare Systems

Suhas Ravindra Kokate¹; Dr.S.S.Ohol²; Dr.K.P.Wani³

^{1,2}COEP Technological University, Pune, India

³Automotive Research Association of India, Pune, India
kokatesr23.mech@coeptech.ac.in

Abstract The core issue addressed in this research is the excessive transmission of random vibrations in high-speed military vehicles, which can compromise the performance and reliability of vehicle-mounted electronic warfare (EW) systems. Insufficient damping can lead to misalignment, inaccurate signal detection, and potential system failures. This project aims to design and develop advanced damping solutions to mitigate the effects of shocks, bumps, and vibrations encountered by military vehicles, thereby enhancing the performance and lifespan of critical EW components. The study focused on creating dampers using five different materials: nitrile, neoprene, Sorbothane, EPDM, and silicone rubber. The simulations were conducted in ANSYS, following the MIL-STD-810 H standards, including static, modal, harmonic, and random vibration analyses. Based on simulation metrics such as acceleration response and transmissibility, Nitrile rubber with Shore A60 hardness demonstrated superior vibration reduction and mechanical integrity. These dampers were fabricated in various shapes and dimensions, fitted to the R-system, detection unit, and jammer system. The fabricated Nitrile dampers were integrated into the EW system and subjected to real-world random vibration lab testing. The experimental findings were closely related to the simulation results, demonstrating the models' accuracy and the material's performance. This study confirms that Nitrile rubber is the most effective damping solution among all materials, offering a robust, reliable method to minimize vibrations and improve EW system stability in military vehicle environments.

Keywords: Random Vibration, Rubber performance, Damping material selection

1. Introduction

Heavy military vehicles driving on rough grounds produce large low-frequency vibrations (up to 20 Hz), which negatively impact the performance of vehicle-mounted radar, masts and electronic warfare (EW) systems, leading to physical displacement, misalignment, and reduced signal accuracy. To prevent these problems, this research examines the application of elastomer-based dampers, namely Nitrile rubber, for vibration control in radar support structures. Previous studies, such as Younes et al. (2020), illustrated that shock hammer excitation coupled with accelerometer measurement provides consistent data for modal damping analysis. Also, research by Premnath and Gopi (2023) and Chen et al. (2022) confirmed the better damping performance of materials such as Nitrile, Neoprene, and EPDM compared to conventional metal dampers, particularly under random and low-frequency excitation typical of military environments.

The work entails the development and manufacturing of rubber-based dampers, CAD-modeled and analyzed with ANSYS for vibration characteristics under load conditions. Experimental testing was performed with a vibration shaker rig, according to MIL-STD-810H standards, using accelerometers to record performance parameters. Results indicated a strong connection between simulation and experimental outcomes, confirming the accuracy of the nitrile. Dampers are highly effective in minimizing destructive vibrations, improving radar stability, and electronic equipment reliability is increasing in the presence of challenging terrain conditions.

2. Model Of System

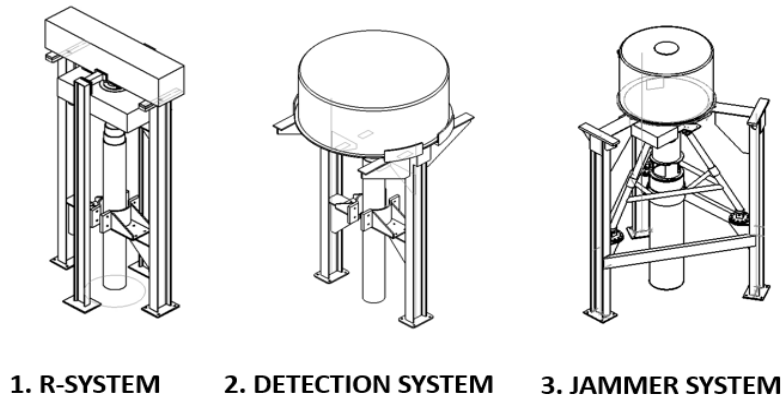


Fig 2: CAD Model of EW System

3. Mathematical Modeling

The dynamic behavior of mechanical systems subject to vibrations can be described by the classical second-order differential equation of motion for a damped single-degree-of-freedom (SDOF) system:

$$[m]\{\ddot{x}(t)\} + [c]\{\dot{x}(t)\} + [k]\{x(t)\} = \{F(t)\} \quad (1)$$

Where $[m]$ is the mass matrix (kg), $[c]$ is the damping matrix (n·s/m), and $[k]$ is the stiffness matrix (N/m), and $x(t)$ is the displacement vector a function of time, $c\dot{x}(t)$ is used to represent the energy lost due to damping, while $kx(t)$ represents the restoring force resulting from the system's stiffness.

Table 1. Material and Geometric Properties of Pneumatic Mast

Material	Young's Modulus (Gpa)	Yield Strength (Mpa)	Outer Diameter (mm)	Inner Diameter (mm)	Length (mm)
Aluminium 6063	69	276	150	138	1600

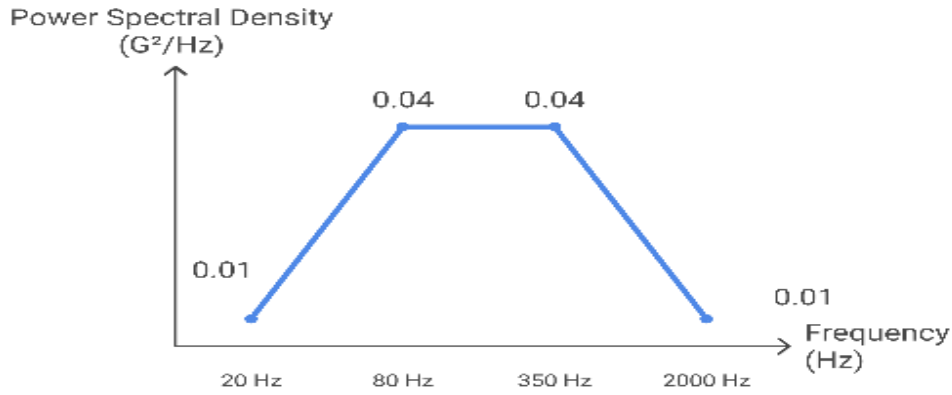
To ensure structural integrity, the load capacity of the pneumatic mast (Aluminium 6063) was evaluated under a compressive load of **1075 N**.

Moment of Inertia (I) formula $I = \frac{\pi}{64}(D^4 - d^4) = 6.875 \times 10^{-6} m^4$ (2)

Buckling Load (P_{cr}): $P_{cr} = \frac{\pi^2 EI}{(KL)^2} = 580,078 N$ (3)

Yield Load (P_{yield}): $A \times \sigma_y$; $A = \frac{\pi}{4}(D^2 - d^2) = 750,000 N$ (4)

The mast is structurally safe under the applied load (1075 N), as it is far below both the buckling and yield thresholds. The excitation input for the random vibration analysis was defined using a Power Spectral Density (PSD) curve specified in MIL-STD-810H, simulating the unpredictable vibration environment encountered by military vehicle-mounted systems. This standardized PSD profile was set in ANSYS as the base excitation to assess how various elastomeric materials behave on a dynamic scale, particularly their damping effectiveness under random vibrations.



Power Spectral Density vs. Frequency

Fig 3: PSD Graph

The Calculation of G_{rms} Using the Trapezoidal Integration Method:

$$\int_{f_1}^{f_2} PSD(f)df = \sum \frac{PSD_i + PSD_{i+1}}{2} \times f_{i+1} - f_i \quad (5)$$

4. FEA and Structural Performance Evaluation

A finite element-based modal analysis was conducted in ANSYS on the R-system to extract its fundamental dynamic characteristics, including natural frequencies and associated mode shapes. To measure the system's response to operational excitations, the first nine natural frequencies were calculated. The CAD model of the structure included a defined set of mechanical boundary conditions along with the material properties of Aluminum 6063, allowing for precision during predictions of the structure's resonance characteristics.

Table 2. Natural Frequencies of R-system model

Modes shape	1	2	3	4	5	6	7	8	9
Frequency [Hz]	9.7	12.4	21.5	21.6	37.3	42.1	45.0	53.4	54.3

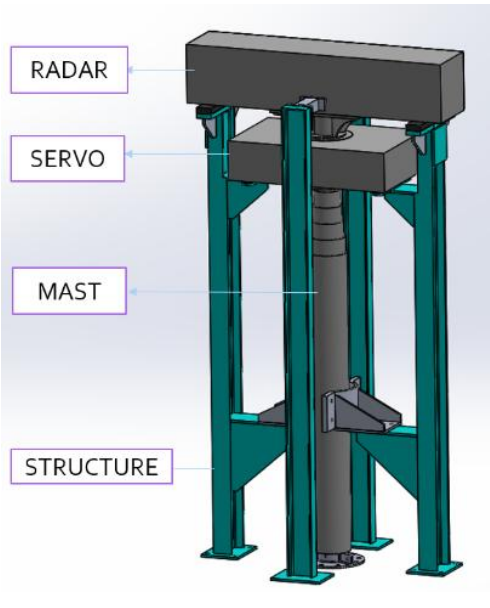


Fig.4 R-System Without Damper

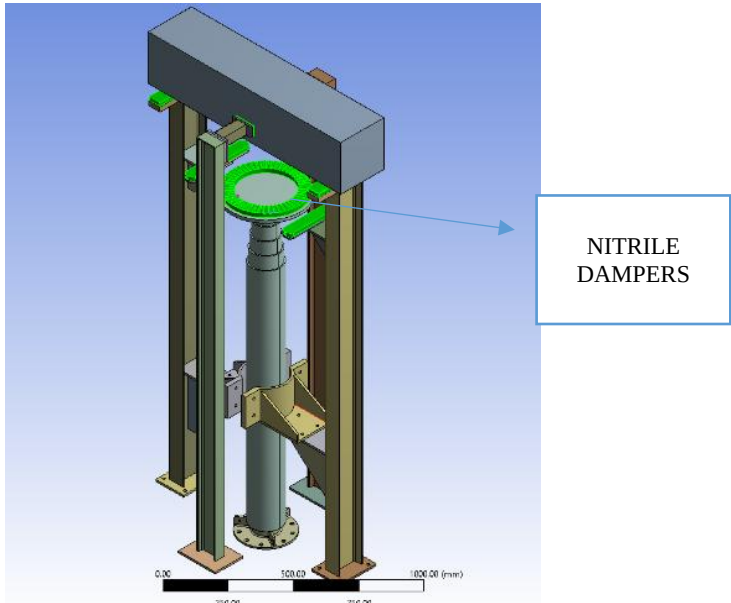


Fig.5 R-System with Damper

Table 3: Comparative Analysis Results: R-System with and Without Damper

MATERIAL	WITHOUT DAMPER	NITRILE	NEOPRENE	SORBPTHANE	EDPM	SILICON
Max Stress (MPa)	75.45	2.67	2.90	3.099	3.05	3.49
Displacement (mm)	2.03	0.08	0.10	0.11	0.11	0.16

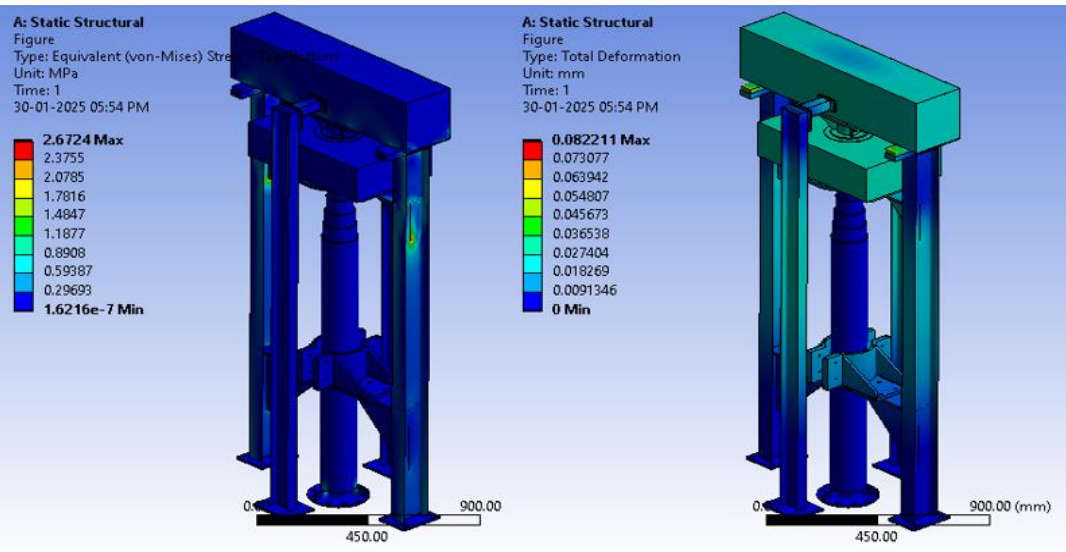


Fig.6 Stress–Displacement Behaviour with Nitrile Damper

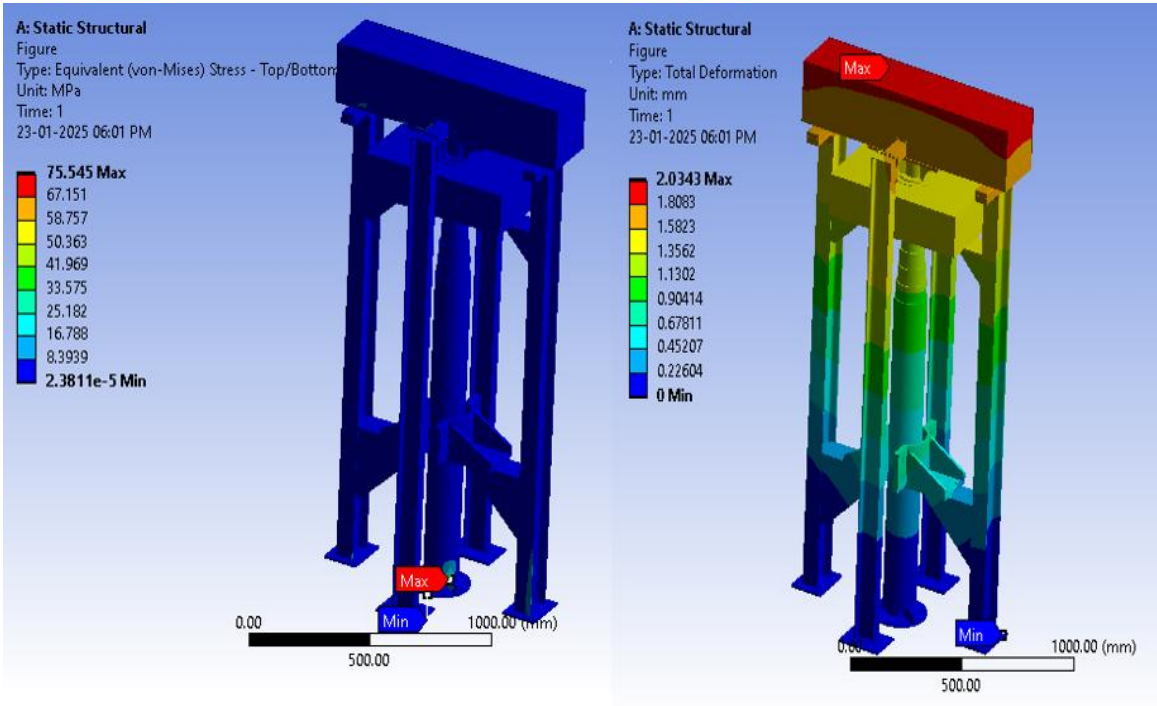


Fig.7 Stress–Displacement Behaviour Without Damper

4.1. Random Vibration Analysis

A random vibration analysis was performed on different damping materials under identical boundary conditions and excitation profiles. The Power Spectral Density (PSD) input was kept the same for each material to ensure a uniform comparison. The output acceleration response PSD was obtained from the analysis for each material. The Root Mean Square (RMS) acceleration G_{rms} was calculated by integrating the PSD curve across the frequency range of interest using the trapezoidal integration method

Table 4: Comparative Output G_{rms} Results Response

MATERIAL	WITHOUT DAMPER	NITRILE	NEOPRENE	SORBOTHANE	EPDM	SILICON
G_{rms} (g)	27.88	5.85	7	6.3	6.3	6

Each damping material was obtained through the use of response PSDs, with random vibration analysis performed. The response spectra provided were used to obtain the RMS acceleration values, which serve to measure the amount of vibration passed through the system. All materials tested, the lowest RMS response was obtained from nitrile rubber, which suggests it has the greatest energy-damping capabilities relative to the minimization of dynamic excitation. This supports the case for the use of nitrile as a damping material for important electronics in military systems that operate in high-vibration environments.

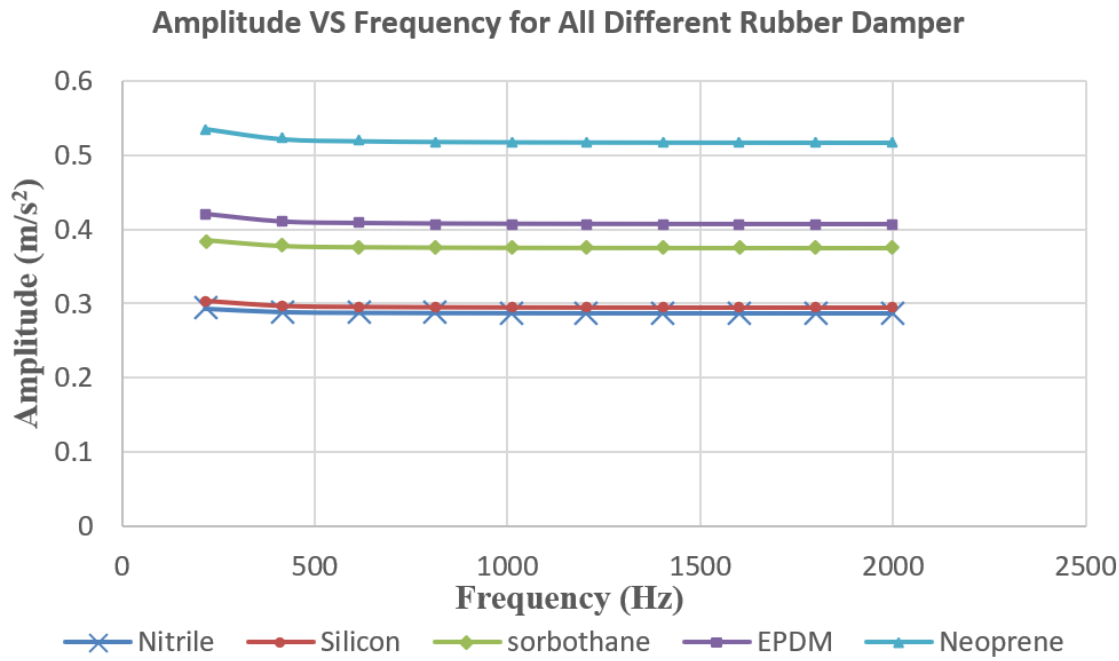


Fig.8 Amplitude VS Frequency for All Different Rubber Dampers

Harmonic response simulation results showed a greater reduction in amplitude of vibration for the system with Nitrile rubber treatment compared to other elastomeric dampers. Lesser amplitudes across the frequency spectrum, particularly at the resonant one, prove that Nitrile dampers effectively suppress the transfer of vibrational energy and help in attaining lower chances of structural fatigue and hence are better in terms of operation.

5. Fabrication of Nitrile Rubber Dampers

Based on the simulation analysis, Nitrile rubber was selected as the optimal damping material. Dampers were subsequently fabricated in various sizes and geometries tailored to meet the specific dimensional and mounting requirements of the R-System.

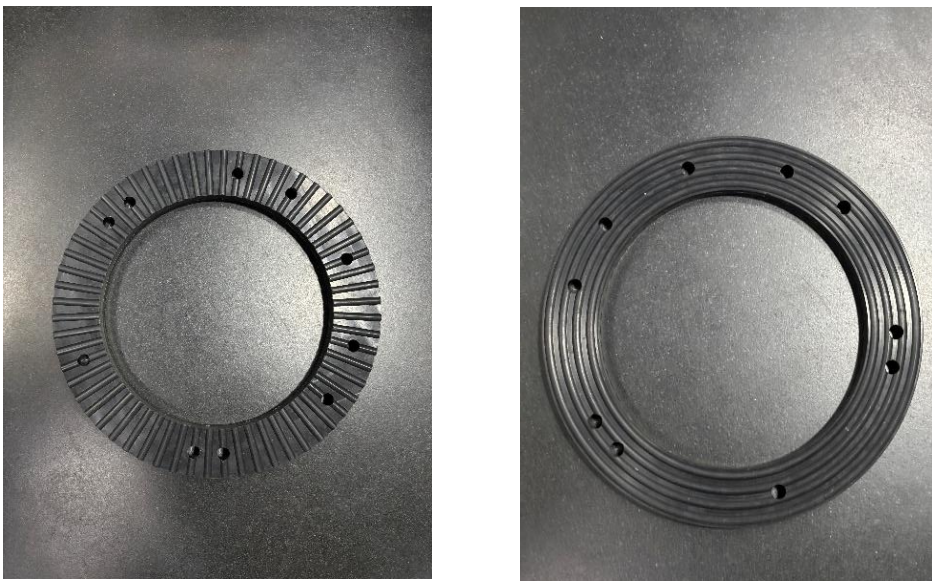


Figure 9. Fabricated Nitrile Damper (A50 Shore)

Table 6: Mechanical Parameters of Selected Rubbers

MATERIAL	Density (kg/m ³)	Young's Modulus (MPa)	Poisson's Ratio	Damping Ratio
Nitrile	1200	6	0.49	0.04
Neoprene	1200	9	0.48	0.03
Sorbothane	1200	5	0.49	0.5
Epdm	1100	7	0.47	0.03
Silicon	1100	5	0.47	0.04

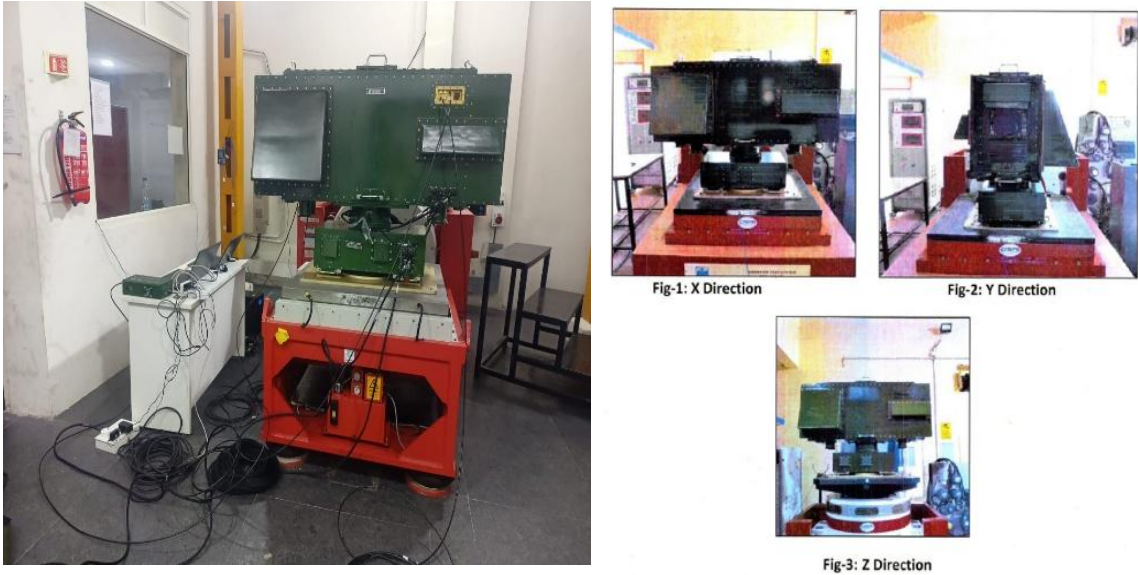


Fig 10. Experimental Testing Setup

6. Random Vibration Testing

Random vibration testing was performed following MIL-STD-810H, with the fabricated nitrile rubber dampers integrated into the system to evaluate their vibration attenuation performance under standardized environmental conditions.

Table 8: Testing G_{rms} Results	
DEMAND	7.32 G RMS
CONTROL	6.044 G RMS

Final output results from lab testing showed a demand of 7.32 Grms and control of 6.044 Grms. No structural deformities or damage were observed during or after the test, and the performance of the tested units was satisfactory

7. Conclusion

This research focused on reducing random Vibration in Military Vehicle-Mounted EW Systems. The evaluation included simulations in ANSYS software as well as lab testing following MIL-STD-810H guidelines. Among the

five materials, Nitrile Rubber (Shore A60) showed the most effective damping to the system, providing a well-balanced combination of damping efficiency and structural stability. The fabricated dampers were integrated into the R-system, and random vibration lab testing was done following MIL-STD810H standards. The close response between simulation and experimental data shows the accuracy and dependability of the modelling approach. The performance was found to be satisfactory during lab testing. With no structural damage or deformation observed. This work not only improves the stability and operational effectiveness of EW systems in challenging military conditions but also provides a proven approach for choosing and implementing damping solutions. Future research will focus on the use and integration of machine learning to optimize the selection and performance of damping materials.

8. References

- 1) MIL-STD-810H, *Test Method Standard for Environmental Engineering Considerations and Laboratory Tests*, U.S. Department of Defense, Jan. 2019.
- 2) 2.T.D. Gillespie, *Heavy Truck Ride*, SAE Technical Paper 850001, 1985. <https://doi.org/10.4271/850001>.
- 3) M. Younes, N. Abed, and D. Al-Nimr, “Experimental investigation of natural frequencies and damping using three modal testing techniques,” *Journal of Vibration and Control*, vol. 26, no. 15–16, pp. 1394–1407, 2020. <https://doi.org/10.1177/1077546319880773>.
- 4) R. Premnath and S. Gopi, “Dynamic behavior analysis of metallic and elastomeric dampers under vibratory excitation,” *Materials Today: Proceedings*, vol. 79, pp. 313–319, 2023. <https://doi.org/10.1016/j.matpr.2023.02.358>.
- 5) K. Zaman, M. Alam, and T. Islam, “Comparative study of vibration-damping materials: An experimental analysis,” *Journal of Mechanical Science and Technology*, vol. 37, no. 3, pp. 1183–1190, 2023. <https://doi.org/10.1007/s12206-023-0221-7>.
- 6) B. Chen, J. Dai, T. Song, and Q. Guan, “Research and Development of High-Performance High-Damping Rubber Materials for High-Damping Rubber Isolation Bearings: A Review,” *Polymers*, vol. 14, no. 15, p. 3022, 2022.
- 7) A Jaskot, *Analysis of vibrations in a dynamic damper*, IOP Conf. Series: Materials Science and Engineering, **1252** (1), 2022.
- 8) D.A. Oshmarin, N.A. Iurlova, N.V. Sevodina, M.A. Iurlov, *A comparative analysis of two approaches to damping of vibrations*, Proc. of the European Conference on Fracture, Institute of Continuous Media Mechanics UB RAS, Russia.
- 9) T. Imthiyaz Ahamed et al., *Implementation of Magneto-rheological dampers in bumpers of Automobiles*, Procedia Engineering, 97, 2014.

Automated Drone Detection Using YOLOv11 Object Detection Framework

Pranav Kaingade¹, Dr.A.G.Thosar² and Prof.S.U.Ghunage¹

1.Dept. of Manufacturing Engineering and Industrial Management, COEP Technological University Pune, India.

2.Dept. of Electrical Engineering, COEP Technological University Pune, India

Abstract

The widespread adoption of drones for leisure and commercial purposes has led to increasing security risks, particularly in unauthorized or no-fly zones. To address this growing concern, this paper proposes a fully automated drone identification system based on the latest YOLOv11 (You Only Look Once, version 11) object detection architecture. Unlike traditional object detectors, the presented model is fine-tuned to detect drones with enhanced precision, even in challenging scenarios that involve busy visual backgrounds and miniature aerial objects. The detection framework is trained using a well-constructed dataset featuring a broad range of drone categories, dimensions, and viewpoints to ensure consistent performance across diverse operational conditions. By incorporating architectural advancements in YOLOv11, the system achieves real-time inference while maintaining high accuracy, making it ideal for surveillance, airspace control, and the safeguarding of high-security zones. Benchmark results show that the proposed model surpasses earlier YOLO versions and competing methods in terms of detection accuracy, recall rate, and processing efficiency. This research demonstrates the viability of YOLOv11 as a dependable and scalable solution for drone monitoring applications.

Keywords-Yolov11, Drone detection, Security, Object Detection

1. Introduction

The rapid advancement and widespread adoption of drone technology have revolutionized sectors like logistics, agriculture, surveillance, and entertainment. However, the accessibility and affordability of drones have also introduced major challenges, particularly in terms of security, privacy, and safety. Unauthorized drone activities near restricted areas, critical infrastructure, and private spaces have raised concerns about surveillance, smuggling, and airspace breaches. As a result, the development of efficient and real-time drone detection systems has become crucial.

Traditional detection approaches such as radar, acoustic sensors, and radio frequency (RF) analysis are often limited by accuracy, scalability, and environmental adaptability. In contrast, deep learning based computer vision methods offer superior accuracy and responsiveness in detecting drones under various conditions. Among these, the YOLO (You Only Look Once) family of object detection algorithms has emerged as a powerful solution due to its real-time speed and high detection accuracy.

A range of studies have explored drone detection using deep learning and machine learning techniques. For instance, improved detection of micro and mini drones was achieved using YOLOv3 by increasing input resolution and optimizing anchor boxes [1], while another study utilized YOLOv3 with transfer learning to achieve an average confidence score of 88.9% in real-time deployment on NVIDIA Jetson TX2 [2]. Other researchers proposed hybrid approaches combining Haar cascade classifiers and CNNs to identify drone vendor models, achieving 89% accuracy with limited training data [3]. A YOLOv3-based embedded system achieved an exceptional accuracy of 99.99% for small drone detection under complex conditions, using 106 convolution layers and up sampling techniques [4].

In terms of datasets, the introduction of the Drone-type dataset [5] provided a benchmark for classifying various drone models, with YOLOv3 achieving a mAP of 99.5%. Broader machine learning-based surveys have assessed

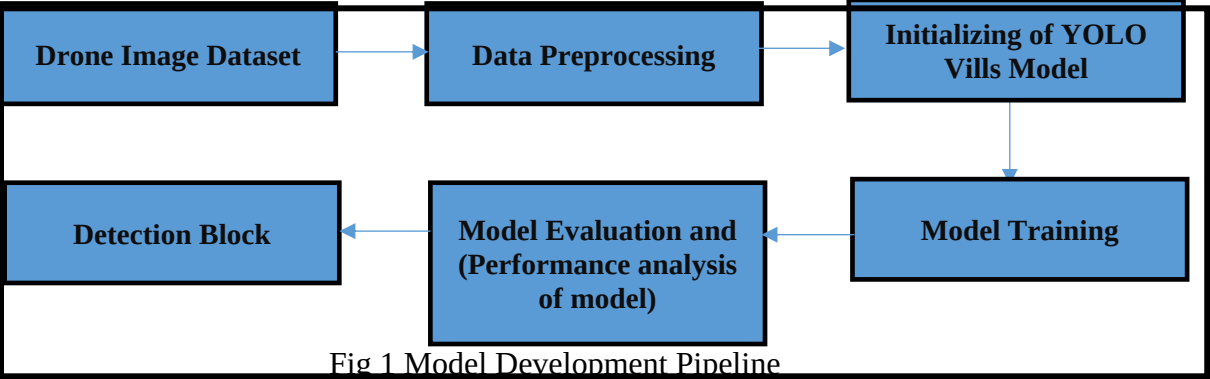
classifiers like CNNs, SVMs, and LSTMs across multiple sensing domains highlighting the need for standardized datasets and collaborative multimodal detection approaches [6,13].

Deep learning models such as YOLOv4 have also shown strong performance in real-time drone detection with a mAP of 74.36%, effectively distinguishing drones from similar objects like birds [7]. Surveys have highlighted the dominance of one-stage detectors like YOLO and SSD for real-time applications due to their superior speed and accuracy [8]. In response to the challenges of detecting fast-moving small drones, recent studies have introduced modified YOLOv8-based models using multi-scale image fusion and feature enhancements to achieve high fps and detection precision [9,10,12].

This research builds upon these advances by proposing a drone detection system based on YOLOv11[14,15], the latest in the YOLO family. Designed for enhanced accuracy, speed, and small object detection, YOLOv11 aims to address existing gaps by handling varying drone sizes, speeds, and challenging visual conditions. This work contributes to the ongoing effort to develop practical, scalable, and intelligent drone surveillance solutions. This paper is arranged as: section 2 introduces proposed system. Data acquisition and processing is discussed in section 3. Model training discussed in section 4. Implementation of system explained in section 5. The results and evaluation discussed in section 6. Finally, the work is concluded in section 7.

2. Proposed system

The proposed system for drone detection leverages a YOLOv11-based deep learning framework and comprises two main stages: the model development pipeline and the detection block. The process commences with the collection of a drone image dataset, which is subjected to a data preprocessing phase. This phase involves operations such as image resizing, normalization, and annotation formatting to ensure the dataset is compatible with the training pipeline and to enhance the model’s generalization capabilities. Following preprocessing, the YOLOv11 model is initialized with suitable configurations and hyperparameters, specifically tuned for the drone detection task. The model is then trained using the pre-processed dataset, allowing it to learn the distinguishing features of drones across varying scenarios. Upon completion of training, the model's performance is rigorously evaluated using key metrics such as precision, recall, F1-score, and mean average precision (mAP) to assess its detection accuracy and reliability. The best-performing model is subsequently integrated into the real-time detection block.



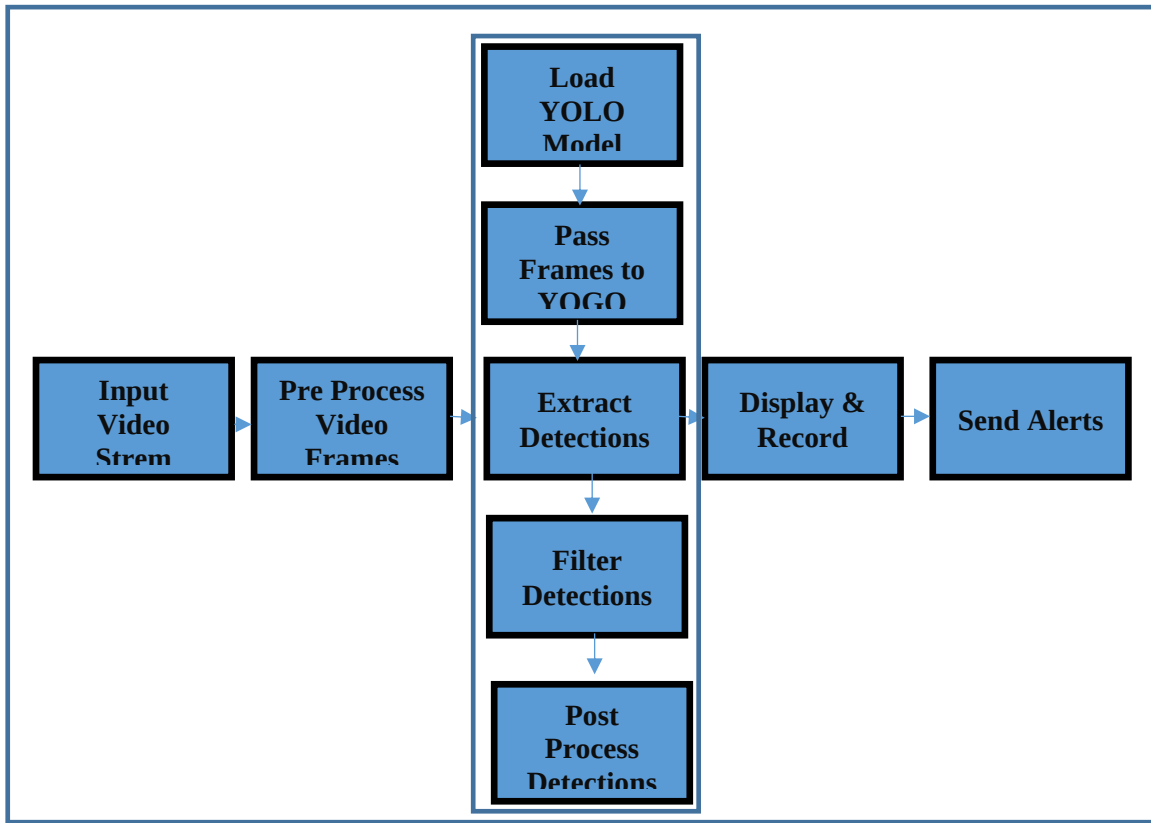


Fig.2. Detection block

The detection block handles live video input for drone detection. It begins by capturing a continuous video stream, which is processed frame by frame. Each frame undergoes preprocessing and is fed into the trained YOLOv11 model for inference. The model performs object detection to identify potential drones, and the results are filtered using thresholds and non-maximum suppression techniques to eliminate redundant or false detections. Post-processing steps refine the final output, which is then displayed to the user. Additionally, the system is equipped with functionalities to record the detection results and issue alerts upon identifying drones. This end-to-end approach provides a robust, accurate, and scalable solution for real-time automated drone detection using the advanced capabilities of the YOLOv11 architecture.

3.Data acquisition and processing

The images of drones needs to be collected to train proposed model. The dataset used in this was collected from various open-source platforms to make sure it covers different visuals and improves the model's performance during training, validation, and testing. It includes 9376 images sourced from platforms like Open Images, Kaggle, and Roboflow carefully selected to cover different lighting conditions, environments, and drone types, helping the model to work better in real-world situations. The dataset was divided into three parts for effective training and evaluation: the training set has 7,016 images (75%) to help the model learn to identify drones in different settings; the validation set includes 1,905 images (20%) used to fine-tune the model and avoid overfitting; and the test set contains 455 images (5%) to evaluate how well the model works on new data. The dataset includes various drone sizes and angles in different backgrounds like urban, rural, and indoor environments. Each image is properly annotated with bounding boxes and class labels to ensure accurate detection, even in tough conditions like low light, motion blur, and occlusion. The annotations were done using tools like Labelmg (for marking bounding boxes) and Roboflow (for managing and enhancing the dataset). The dataset was gathered from Open Images (offering a wide range of drone images), Kaggle (with additional datasets from the community), and Roboflow (providing well-organized and pre-

processed drone datasets). This well-balanced and diverse dataset ensures that the model performs well in real-world situations, making it suitable for use in surveillance, security, and airspace monitoring.

4. Model Training

The YOLOv11-S model was successfully trained for drone detection using a pre-trained model with a properly organized dataset. The training was carried out for 80 epochs with an image size of 640x640 pixels. The training process generated performance graphs, including loss curves and precision-recall charts, which helped in understanding the performance of newly trained model.

The model was trained on Google Colab using a T4 GPU for about 3 hours, which worked efficiently considering the dataset size and model complexity. During training, the model adjusted its weights based on the dataset, improving its ability to detect drones in different conditions. Performance indicators like loss, precision, recall, and mAP (mean Average Precision) improved steadily, showing that the model was learning effectively without major issues like overfitting or underfitting.

The YOLOv11-S variant was chosen because it strikes a good balance between speed and accuracy. Since it is a lighter version, YOLOv11-S is well-suited for real-time object detection, making it ideal for tasks where quick response and low delay are crucial, like drone detection. Its smaller size reduces the computational load while maintaining high detection accuracy, which is especially useful for deploying on edge devices or low-power systems.

The final model weights were saved after training, ready for use in detecting drones or for further fine-tuning. The successful training of the YOLOv11 model confirms that it can accurately detect drones even in complex environments with high accuracy and efficiency.

5. Real time Drone Detection Implementation:

The system is implemented by using camera as source and Laptop as processing unit the drones were brought in front of the camera while flying in different environmental conditions. The below are the images demonstrated drone detection under various weather conditions, showcasing high confidence scores. Confidence score represents estimated probability of detection. Real-time detection is effectively achieved even with multiple drones present in the scene.



Fig 3: Drone Detection in cloudy weather

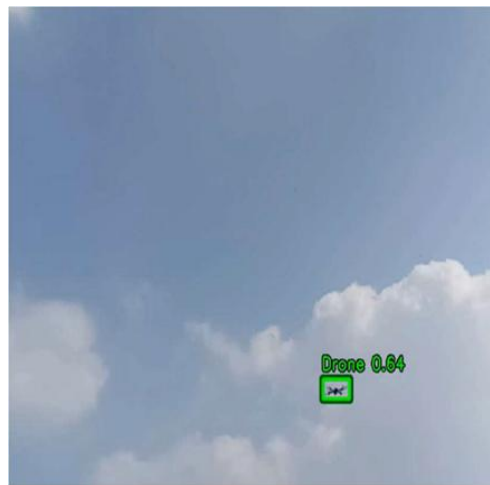


Fig 4: Drone Detection in Sunny weather

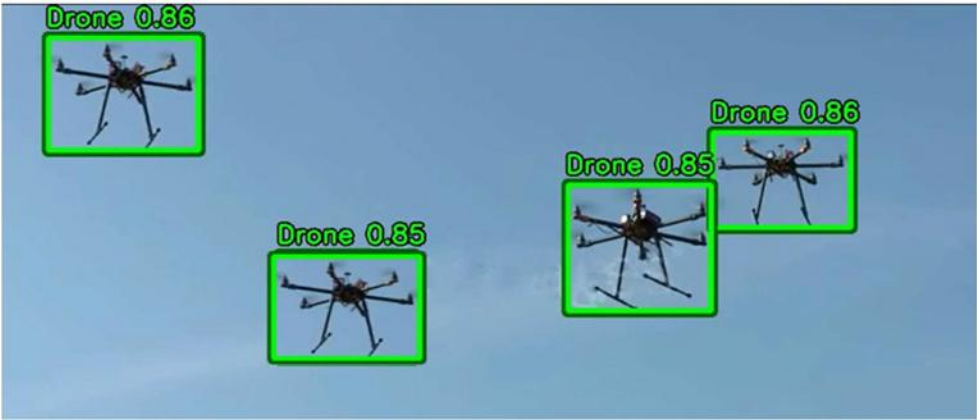


Fig 5:Multiple Drone Detection in clear weather

6. Results and Discussion

The performance of model is evaluated using parameters like precision ,recall,f1-score and mean average precision (mAP).The model prediction error is represented by loss, which quantifies the discrepancy between predicted and actual values during training. Precision refers to the proportion of true positive detections among all positive predictions, reflecting the accuracy of the model’s positive calls. Recall indicates the proportion of true positive detections among all actual positives, demonstrating the model’s ability to locate every relevant instance. Finally, mean Average Precision (mAP) measures overall detection performance by averaging precision values across different recall levels and classes.F1- score is the harmonic mean of precision and recall.

Table 1: Performance Metrics for Drone Detection

Model	Yolo V11-S
Image size	640
Epochs	80
total Parfameters	80
Flops (G)	21.5 Glops
Precision (%)	96.1%
Recall (%)	96%
F1-Score (va) (%)	97.9%
aAP 50-95 (val) (%)	72.1%
Av. Latency (Val) (ms)	113ms

The YOLOv11-S model, trained on 640×640 images for 80 epochs, has 9.4 million parameters and operates at 21.5 GFLOPs. It achieves high accuracy with 96.1% precision, 96.4% recall, and a 96% F1-score. The model performs strongly with a mAP@50 of 97.9% and mAP@50-95(overlap) of 72.1%. Its average detection latency is just 113 ms, making it suitable for real-time applications.

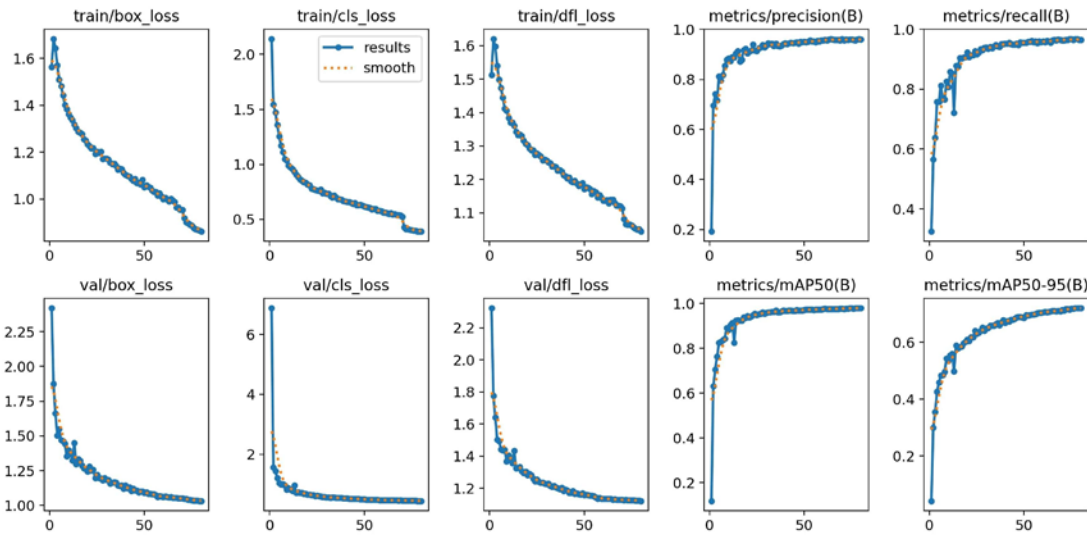


Fig 6:Graphical Representation of Performance Analysis for YOLOv11-S model

The above image fig.6 illustrates the training and validation performance metrics of the YOLOv11-S model over 80 epochs. As the epochs increase, the losses are getting decreased while a significant increment in precision, recall, mAP. This indicates effective learning and strong detection capability of the newly trained model.

7. Conclusion

The proposed automated drone detection system exhibits decent performance, validating its applicability in real-world scenarios such as security surveillance, airspace monitoring, and drone traffic management. The model achieves good precision at higher confidence thresholds, ensuring that detected drones are almost always correctly classified. The normalized confusion matrix shows approximately 98% detection accuracy for drones and complete accuracy for background objects, underscoring the system's strong discriminatory capabilities.

Additionally, a high recall score at very low confidence thresholds highlights the model's robustness in detecting nearly all drone instances, a crucial feature in scenarios where missed detections could pose serious risks. Over the course of 80 training epochs, the model demonstrated consistent improvements, characterized by declining box, classification, and distribution focal losses, along with increasing values of precision, recall, and mean Average Precision (mAP).

Moreover, the Precision-Confidence and Recall-Confidence curves demonstrate the model's consistent performance across a wide range of confidence thresholds, reinforcing its reliability in dynamic and uncertain environments. Importantly, the system achieves low detection latency, which significantly enhances its suitability for real-time applications where timely responses are critical. With a balanced combination of accuracy, responsiveness, and adaptability, the proposed system presents a strong foundation for practical deployment. Further refinements can focus on optimizing performance under varying environmental conditions and expanding its capabilities for multi-drone detection and tracking.

Looking ahead, the system can be extended for real-time multi-camera support on edge devices, integration with UAV traffic management and geofencing systems, and drone classification with intent prediction. Enhancing robustness using diverse datasets and synthetic data, and enabling onboard deployment or counter-drone responses, can further evolve it into a complete aerial threat mitigation solution.

References

- 1) T. Delleji, H. Fekih, and Z. Chtourou, Deep Learning-Based Approach for Detection and Classification of Micro Mini Drones, Proceedings of the 4th International Conference on Advanced Systems and Emergent Technologies, 2020.
- 2) D. T. W. Xun, Y. L. Lim, and S. Srigrarom, Drone Detection Using YOLOv3 with Transfer Learning on NVIDIA Jetson TX2, International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics, 2021.
- 3) D. R. Lee, W. G. La, and H. Kim, Drone Detection and Identification System Using Artificial Intelligence, IEEE, 2018.
- 4) K. Madasamy, V. Shanmuganathan, V. Kandasamy, M. Y. Lee, and M. Thangadurai,
- 5) Embedded System-Based Object Surveillance Detection System with Small Drone Using Deep YOLO, EURASIP Journal on Image and Video Processing, vol. 2021, no. 1, 2021.
- 6) K. AlDosari, A. Osman, O. Elharrouss, S. Al-Maadeed, and M. Z. Chaari, Drone-type-Set: Drone Types Detection Benchmark for Drone Detection and Tracking, Proceedings of the International Conference on Intelligent Systems and Computer Vision, 2024.
- 7) S. Singha and B. Aydin, Automated Drone Detection Using YOLOv4, MDPI Drones, vol.5, no. 3, 2021, pp. 95.
- 8) M. Sliti and M. Garai, Drone Detection and Classification Approaches Based on ML Algorithms, Proceedings of the 28th Asia Pacific Conference on Communications (APCC), 2023.
- 9) Z. Tan and M. Karakose, Survey and Comparative Study for Drone Detection Using Deep Learning, Proceedings of the International Conference on Data Analytics for Business and Industry, 2022.
- 10) J.-H. Kim, N. Kim, and C. S. Won, High-Speed Drone Detection Based on YOLOv8, IEEE.
- 11) International Conference on Acoustics, Speech and Signal Processing, 2023.
- 12) B. Yilmaz and U. Kutbay, YOLOv8-Based Drone Detection: Performance Analysis and Optimization, MDPI Computers, vol. 13, no. 2, 2024.
- 13) N. Al-Iqubaydhia, A. Alenezia, T. Alanazia, A. Senyora, and N. Alanezia, Deep Learning for Unmanned Aerial Vehicles Detection, Elsevier, vol. 1574-0137, 2023.
- 14) Z. Zhang, “Drone-YOLO: An Efficient Neural Network Method for Target Detection in Drone Images” Drones 2023.
- 15) B. Taha and A. Shoufan, Machine Learning-Based Drone Detection and Classification, IEEE Access, vol. 7, 2019, pp. 49610–49622.
- 16) G. Jocher and J. Qiu, Ultralytics YOLOv11, Version 11.0.0, 2024. [Online]. Available: <https://github.com/ultralytics/ultralytics>.
- 17) A..Rasheed and M. Zarkoosh, YOLOv11 Optimization for Efficient Resource Utilization, arXiv preprint arXiv:2412.14790, 2024. [Online]. Available: <https://doi.org/10.48550/arXiv.2412.14790>.

Spatial Attention Encoder-Decoder for Plant Disease Detection

Debajyoti Dhar¹, Devshri Satyarthi², and K.V. Arya²

debajyotidhar111@gmail.com

devshri@iiitm.ac.in

kvarya@iiitm.ac.in

Multimedia and Information Security Research Group, Department of Computer
Science and Engineering, ABV-IIITM Gwalior-474015, India

Abstract. Plant diseases are a big problem for food security around the world, so we need quick and accurate ways to find them. Traditional manual inspection methods take a long time, make mistakes, and can't be scaled up. Also, current deep learning methods, such as Convolutional Neural Networks and Vision Transformers, have problems with how well they work with computers and how well they can change to real-world situations. It suggests a multidimensional encoder-decoder architecture that uses multi-kernel convolutional layers to improve feature extraction at different scales. It also suggests a new attention mechanism that ensures attention across multiple channels of an intermediate image, which leads to better disease recognition. The model goes through a lot of preprocessing, enhancement, and training on the PlantVillage dataset to deal with problems like class mismatch and domain adaptation. Comparative study with baseline models like AlexNet, VGG, SVM, and Random Forest shows that our method works, as shown by the 97.66% test accuracy. As compared to the current state of models, this design can achieve very high accuracy and precision while having very few parameters. This means that it can also be used for events that happen in real time.

Keywords: Plant Disease Detection · Multi-Kernel CNN · Encoder- Decoder · Attention Mechanism · Smart Agriculture

1 Introduction

Timely and accurate detection of plant diseases is vital for food security and sustainable agriculture. Traditional methods like ELISA and PCR are accurate but slow, costly, and unscalable [16]. With advancements in imaging and deep learning, CNNs and Vision Transformers (ViTs) have become popular, though CNNs lack generalization and ViTs demand high resources. This paper proposes a lightweight encoder-decoder model with multi-kernel convolutions and a novel Cross-Channel Spatial Attention (CCSA) mechanism, enabling robust, interpretable disease detection suitable for real-time deployment in low-resource agricultural environments. The major contributions of this paper are:

- A novel encoder-decoder model with multi-kernel convolution blocks for effective multi-scale feature extraction from plant disease images.
- The introduction of a Cross-Channel Spatial Attention (CCSA) mechanism that jointly attends to spatial and channel-wise features to enhance interpretability and classification accuracy.
- A comprehensive evaluation on the PlantVillage dataset with detailed ablation studies showcasing the superiority of the proposed model over traditional deep learning baselines.
- A lightweight architectural design with minimal computational overhead, making it suitable for real-time deployment in resource-constrained agricultural environments.

The rest of the research paper is organized in the following way:

Section 2 presents a review of existing literature on traditional and modern plant disease detection methods, highlighting the need for improved interpretability and computational efficiency.

Section 3 details the proposed methodology, including dataset preprocessing, the architecture of the encoder-decoder model, and the design of the multi-kernel blocks and CCSA module.

Section 4 discusses experimental results and compares the proposed model's performance with existing models using various evaluation metrics.

Section 5 concludes the paper by summarizing the contributions and outlining directions for future work, particularly concerning deployment on edge devices and potential for multimodal fusion.

2 Related Work

Early plant disease detection relied on visual inspection and lab diagnostics like PCR [21] and ELISA [8], which, despite their accuracy, lacked scalability and affordability. Traditional machine learning approaches were later adopted but suffered from poor generalization and robustness.

Deep learning, particularly Convolutional Neural Networks (CNNs) [14], addressed these issues by learning hierarchical image features. Models like AlexNet [10], VGG [17], ResNet [6], and EfficientNet [18] have been widely used. However, their need for large labeled datasets and limited interpretability has driven interest in methods like data augmentation and transfer learning.

Vision Transformers (ViTs) [20,5] introduced global attention mechanisms beneficial for identifying subtle, overlapping disease symptoms. Despite high accuracy, their computational demands limit real-time deployment. Hybrid CNNTransformer models have shown improved efficiency and accuracy [3,23].

Encoder-decoder models are effective for multiscale feature extraction and spatial localization [2]. Architectures such as SegNet [4] and variants used in hyperspectral and LiDAR domains [7] show promise for lightweight, interpretable, real-time applications [12].

Attention mechanisms further refine feature focus. ViTs use self-attention for global context [22], while CNNs apply spatial attention for localized emphasis [13].

Finally, AI's integration into smart agriculture enables real-time disease detection and targeted treatment strategies [9], with models like CNNs and YOLOv7 successfully applied to tomato leaf disease classification.

3 Methodology

3.1 Dataset and Preprocessing

In this work, used a curated subset of the PlantVillage dataset comprising RGB images of tomato, potato, and pepper bell crops. Each class includes healthy and diseased samples, posing a multi-class classification challenge. The dataset has a class imbalance, with tomato diseases being overrepresented. To address this, we applied geometric augmentations (flipping, rotations, scaling), along with CutMix, MixUp, and Gaussian blur, to simulate real-world variability and enhance generalization.

3.2 Model Architecture

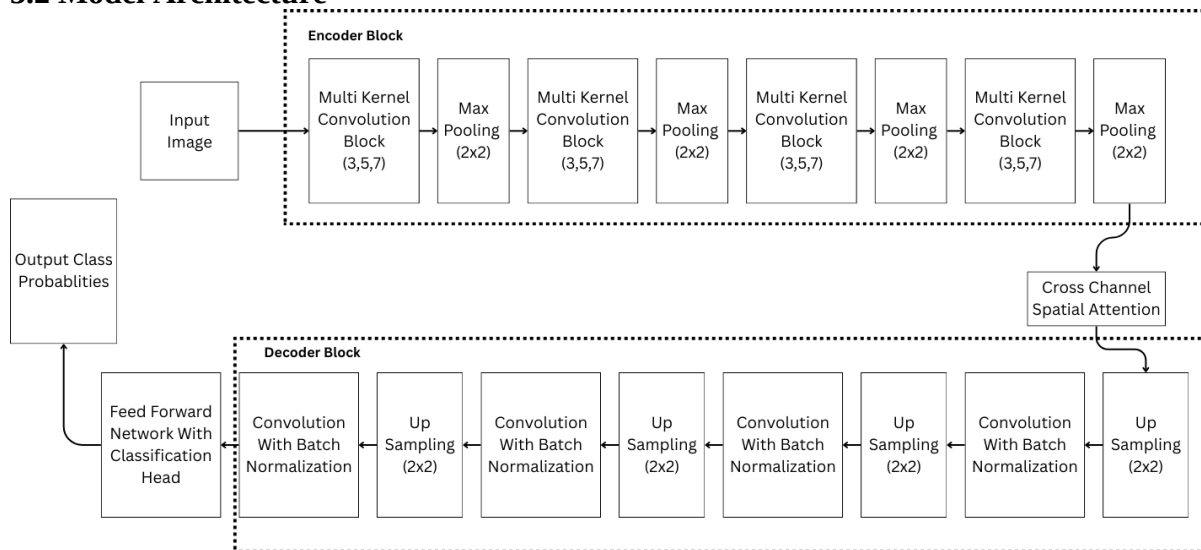


Fig. 1: Proposed Encoder-Decoder Model Architecture

The proposed model as shown in Fig. 1 follows an encoder-decoder design with multi-kernel convolution blocks to capture features at various scales using parallel 3×3 , 5×5 , and 7×7 filters. Skip connections link encoder and decoder layers to retain fine-grained details essential for accurate disease spot localization. A final softmax layer enables multi-class classification

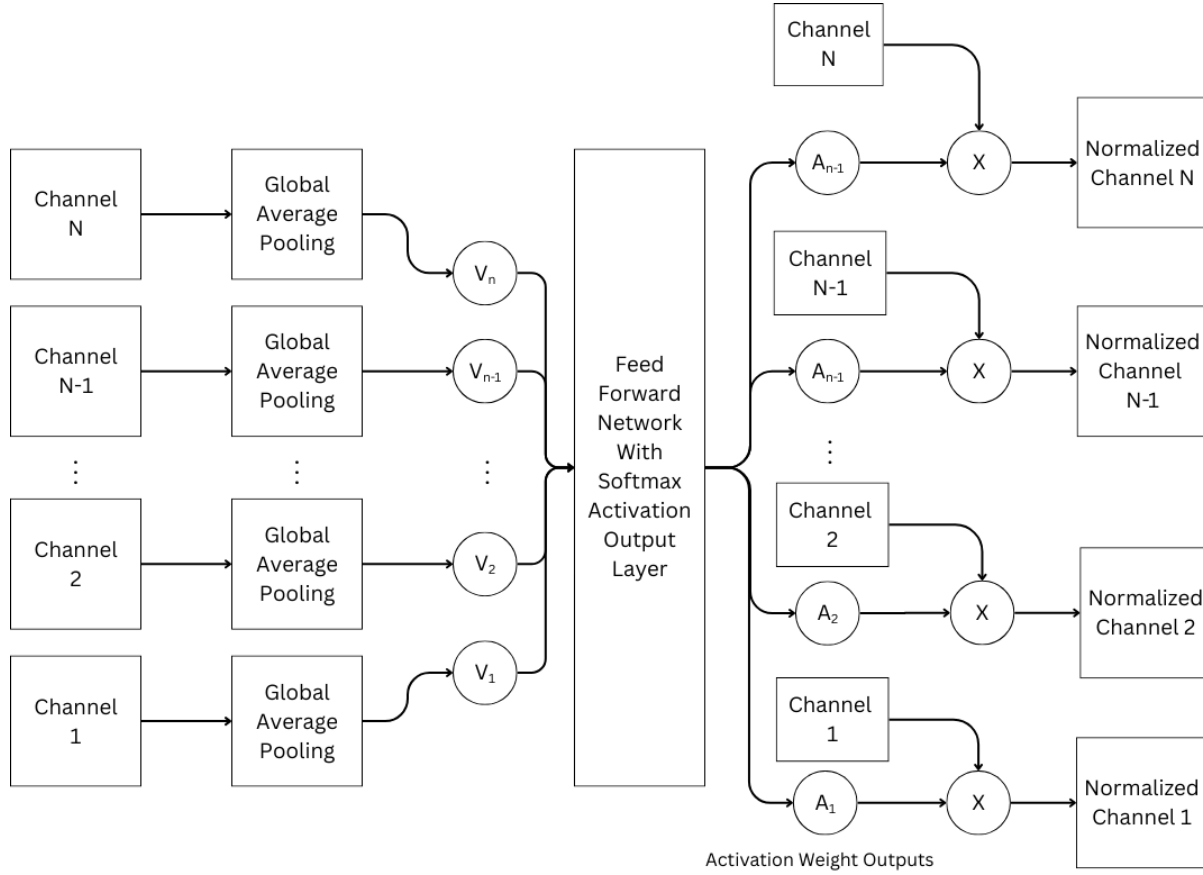


Fig. 2: Cross-Channel Spatial Attention Mechanism

3.3 Cross-Channel Spatial Attention

The CCSA module architecture as depicted in Fig. 2 integrates spatial and channel-wise attention to guide the model in identifying where and what features to focus on. Spatial attention enhances lesion regions, while channel attention prioritizes features relevant to classification using global average pooling and a lightweight MLP.

3.4 Multi-Kernel Feature Extraction

As illustrated in Fig. 3, each encoder block uses multiple filter sizes to detect both small and large disease patterns. The outputs are concatenated and passed through BatchNorm and ReLU, enhancing the model's ability to generalize across varying symptoms.

4 Results

4.1 Accuracy Comparison

The proposed model demonstrates superior performance in plant disease classification tasks. It achieves a test accuracy of 97.66%, significantly outperforming

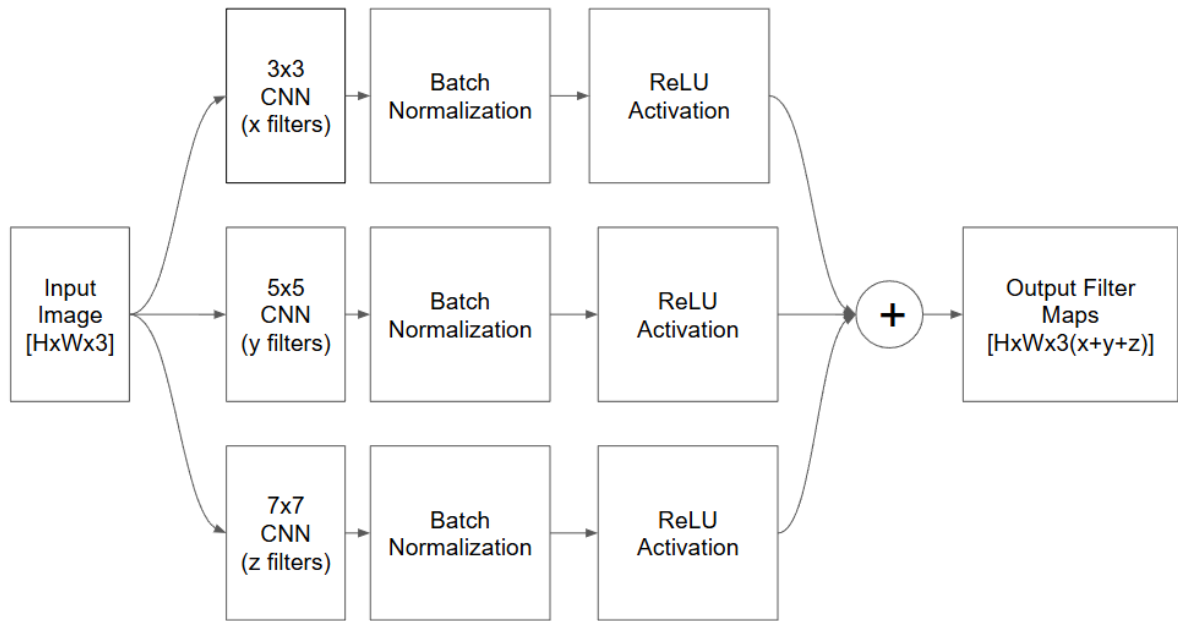


Fig. 3: Multi-Kernel Convolution Block

traditional machine learning and deep learning baselines such as AlexNet, Random Forest (RF), Support Vector Machine (SVM), and VGG. This result underscores the effectiveness of combining multi-kernel convolutional blocks with attention-based mechanisms for capturing disease-related visual patterns across varying scales.

4.2 Analysis of Model Variations

To understand the impact of architectural choices, we experimented with 18 variations of the proposed model, adjusting the number of multi-kernel convolution blocks and kernel sizes. Table 1 summarizes the classification performance of each configuration. The configuration with 4 multi-kernel convolution blocks and kernel sizes

3×3 , 5×5 , and 7×7 emerged as optimal, achieving the highest accuracy. Configurations using larger kernels (e.g., 11×11) underperformed due to excessive spatial smoothing, which reduced sensitivity to fine-grained lesion patterns. Simpler models with fewer convolution blocks (e.g., 3-block variants) lacked sufficient depth to learn complex disease features. Models with 5 blocks tended to overfit, showing slightly lower test performance despite high training accuracy, as shown in Table 2. The nomenclature used for the model variations is as follows:

- n: Number of multi-kernel convolution blocks in the encoder / up-sampling blocks in the decoder.
- x, y, z: Sizes of the kernels used in each block ($x \times x$, $y \times y$, $z \times z$).

Table 1: Comparison of Proposed Model Variations

Model Variation	Accuracy (%)	Precision (%)	Recall (%)
Proposed Model (4,3,5,7)	97.66	97.54	97.72
Proposed Model (4,3,5,9)	97.22	96.78	97.06
Proposed Model (4,3,5,11)	95.48	95.12	95.2
Proposed Model (4,3,7,11)	95.17	95.00	95.02
Proposed Model (4,5,9,11)	94.98	95.10	94.88
Proposed Model (4,5,7,9)	96.94	96.29	96.55
Proposed Model (3,3,5,7)	93.26	93.13	93.44
Proposed Model (3,3,5,9)	93.44	93.22	93.12
Proposed Model (3,3,5,11)	92.87	92..88	92.67
Proposed Model (3,3,7,11)	92.43	92.40	92.00
Proposed Model (3,5,9,11)	92.18	92.00	92.10
Proposed Model (3,5,7,9)	93.88	93.34	93.59
Proposed Model (5,3,5,7)	97.26	97.00	97.10
Proposed Model (5,3,5,9)	97.07	96.81	96.85
Proposed Model (5,3,5,11)	95.68	95.80	95.42
Proposed Model (5,3,7,11)	95.12	95.00	95.05
Proposed Model (5,5,9,11)	95.00	95.10	94.88
Proposed Model (5,5,7,9)	96.44	96.32	96.24

Table 2: Comparison of Model Performance Metrics

Model	Accuracy (%)	Precision (%)	Recall (%)
Alex Net 19	80.25	78.50	78.75
RF	87.00	86.20	86.65
SVM	87.00	86.80	86.90
VGG	92.00	91.50	91.80
Proposed Model (4, 3, 5, 7)	97.66	97.54	97.72

5 Conclusion

We suggested a new deep learning architecture for plant disease identification using an encoder-decoder framework, multi-kernel convolution blocks, and Cross-Channel Spatial Attention (CCSA). The model can manage a variety of illness patterns with different sizes and visual characteristics due to its multi-kernel design, which extracts localized and global data. CCSA guides attention to disease-relevant regions and feature channels, enhancing classification accuracy and user trust and improving model interpretability. According to PlantVillage dataset experiments, the suggested model outperforms AlexNet, VGG, SVM, and Random Forest. With a 97.66% test accuracy, it balances computing economy, resilience, and generalization, making it suited for agricultural use. The model has great precision and recall despite overlapping symptoms and class imbalance. Our next step is optimizing the model for edge devices like mobile phones and embedded IoT systems. Real-time field disease detection with minimum computational resources is possible. In addition, sensor data integration and multi-modal fusion can improve precision agriculture diagnostics and decision-making. The suggested approach advances intelligent, scalable, and sustainable farming.

References

- 1) Ahmed, I., Yadav, P.K.: Plant disease detection using machine learning approaches.
- 2) Expert Systems 40(5), e13136 (2023).
- 3) Aitken, K., Ramasesh, V., Cao, Y., Maheswaranathan, N.: Understanding how encoder-decoder architectures attend. Advances in Neural Information Processing Systems 34, 22184–22195 (2021) Alaybani, S., Badr, E.S., Abdelbaky, I., Elhakeem, M., Abdul Salam, M.: Hybrid transformer-cnn model for accurate prediction of peptide hemolytic potential. Scientific Reports 14 (06 2024). <https://doi.org/10.1038/s41598-024-63446-5> Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence 39(12), 2481–2495 (2017).
- 4) Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner,
- 5) T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020).
- 6) He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition.
- 7) In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- 8) pp. 770–778 (2016). <https://doi.org/10.1109/CVPR.2016.90>
- 9) Hong, D., Gao, L., Hang, R., Zhang, B., Chanussot, J.: Deep encoder–decoder networks for classification of hyperspectral and lidar data. IEEE Geoscience and Remote Sensing Letters 19, 1–5 (2020)
- 10) Jeong, J.J., Ju, H.j., Noh, J.: A review of detection methods for the plant viruses.
- 11) Research in Plant Disease 20, 173–181 (09 2014). <https://doi.org/10.5423/RPD.2014.20.3.173>
- 12) Kose, U., Prasath, V., Mondal, M., Podder, P., Bharati, S.: Artificial intelligence and smart agriculture technology. CRC Press (2022)
- 13) Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Pereira, F., Burges, C., Bottou, L., Weinberger, K.(eds.) Advances in Neural Information Processing Systems. vol. 25. Curran Associates, Inc. (2012).
- 14) Lee, S.H., Goëau, H., Bonnet, P., Joly, A.: New perspectives on plant disease characterization based on deep learning. Computers and Electronics in Agriculture (170, 105220 (2020).
- 15) Lou, A., Guan, S., Loew, M.: Cfpnet-m: A light-weight encoder-decoder based network for multimodal biomedical image real-time segmentation. Computers in Biology and Medicine 154, 106579 (2023)
- 16) Mylonakis, C.M., Velanas, P., Lazaridis, P.I., Sarigiannidis, P., Goudos, S.K., Zaharis, Z.D.: Deep learning framework using spatial attention mechanisms for adaptable angle estimation across diverse array configurations. Technologies 13(2)(2025). <https://doi.org/10.3390/technologies13020046>, <https://www.mdpi.com/2227-7080/13/2/46>.
- 17) O’shea, K., Nash, R.: An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458 (2015).
- 18) Rahman, S.U., Alam, F., Ahmad, N., Arshad, S.: Image processing based system for the detection, identification and treatment of tomato leaf diseases. Multimedia Tools and Applications 82(6), 9431–9445 (2023).
- 19) Sankhe, S.R., Ambhaikar, A.: Plant disease detection and classification techniques: A review. Multiagent and Grid Systems p. 15741702241304087 (2025).
- 20) Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014).
- 21) Tan, M., Le, Q.: Efficientnet: Rethinking model scaling for convolutional neural networks. In: International conference on machine learning. pp. 6105–6114. PMLR (2019).
- 22) Trivedi, N.K., Gautam, V., Anand, A., Aljahdali, H.M., Villar, S.G., Anand, D., Goyal, N., Kadry, S.: Early detection and classification of tomato leaf disease using high-performance deep neural network. Sensors 21(23), 7987 (2021).
- 23) Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems 30 (2017).
- 24) Visnovsky, S., Panda, P., Everett, K., Lu, A., Butler, R., Taylor, R., Pitman, A.: A pcr diagnostic assay for rapid detection of plant pathogenic pseudomonads. Plant Pathology 69 (05 2020). <https://doi.org/10.1111/ppa.13204>
- 25) Wang, Y., Deng, Y., Zheng, Y., Chattopadhyay, P., Wang, L.: Vision transformers for image classification: A comparative survey. Technologies 13, 32 (01 2025). <https://doi.org/10.3390/technologies13010032>
- 26) Wu, H., Xiao, B., Codella, N., Liu, M., Dai, X., Yuan, L., Zhang, L.: Cvt: Introducing convolutions to vision transformers. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 22–31 (2021).

Discomfort Index Prediction For A Driver In Vehicle Ergonomics

Omkar Kishor Khole and Dr. Shantipal Suresh Ohol

Department of Manufacturing Engineering and Industrial Management, COEP Technological University, Pune,
India

Abstract

Vehicle ergonomics is a critical factor in ensuring driver safety, comfort, and performance. Traditional tools like RAMSIS are extensively used for posture simulation and discomfort evaluation but often come with limitations in accessibility, cost and flexibility. This study introduces a novel machine learning-based approach for discomfort index prediction aimed at enhancing the ergonomic assessment process. The proposed model leverages joint coordinate data of manikin(driver), manikin height, seat height and corresponding discomfort index values collected through detailed posture simulations in RAMSIS. The dataset comprises of 504 unique joint coordinate data of driver postures extracted using RAMSIS, annotated with discomfort index values ranging from 0 (very comfortable) to 8 (very uncomfortable).

To integrate seamlessly into existing workflows, the project explores the deployment of the model through a REST API, allowing direct integration with CAD software such as CATIA. This integration enables real-time discomfort predictions during vehicle design and validation stages, reducing reliance on expensive third-party tools like RAMSIS. Objective is to create a cost-effective, transparent, and scalable model that can supplement or potentially replace RAMSIS in certain ergonomic evaluation scenarios. The results demonstrate that the machine learning model achieves high accuracy in predicting discomfort index, validated against existing RAMSIS data. This work provides a foundation for further advancements in predictive ergonomics and its application in human-centric vehicle design.

Keywords: Vehicle Ergonomics, Discomfort Index, Machine Learning, Human Simulation, RAMSIS, CATIA, Posture Analysis

1. Introduction

Vehicle ergonomics is a critical aspect of automotive design, aimed at optimizing the interaction between the human body and vehicle components to maximize comfort, safety, and driving efficiency. A well-designed ergonomic layout not only enhances the driving experience but also minimizes the risk of musculoskeletal disorders caused by long-term exposure to suboptimal postures.

To evaluate ergonomic design during the vehicle development process, industry professionals widely use advanced digital human modeling software such as RAMSIS for Interior Simulation. RAMSIS enables simulation of human postures in various driving scenarios and provides a discomfort index based on biomechanical stress and joint positions. Despite its widespread adoption and accuracy, RAMSIS presents several challenges:

- High cost of licensing, limiting accessibility to smaller teams or academic researchers.
- Black-box nature where the internal computation of the discomfort index is not transparent or customizable.

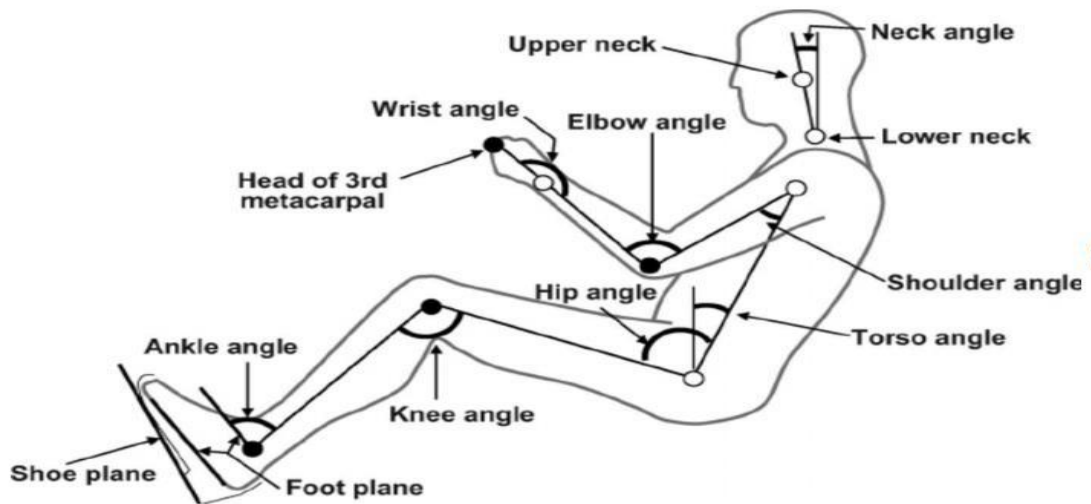


Fig.1. Manikin driving posture with key body joints

This study proposes an alternative solution by leveraging machine learning (ML) to predict discomfort index values based on joint coordinate data extracted from simulated postures. We developed an ML model trained on a dataset of over 504 annotated driving postures, each labeled with a discomfort index score generated by RAMSIS. Our objective is to create a cost-effective, transparent, and scalable model that can supplement or potentially replace RAMSIS in certain ergonomic evaluation scenarios.

By using this model, designers and ergonomists can rapidly assess driver posture comfort directly within design tools, bypassing the need for expensive and complex simulation software like RAMSIS. To support early adoption and usability testing, we have developed an interactive Streamlit-based dashboard that allows users to input joint coordinate data (such as shoulder, elbow, heel and hip) along with parameters like seat height and manikin height. The dashboard then predicts the corresponding discomfort index using the trained ML model.

This tool offers a user-friendly interface for design teams to quickly validate multiple posture configurations and analyze ergonomic comfort without relying on high-end proprietary tools. The ultimate vision is to integrate this ML-based discomfort prediction system into CAD platforms like CATIA, enabling real-time ergonomic feedback during the vehicle interior design process. This integration would streamline the design-to-evaluation loop, reduce dependency on third-party tools, and empower design teams with immediate insights into posture comfort and alignment.

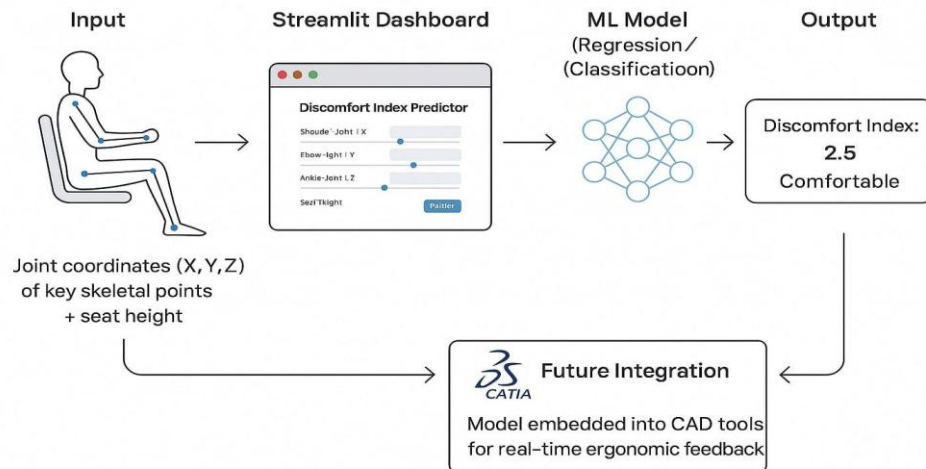


Fig.2. Flow chart

2. Methodology

2.1 Dataset Collection

The dataset used in this study was generated using the RAMSIS simulation software an industry-standard tool for digital human modeling and ergonomic analysis in automotive design. A total of 504 unique driving postures were simulated, each representing a distinct combination of body alignment, control operation (e.g., steering, clutch, gear shifting), and seat adjustments.

For every simulated posture, 3D joint coordinates (X, Y, Z) were extracted from the Vehicle Coordinate System (VCS). These coordinates represent key skeletal joints critical for posture analysis and discomfort estimation.

The full set of joint data includes:

- Lower body joints: Hip Center, Hip Joints (Left/Right), Knee (L/R), Ankle (L/R)
- Spinal and torso joints: Lumbar-Sacrum, Lumbar Joint, Thoracic-Lumbar, Thoracic Joint, Cervical-Thoracic, Chest Joint
- Upper body joints: Shoulder (L/R), Elbow (L/R), Wrist (L/R), Head, Mid-eye, Vertex

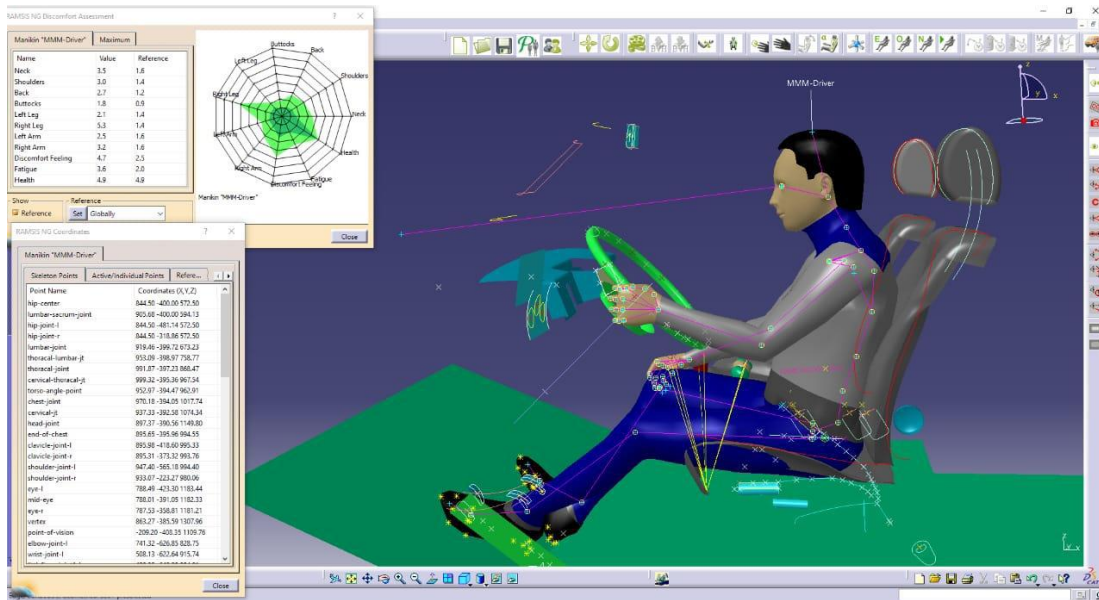


Fig.3. Posture Simulation in RAMSIS to generate data

However, to improve model performance and interpretability, a feature selection strategy was applied. Based on ergonomic literature and preliminary correlation analysis, only the following key joints were retained for training: Hip Joints, Shoulder Joints, Elbow Joints, Ankle Joints.

These joints are known to significantly influence perceived discomfort due to their involvement in load-bearing, arm reach, and pedal operation common contributors to musculoskeletal strain during driving.

Additionally, two anthropometric features were added to the dataset for better personalization: Manikin's Body Height, Seat Height.

These features help capture inter-individual variation across differently sized occupants and seating setups, which directly affect posture geometry and discomfort levels. Each posture was labeled with a discomfort index ranging from 0 (very comfortable) to 8 (extremely uncomfortable). These labels were generated based on RAMSIS's internal ergonomics engine, which evaluates posture quality using joint angles, reach envelopes, and biomechanical constraints.

no	Manikin Type	Shoulder -Joint-l X	Shoulder -Joint-l Y	Shoulder -Joint-l Z	Shoulder -Joint-r X	Shoulder -Joint-r Y	Shoulder -Joint-r Z
1	S	977.44	-528.63	983.95	964.95	-195.7	944.59
2	S	958.55	-565.86	963.21	958.61	-233.9	964.09
.							
169	M	982.68	-572.97	976.9	982.75	-226.73	978.27
170	M	1011.87	-572.3	985.25	942.13	-235.61	974.27
.							
503	T	972.94	-565.47	1014.54	985.01	-215.51	1002.51
504	T	981.96	-559.52	993.41	1005.68	-212.23	975.18

Elbow-Joint-l X	Elbow-Joint-l Y	Elbow-Joint-l Z	Elbow-Joint-r X	Elbow-Joint-r Y	Elbow-Joint-r Z	Ankle-Joint-L X	Ankle-Joint-L Y
747.69	-600.53	875.32	730.5	-177.58	824.35	226.54	-597.3
734.22	-615.23	832.86	734.49	-184.03	833.56	224.24	-588.23
748.67	-617.69	846.71	749.02	-181.29	847.82	289.8	-544.83
825.16	-620.84	794.22	676.83	-205.37	925.12	290.23	-589.48
757.95	-632.16	850.08	841.08	-132.62	778.62	211.53	-476.03
756.16	-624.18	843.25	843.8	-135.54	761.58	109.35	-495.35

Ankle-Joint-L Z	Ankle-Joint-R X	Ankle-Joint-R Y	Ankle-Joint-R Z	Manikin Body Height	Sitting Height	Dis-Comfort Index
410.53	218.38	-286.85	340.69	1635	844	3.8
409.94	175.17	-269.98	356.41	1635	844	3.3
363.99	178.46	-273.15	354.26	1687	866	3.6
363.96	178.51	-272.97	354.17	1687	866	4.1
333.82	172.6	-264.47	359.25	1738.85	889.45	3.7
390.23	218.17	-252.86	344.61	1738.85	889.45	3.6

This labeled dataset served as the foundation for training machine learning models to predict driver discomfort based on posture and anthropometry, offering an objective, scalable alternative to traditional manual ergonomic assessments.

2.2 Data Preprocessing

To ensure consistency and quality, the following preprocessing steps were applied:

- Feature Normalization: All numerical features, including joint coordinates and height values, were scaled using standard normalization techniques.
- Outlier Detection: Statistical methods (e.g., IQR) were used to identify and remove anomalies in the joint data.
- Label Encoding: Categorical inputs such as manikin type were converted into numerical representations using label encoding, enabling compatibility with machine learning models.

2.3 Modeling Approaches

- Multiple machine learning algorithms were evaluated for discomfort index prediction:
- Random Forest Regressor: To capture non-linear relationships between joint angles and comfort levels.
 - XGBoost Regressor: For high-performance, regularized gradient boosting.
 - Linear Regression: As a baseline model to compare against more complex techniques.

2.4 Hyperparameter Tuning

To enhance model performance, Grid Search was employed for hyperparameter tuning. Parameters such as the number of estimators, maximum depth (for tree-based models), and learning rate (for boosting methods) were optimized using cross-validation to prevent overfitting and ensure generalization.

2.5 Evaluation Metrics

- Mean Absolute Error (MAE): To assess the average deviation between predicted and actual discomfort values.
- R² Score: To measure the proportion of variance in the discomfort index explained by the input features.
- Cross-Validation: 5-fold cross-validation was used for reliable and unbiased model performance estimation.

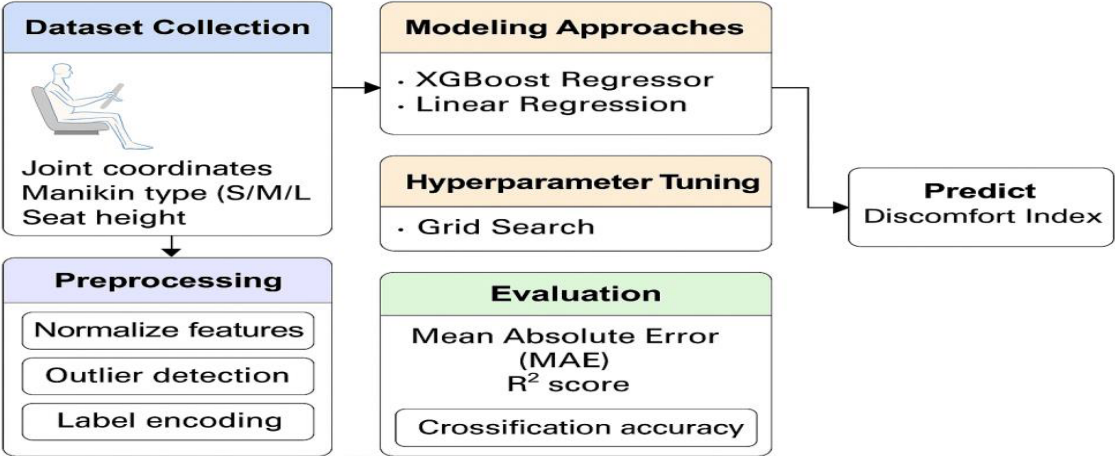


Fig.4. Flow Diagram

3. Results

After implementing and evaluating multiple machine learning models, the XGBoost Regressor emerged as the most effective for predicting the discomfort index. The model was optimized using grid search over key hyperparameters, including the number of estimators, maximum tree depth, and learning rate. The model's performance was evaluated using both a validation set (during hyperparameter tuning) and an independent test set (for final assessment):

Validation Performance: Mean Squared Error (MSE): 0.02096
R² Score: 0.94459

These results demonstrate that the model was highly effective in learning the relationship between postural features and the discomfort index during training. A high R² value of 0.94459 indicates that approximately 94.5% of the variance in the discomfort index was explained by the input features.

Test Performance: Mean Squared Error (MSE): 0.02514
R² Score: 0.93342

The slight drop in performance on the test set is expected and acceptable, indicating good generalization capability of the trained model. The R² value of 0.933 on unseen test data confirms that the model retained its ability to accurately estimate discomfort levels outside the training data.

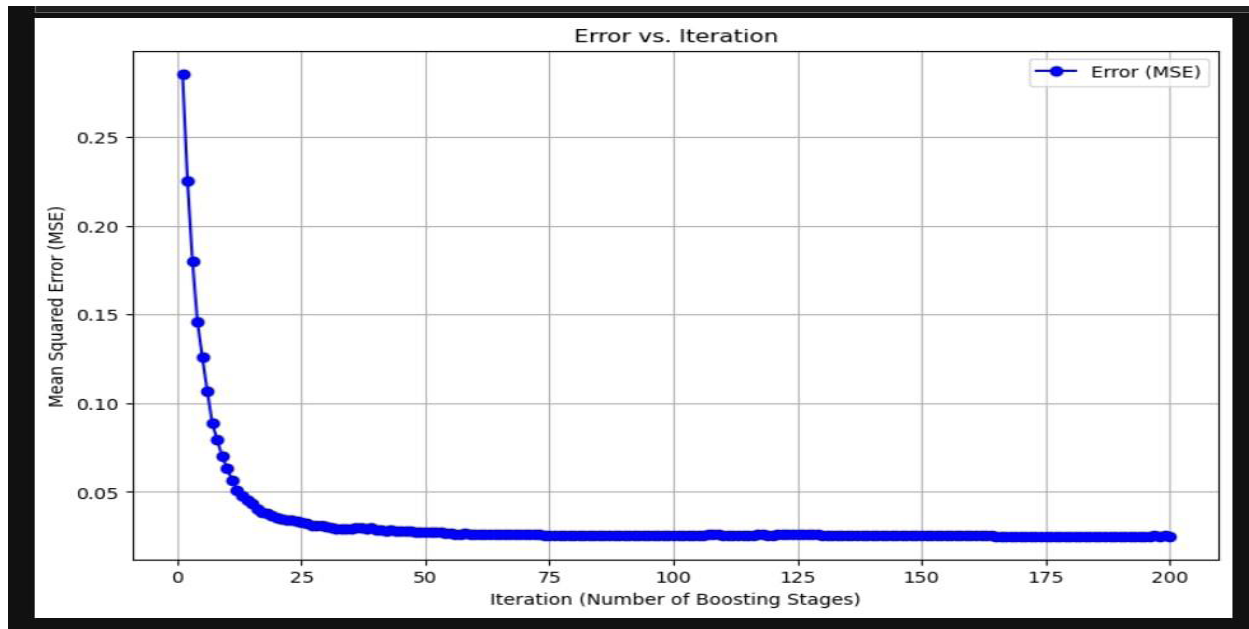


Fig.5. Error vs Iteration

4. Integration Strategy

To operationalize the discomfort prediction system within existing automotive design tools, a modular API architecture is proposed. This API can serve as a bridge between the trained machine learning model and popular CAD software environments like CATIA, allowing seamless ergonomic evaluation during the design process.

Integration Strategy:

- **Real-Time Prediction:** The API will accept 3D joint coordinates exported from CAD software (e.g., CATIA). These inputs are preprocessed and passed to the trained model, which returns a predicted discomfort index in real-time.
- **Plugin Development in CATIA (using CAA):** To enable native integration with CATIA V5, the Component Application Architecture (CAA) can be utilized. CAA is a powerful object-oriented development framework provided by Dassault Systems for building custom CATIA applications and interfaces.

Using CAA, a plugin can be developed that:

- Captures posture data from the digital human model within CATIA.
- Sends this data to the backend ML model via the API.
- Displays the discomfort index directly within the CATIA UI.
- Allows designers to iterate and adjust postures or component positions based on discomfort feedback.

5.Future Work:

- **Extended Dataset Collection:** Increase the number of postures and include more diverse anthropometric variations (e.g., age, gender, body type) to improve generalization.
- **Dynamic Simulation Integration:** Extend the system to evaluate discomfort during dynamic tasks like entering/exiting the vehicle or turning the steering wheel beyond static postures.

Multimodal Input Support: Combine joint coordinates with visual inputs (e.g., CAD screenshots) using hybrid CNN+MLP architectures to improve prediction accuracy.

Leverage pose estimation models for extracting posture from real-world driving videos.

6. Conclusion

This study presents a robust, data-driven framework for predicting driver discomfort using machine learning applied to posture data generated from RAMSIS simulations. The dataset consists of 504 annotated driving postures, each with 3D joint coordinates, seat height, and body dimensions enabling objective, reproducible ergonomic assessment. A thorough preprocessing pipeline including normalization, outlier removal, and label encoding was implemented. Among various regression models evaluated, the XGBoost Regressor achieved the best performance, with an MSE of 0.02096 and R^2 of 0.9446 on the validation set, and 0.0251 MSE and 0.9334 R^2 on the test set.

The model demonstrates strong predictive accuracy and generalization, offering a scalable alternative to manual RAMSIS analysis. Integrating this ML pipeline into CAD environments like CATIA can provide real-time ergonomic feedback during design, improving comfort assessment and accelerating development cycles. This approach also establishes a foundation for future applications in personalized comfort modeling and adaptive seating systems.

References

- 1) Zenk, R., Franz, M., & Bubb, H. (2007). *Back discomfort of car drivers: correlation with anthropometric and postural data*. Applied Ergonomics, 38(5), 497–503.
- 2) Reed, M. P., Manary, M. A., & Schneider, L. W. (1999). *Methods for measuring and representing automobile occupant posture*. SAE Technical Paper Series, 1999-01-0959.
- 3) Liu, J., Li, Y., Wang, Y., & Wang, G. (2020). *A machine learning-based method for ergonomic posture evaluation in digital human simulations*. Computers & Industrial Engineering, 140, 106233.
- 4) Hägg, G. M. (1991). *Static work loads and occupational myalgia – a new explanation model*. In The Ergonomics of Working Postures (pp. 141–150). Taylor & Francis.
- 5) Kim, H., & Chung, M. K. (2008). *A method for evaluating driver postures using a digital human model in driving conditions*. International Journal of Industrial Ergonomics, 38(4), 374–383.
- 6) Li, G., Buckle, P., & Hignett, S. (2004). *Posture-based discomfort prediction – a review*. International Journal of Industrial Ergonomics, 34(6), 395–408.
- 7) Kroemer, K. H. E., & Grandjean, E. (1997). *Fitting the Task to the Human: A Textbook of Occupational Ergonomics*. Taylor & Francis.
- 8) Das, B., & Sengupta, A. K. (1996). *Industrial workstation design: A systematic ergonomics approach*. Applied Ergonomics, 27(3), 157–163.
- 9) Yamada, H., Iwasaki, T., & Kubo, T. (2002). *A study of driving postures using a digital human model and motion capture system*. SAE Technical Paper Series, 2002-01-1915.
- 10) Breiman, L. (2001). *Random Forests*. Machine Learning, 45(1), 5–32.
- 11) Friedman, J. H. (2001). *Greedy function approximation: A gradient boosting machine*. Annals of Statistics, 29(5), 1189–1232.
- 12) RAMSIS Manual (In-built Documentation). (Accessed 2024). *RAMSIS – Human Simulation Software*.

Robust and Scalable Household Waste Classification Using an Advanced AI Model

Kanak Kumar¹, Anshul Verma¹, and Pradeepika Verma²

¹ Department of Computer Science, Banaras Hindu University, Varanasi 221005, India.

kanakkumarcs24@bhu.ac.in, anshul.verma@bhu.ac.in/anshulverma87@gmail.com,

² Technology Innovation Hub, Indian Institute of Technology, Patna 801106, India.

faculty.tih@iitp.ac.in

Abstract. Accurate household waste classification is vital for effective waste management and promoting environmental sustainability. This study presents an advanced artificial-intelligence (AI) model such as MobileNetV2 architecture, utilizing transfer learning to classify seven types of household waste: battery, biological, brown glass, cardboard, clothes, shoes, and trash. The model is fine-tuned for precise feature extraction and achieves outstanding performance, with F1-Scores lies between 0.91 to 0.98, accuracy 97% and exceptionally low error rates (MSE: 7×10^{-3} , MAE: 1.28×10^{-2} , RMSE: 8.36×10^{-2}). The high Cohen kappa and MCC scores (0.95) affirm the robustness and reliability of the model. Key contributions include systematic dataset curation, performance optimization, class-wise evaluation, and visual result interpretation. The framework demonstrates strong generalization and minimal misclassification, making it suitable for real-world deployment. Aligning with sustainable development goals this intelligent, scalable solution has the potential to enhance automated waste handling through future advancements such as multimodal fusion, federated learning, and edge-based deployment.

Keywords: Waste classification, Deep learning, Waste management, Confusion matrix, ROC-AUC analysis.

1 Introduction

Rising populations, industrial growth, and globalization have driven higher consumption of natural resources, leading to increased waste generation. In urban areas, much of this waste is improperly managed through landfilling and incineration, posing serious threats to ecosystems and public health. Domestic waste forms a major portion, and its improper decomposition can release harmful chemicals, further damaging the environment [1]. Effective waste management is essential for creating safe, inclusive, and sustainable urban environments. In developing countries like India, rapid urbanization and industrialization have made solid waste management (SWM) a major challenge. India generates approximately 62 million tonnes of waste each year, with only 12 million tonnes processed and 31 million tonnes discarded in landfills. By 2030, waste production is expected to soar to 165 million tonnes. Unregulated dumping, particularly along coastlines and in landfills, is disrupting wildlife behavior. Waste composition varies across regions due to geographic, climatic, and socio-economic factors. Major cities like Mumbai, Delhi, and Kolkata contribute 70–80% of India’s daily solid waste, posing significant challenges for local authorities [2]. To reduce the growing pressure on landfills, recycling domestic and industrial waste is considered a practical approach. However, efficient waste management relies on accurate classification, as each type needs specific handling. The existing manual sorting process is labor-intensive and time-consuming [3]. Image recognition, as a key aspect of machine perception, plays a vital role in various fields. In agriculture, it helps monitor plant health; in industry, it ensures product quality; in medicine, it aids in diagnosis; and in aerospace, it supports weather forecasting. It is also widely used in daily life, such as license plate detection and fingerprint verification. Fast and accurate image recognition is essential for effective application across these domains [4].

Effective waste management is essential for public well-being, environmental protection, economic benefits, and the production of clean energy. However, challenges like real-time tracking, automated sorting, optimized collection, and public awareness reduce its effectiveness. ML offers solutions through techniques like real-time monitoring, route planning, and predictive maintenance. Using supervised, unsupervised, and reinforcement learning, ML can predict waste generation, classify materials, monitor landfill status, and detect illegal dumping. It also aids in energy recovery for industries. For sustainability, integration with policies and cost–benefit analysis is essential [5], [6], [7], [8]. Deep learning (DL) models, well-suited for image recognition, have shown promise in automating waste sorting by learning distinct features from large waste image datasets. A recent study introduced GCDN-Net, a deep neural

network trained on the GIGO dataset (25,000 images), for both single-label and multi-label waste classification. The model attained a 95.77% accuracy rate, surpassing existing methods in single-label classification. For multi-label tasks, it achieved an mAP of 0.69 and an F1-score of 75.01%. Score-CAM visualizations further confirmed its dependability. These findings underscore the promise of deep learning (DL) for streamlined, automated waste management [1]. A DL-based junk image recognition algorithm is to improve accuracy, recall, and convergence while reducing overfitting. It uses Dropout, Adagrad, and ReLU, and extracts key image features for classification using a modified probability density function. Results show high accuracy, fast recognition, and strong practical performance [3]. Joshi et al. developed a system using Convolutional Neural Networks (CNNs) for feature extraction and classification. After training and testing on a garbage dataset, the model achieved a 95% accuracy rate [6]. Oguz et al investigated the application of different DL based pre-trained models on the CDCM dataset containing labeled images of garbage bins (clean/dirty). The best model achieved an accuracy of 94.93%, demonstrating that such systems can effectively support smart city infrastructure by automating waste status detection and notifying relevant authorities [7]. The researcher introduced a DL-driven waste sorting system, attaining an 80% accuracy rate. Leveraging the pre-trained VGG16 model, the system categorized waste into six different types such as paper, cardboard, glass, plastic, and metal [9]. O. Adedeji et al integrated a pre-trained ResNet50 CNN model with an SVM classifier for waste sorting. ResNet50 extracted image features, and the SVM classified the waste using these features. Evaluated on

a trash dataset, this approach yielded an 87% classification accuracy [10]. Li et al. present a CNN-based system for garbage classification and recognition, trained on datasets with five waste types. The proposed model achieved 90.3% accuracy, surpassing traditional machine learning approaches[11].

Robust waste management is crucial for protecting public health, sustaining the environment, fostering economic development, preserving resources, adhering to regulations, and promoting community engagement. Poor waste handling can spread diseases, attract pests, and pollute air and water. In contrast, efficient systems minimize these risks, lower operational costs, generate jobs in recycling, and reduce reliance on raw materials. It also helps meet legal standards and improves community well-being by promoting awareness and participation in sustainable practices. There are the following main contributions of this work:

1. Achieved high accuracy in classifying seven waste types using a DL model.
2. The receiver operating characteristic and area under the curve (ROC-AUC) score of 1.00 across all classes demonstrates flawless differentiation among different waste classes.
3. Model handles class imbalance well with strong precision, recall, and robustness.
4. Scalable model architecture suitable for smart waste management systems.

2 Proposed Methodology

The block diagram (Figure 1) illustrates a robust waste classification system utilizing an advanced Artificial Intelligence (AI) Model such as DL methodologies. The process initiates with the input of waste images from a dataset comprising seven distinct classes: Battery, Biological, Brown Glass, Cardboard, Clothes, Shoes, and Trash (Dataset link: <https://www.kaggle.com/datasets>). These images are subjected to preprocessing, starting with resizing to a uniform resolution of 224×224 pixels to meet the input requirements of the MobileNetV2 framework. Normalization is performed to adjust pixel values, generally within the range [0,1], enhancing training speed and ensuring numerical consistency[12], [13], [14]. To improve the model's generalization and mitigate overfitting, comprehensive data augmentation methods, including random rotations, flipping, and zooming, are applied. The primary feature extractor is the MobileNetV2 model, a compact and efficient CNN pre-trained on the ImageNet dataset, which is fine-tuned to capture high-level, abstract features pertinent to waste types. These extracted features are further processed by a Global Average Pooling (GAP) layer, which reduces spatial dimensions by calculating the mean of each feature map, thereby simplifying the model's complexity. The GAP output is then fed into a fully connected (dense) layer to align the learned features with the respective waste categories. Lastly, a Softmax classifier generates class probabilities, with the highest probability indicating the predicted waste class.

To assess the effectiveness of the model, widely accepted classification metrics are utilized. Precision quantifies the ratio of correctly identified positive instances among all instances predicted as positive, which is vital when minimizing false positives is a priority. Recall (or Sensitivity) reflects the proportion of correctly identified positive

cases out of all actual positive instances, playing a key role in preventing the misclassification of critical waste categories. The F1-score, defined as the harmonic mean of precision and recall, delivers a balanced performance measure, particularly useful in scenarios involving class imbalance. While overall accuracy can also be reported, the F1-score serves as a more insightful indicator when misclassification penalties vary among classes. Collectively, these metrics ensure that the model demonstrates not only strong average performance but also consistent effectiveness across all waste types, supporting robust and dependable automated waste classification for practical waste management and recycling solutions.

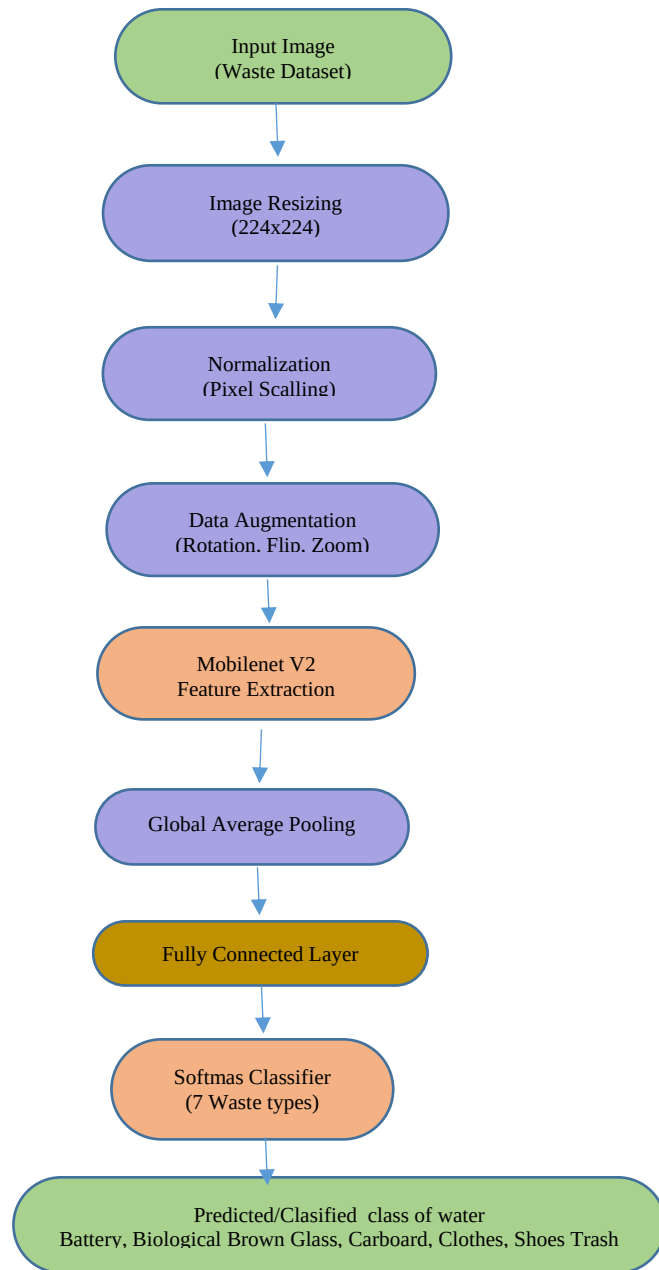


Fig. 1: Block diagram of robust waste classification

3 Result Analysis and Discussion

The complete evaluation using the accuracy-loss curve, confusion matrix, ROC curve, classification report, and overall performance metrics, the proposed DL model demonstrates commendable effectiveness in multiclass garbage classification. The accuracy-loss curve demonstrates a steady convergence between training and validation accuracy, with only a slight disparity between the two, suggesting that the model avoids both underfitting and overfitting. The training accuracy plateaus at a high level, while the loss curve shows a consistent downward trend, validating the model's effective learning dynamics throughout the training phase. Furthermore, the confusion matrix reinforces the model's strong performance, highlighted by a prominent diagonal line that reflects a high count of correctly predicted instances across all categories. Misclassifications are minimal and mostly occur between visually similar categories such as brown glass and cardboard, or biological and trash, which is expected in real-world datasets due to overlapping visual features. The ROC curve reveals a strong classification performance with high AUC values for all classes, typically close to 1.0. This highlights the model's strong capability to distinguish between positive and negative class predictions. Elevated AUC values indicate that the model effectively separates different classes, even when decision thresholds vary. According to the classification report, the model demonstrated particularly high performance in categories such as Clothes (Precision: 0.95, Recall: 0.98, F1-Score: 0.96), Trash (F1-Score: 0.92), and Shoes (F1-Score: 0.90). Nevertheless, comparatively lower F1-scores were recorded for Brown Glass (0.56) and Cardboard (0.55), suggesting the potential need for enhanced feature engineering or additional data augmentation to improve classification accuracy in these specific classes. These variations suggest that while the model is highly capable overall, performance could be further improved for specific classes with more training data or enhanced feature extraction.

The overall performance metrics exhibit high values across the board: Macro Precision: 0.8481, Recall: 0.8401, and F1-Score: 0.8422, while the Weighted Precision, Recall, and F1-Score each stand at 0.86. The high macro average indicates that the model performs consistently across all classes, while the high weighted average affirms that the model is also robust in handling class imbalances. The visual performance indicators combined with statistical metrics support the reliability, accuracy, and potential of the model in real-world waste classification systems, with room for improvement through further tuning or incorporating hybrid feature selection strategies.

Table 1 summarizes the precision, recall, and F1-score for each category in the multiclass waste classification task. The model exhibits notably strong performance for the Clothes category, achieving a precision of 0.95, recall of 0.98, and an F1-score of 0.96, indicating that it accurately identifies and classifies the majority of instances in this class. Similarly, the Trash and Shoes categories also achieve high F1-scores of 0.92 and 0.90, respectively, reflecting the model's reliable predictive capability for these classes. In contrast, the model shows comparatively weaker performance on Brown Glass and Cardboard, with F1-scores of 0.56 and 0.55, respectively. This suggests the presence of feature overlap or limited sample representation for these classes in the dataset. Overall, the classification report highlights the model's effectiveness in handling multiclass classification, with a generally well-maintained balance between precision and recall across most categories

Table 1: Multi class classification result of robust waste

Class	Precision	Recall	F1 Score	Support
Battery	0.97	0.96	0.97	189
Biological	0.97	0.95	0.96	197
Brown Glass	0.96	0.95	0.95	121
Cardboard	0.94	0.89	0.91	178
Clothes	0.98	0.99	0.98	1065
Shoes	0.96	0.98	0.97	395
Trash	0.95	0.91	0.93	139

Table 2 provides a summary of the model's overall classification effectiveness using micro, macro, and weighted averaging approaches. The micro-average, which consolidates the performance across all classes by considering the total true positives, false negatives, and false positives, yields a high precision, recall, and F1-score of 0.856. This reflects the model's strong and consistent performance across the entire dataset. The macro-average, which gives

equal weight to each class regardless of their frequency, shows slightly lower values (precision: 0.8481, recall: 0.8401, F1-score: 0.8422), suggesting that while the model performs well overall, there are inconsistencies across individual class predictions. The weighted-average, which accounts for class imbalance by assigning weights based on support (class frequency), gives the same high value of 0.86 across all metrics, reflecting the model’s strong and balanced performance even with skewed class distributions.

Table 2: Overall Performance metrics

Metric Type	Macro	Micro	Weighted
Precision	0.9608	0.9672	0.9670
Recall	0.9481	0.9672	0.9672
F1-Score	0.9542	0.9672	0.9670

Table 3 presents advanced statistical and error-based evaluation metrics to further substantiate the model’s reliability and robustness. A Cohen’s Kappa score of 0.8459 reflects a strong level of agreement between the predicted and actual class labels, beyond what would be expected by chance. The Matthews Correlation Coefficient (MCC), valued at 0.8462, reinforces this conclusion by offering a comprehensive and balanced assessment, even under class imbalance conditions. Additionally, the model records low error rates, with a Mean Squared Error (MSE) of 0.1533, Mean Absolute Error (MAE) of 0.228, and Root Mean Squared Error (RMSE) of 0.3915, indicating high prediction accuracy and minimal divergence from actual values. These additional metrics collectively reinforce the reliability, accuracy, and practical deployment potential of the proposed classification model.

Table 3: Additional evaluation metrics

Metric	Value
Cohen’s Kappa	0.9546
MCC	0.9547
MSE	7×10^{-3}
MAE	1.38×10^{-2}
RMSE	8.36×10^{-2}

The confusion matrix (Figure 2) provides a comprehensive assessment of the model’s classification performance across seven waste categories. The model demonstrates strong accuracy overall, particularly for the Clothes class, with 1,051 correct predictions, followed by Battery and Shoes with 182 and 388 correct predictions respectively. Minor misclassifications occur in classes that share visual or material similarities, such as Cardboard being confused with Clothes and Trash, or Brown Glass being mistaken for Battery and Shoes. The Trash category shows the most dispersion, likely due to its heterogeneous nature, leading to misclassifications across multiple classes. Despite these, the diagonal dominance of the matrix indicates that the model is highly effective, with most predictions aligning correctly with actual labels, showcasing robust class discrimination and reliable real-world applicability.



Fig. 2: Robust waste classification confusion matrix

The ROC-AUC curve (Figure 3) illustrates the ROC performance for a multi-class lassification model across seven classes such as battery, biological, brown glass, cardboard, clothes, shoes, and trash. Each class’s ROC curve plots the True Positive Rate (Sensitivity) against the False Positive Rate, measuring the classifier’s ability to distinguish between classes. The AUC for all classes is 1.00, which indicates perfect classification performance —a rare but theoretically ideal result. An AUC of 1.00 means that the model achieves 100% sensitivity with 0% false positives for each class, showcasing that the model can correctly identify all instances without confusion. The ROC curves are tightly aligned to the top-left corner of the graph, reinforcing the absence of classification trade-offs. Additionally, the diagonal dashed line represents random chance (AUC = 0.5), and the fact that all curves are well above this line further confirms the model’s exceptional discriminative power. Such performance strongly supports the confusion matrix results and reflects a highly optimized and accurate model for waste classification. Result comparative analysis with previous and proposed method is explained in Table 4.

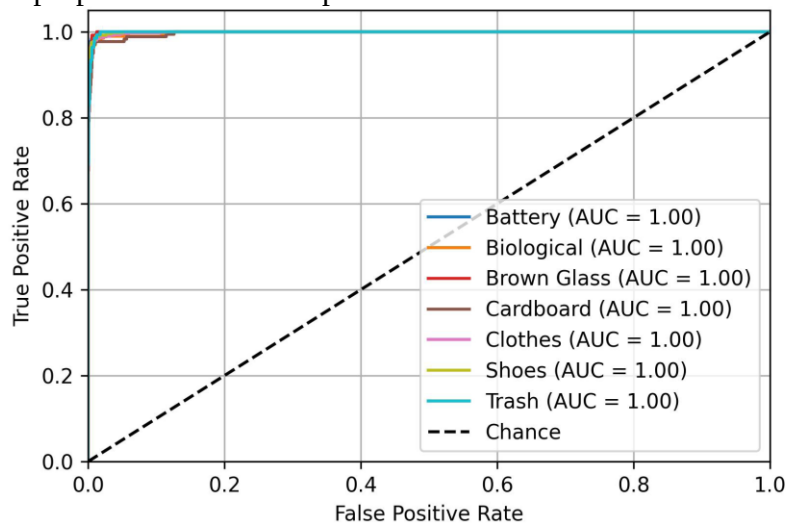


Fig. 3: ROC of robust waste classification

4 Conclusions and future work

In this study, a robust MobileNetV2-based waste classification system was developed to automatically identify seven categories of waste—Battery, Biological, Brown Glass, Cardboard, Clothes, Shoes, and Trash—using image data. The system demonstrated outstanding performance, achieving nearly perfect classification results as evident from the confusion matrix and ROC-AUC analysis. Specifically, the ROC-AUC values for all classes were 1.00, indicating flawless class separation and model generalization. The confusion matrix showed that most samples were correctly classified, with very few misclassifications across categories. This validates the efficiency and reliability of the trained model in distinguishing visually similar waste items. The model’s

Table 4: Result comparative analysis with previous and proposed method

Ref./ Year	Dataset	Models	Accuracy (%)
[1], 2023	GIGO	DL	95.77
[6], 2025	Garbage	CNN	95
[7], 2024	CDCM	Pretrained models	94.93
[9], 2019	Waste	VGG	80
[10], 2023	Waste	Pretrained models and CNN	87
[11], 2023	Waste	CNN	90.3

Proposed method Robust waste Advanced AI Model 97 ability to handle class imbalance and provide high precision and recall across all classes ensures its suitability for real-world applications, especially in automated waste management and smart bin systems. The proposed approach not only enhances the speed and accuracy of waste sorting but also contributes toward sustainable environmental practices by enabling efficient recycling and disposal processes.

This can be extended to include a broader spectrum of waste classes, such as electronic waste (e-waste), hazardous medical waste, and recyclable plastics, to make the model more comprehensive. Incorporating such granular classification would significantly enhance waste segregation efforts, supporting advanced recycling systems. Integrating this model with real-time edge computing based smart bin systems will allow for on-site waste recognition and segregation, reducing human involvement and error. Coupled with IoT infrastructure, these systems can transmit data to centralized monitoring units for decision-making and policy implementation [15], [16]. Further, Introducing multimodal data (e.g., combining visual data with sensor readings like odor or gas emissions) could further refine classification accuracy, especially in ambiguous cases. This would be especially useful in smart cities and industrial Applications. Deploying federated learning could allow multiple waste management units to collaboratively train the model while preserving data privacy, ensuring both scalability and security.

Data availability: <https://www.kaggle.com/datasets> (Garbage dataset) Funding: We gratefully acknowledge the support of the Institutions of Eminence (IoE) Scheme of Banaras Hindu University, Varanasi under Dev. Scheme No. 6031.

References

- 1) Thamarai, M., Naresh, V.S.: Smart self-power generating garbage management system using deep learning for smart cities. *Microprocessors and Microsystems* 98, 104816 (2023). <https://doi.org/10.1016/j.micpro.2023.104816>.
- 2) Flores, M.G., Tan, J.: Literature review of automated waste segregation system using machine learning: A comprehensive analysis. *Int. J. Simul.: Syst. Sci. Technol.* 11 (2019).
- 3) Li, Y., Liu, W.: Deep learning-based garbage image recognition algorithm. *Appl. Nanosci.* 13(2), 1415–1424 (2023). <https://doi.org/10.1007/s13204-022-02731-4> Oğuz, A., Ertuğrul, " O.F.: Determining the fullness of garbage containers by deep learning. *Expert Syst. Appl.* 217, 119544 (2023). <https://doi.org/10.1016/j.eswa.2023.119544>.
- 4) Li, G., Bilal, A., Ibrar, M., Karim, S., Long, H.: Deepscan: Revolutionizing garbage detection and classification with deep learning. In: *International Conference on Artificial Intelligence and Communication Technology*, pp. 177–194 (2023). Springer.

- 5) Joshi, I., Dev, P., Geetha, G.: Garbage classification using deep learning. In: 2023 International Conference on Circuit Power and Computing Technologies (ICCPCT), pp. 809–814 (2023). <https://doi.org/10.1109/ICCPCT57054.2023.10179786>.
- 6) Tang, Z., Wang, L., Qu, M., Sheng, A., Huai, N.: Design of a garbage classification system based on deep transfer learning. *J. Ambient Intell. Hum. Comput.* 16(1), 225–232 (2025). Springer.
- 7) Ren, H., Liu, H., Qin, J.: Deep Neural Networks for Garbage Classification. In: *Proc. Int. Conf. on Industrial IoT, Big Data and Supply Chain (IIoTBDSC)*, pp. 12–17 (2024). IEEE.
- 8) Hossen, M.M., Ashraf, A., Hasan, M., Majid, M.E., Nashbat, M., Kashem, S.B.A., Kunju, A.K.A., Khandakar, A., Mahmud, S., Chowdhury, M.E.H.: GCDN-Net: Garbage classifier deep neural network for recyclable
- 9) Adediji, O., Wang, Z.: Intelligent waste classification system using deep learning convolutional neural network. *Procedia Manufacturing* 35, 607–612 (2019). <https://doi.org/10.1016/j.promfg.2019.06.026>.
- 10) Munir, M.T., Li, B., Naqvi, M.: Revolutionizing municipal solid waste management (MSWM) with machine learning as a clean resource: Opportunities, challenges and solutions. *Fuel* 348, 128548 (2023). <https://doi.org/10.1016/j.fuel.2023.128548>.
- 11) Kumar, K., Verma, A., Verma, P., Tripathi, S.L.: Unveiling Mental Stress: Examining the Impact on College Students' Well-being Using Machine Learning Based Solutions. In: *2024 IEEE International Conference of Electron Devices Society Kolkata Chapter (EDKCON)*, pp. 179–184. IEEE (2024).
- 12) Kumar, K., Verma, A., Verma, P.: MCDPS: Enhancing Clinical Decision Support for Multiple Chronic Disease Prediction Systems Using Ensemble Machine Learning Approaches. In: *International Conference on Intelligent Systems and Pattern Recognition*, pp. 334–345. Springer Nature Switzerland, Cham (2024).
- 13) Verma, A., Verma, P., Pattanaik, K.K., Garg, L. (eds.): *Research Advances in Intelligent Computing*. (2023).
- 14) Kumar, K., Rajput, N.S.: Analysis Space Transformation Based Electronic Nose for Efficient Detection and Monitoring of Volatile Organic Compounds Gases/Odors in Smart Homes. *Eur. Chem. Bull.* 12(7), 513–528 (2023).
- 15) Kumar, K., Verma, A., Verma, P.: IoT-HGDS: Internet of Things integrated machine learning based hazardous gases detection system for smart kitchen. *Internet of Things* 28, 101396 (2024). <https://doi.org/10.1016/j.iot.2024.101396>.

Polyp-EfficientNet: Lightweight CNN for Accurate Polyp Segmentation

Harsha Vardhan Lakavath, Rakesh Kumar Dhayal, Devshri Satyarthi, and
K.V. Arya

vardhanharsha607@gmail.com, rk157082@gmail.com,
devshri@iiitm.ac.in, kvarya@iiitm.ac.in

Multimedia and Information Security Research Group, Department of Computer
Science and Engineering, ABV-IIITM Gwalior-474015, India

Abstract. Polyp segmentation plays a crucial role in the early detection of colorectal cancer, significantly aiding in diagnosis and treatment planning. We propose Polyp-EfficientNet, built on EfficientNet for compactness and optimized for polyp segmentation. With a residual decoder and attention blocks, it ensures accurate and real-time performance. Tested on five datasets, it surpasses state-of-the-art models.

Keywords: Polyp Segmentation · Convolutional Neural Networks · Residual Learning · Attention Mechanism · Colonoscopy

1 Introduction

Colon cancer diagnosis relies heavily on colonoscopy, which enables early detection and removal of polyps. To support this, computer-aided detection (CADe) systems improve outcomes by automating image analysis. A key task is polyp segmentation—identifying polyp boundaries in colonoscopic images. As illustrated in Fig. 1, this is challenging due to variations in polyp size, shape, and similarity to surrounding tissues. Moreover, many deep learning models

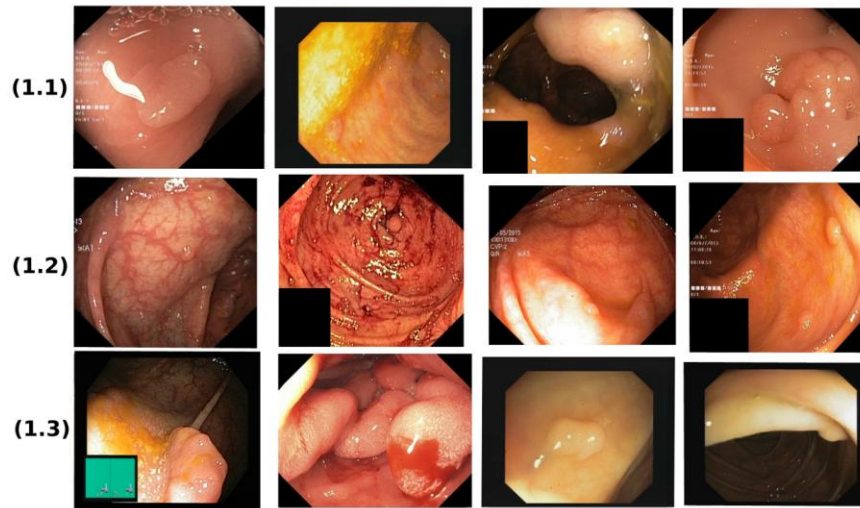


Fig. 1: Various types of polyps found during colonoscopy screening: (1.1) Blurriness, low-quality images and different lighting conditions (1.2) Small size polyps, (1.3) Different shapes and sizes of polyps This research introduces Polyp-EfficientNet, a convolutional neural network (CNN) designed for accurate and efficient polyp segmentation in medical images. The primary contributions of Polyp-EfficientNet are as follows:

- The model adopts a residual decoder architecture that enhances gradient flow and enables multiscale feature fusion, improving the model’s capacity to learn rich representations from complex scenes.
- It integrates channel and spatial attention mechanisms within the decoder to allow the network to emphasize informative features and suppress irrelevant background, thereby improving focus on subtle polyps.
- The overall design maintains a lightweight structure, making it viable for real-time deployment in computer-aided diagnosis systems without compromising accuracy.

The remainder of this paper is organized as follows: Section 2 reviews related work, Section 3 outlines the methodology of Polyp-EfficientNet, Section 4 details the training procedures, Section 5 presents experimental results, Section 6 discusses the findings, and Section 7 concludes the study.

2 Related Work

This section explores existing research directly related to polyp segmentation using deep learning, particularly convolutional neural networks (CNNs), attention mechanisms, and encoder-decoder structures. Early efforts in polyp segmentation focused on traditional computer vision techniques that utilized handcrafted features based on color, texture, and geometric patterns. Karkanis et al. [10] developed a method using color wavelets and linear discriminant analysis, while Iakovidis et al. [7] employed a combination of color histograms and co-occurrence matrices alongside support vector machines (SVMs). These methods offered initial insights into polyp detection by leveraging statistical image features. However, they suffered from significant limitations, including poor generalization across diverse patient populations, high sensitivity to imaging artifacts such as lighting variation or motion blur, and an inability to effectively handle polyps with irregular shapes, small sizes, or textures similar to the surrounding mucosa [15].

To overcome the limited contextual understanding of traditional approaches, transformer-based architectures were introduced in medical image analysis. Models like Polyp-PVT and PVT-CASCADE [19] [20] applied global self-attention and hierarchical feature fusion, enabling effective modeling of long-range dependencies and improving segmentation accuracy, especially for complex polyp structures. These designs marked a shift from purely local feature extraction to holistic scene understanding. Despite their improved performance, these models often came with increased computational overhead, making them less practical for real-time or resource-constrained deployment. Moreover, they sometimes struggled with localizing fine-grained details, especially when polyps had weak boundaries or low contrast. CAFE-Net, which integrated a Transformer-based encoder with convolutional refinements, showed better contextual reasoning and improved segmentation fidelity, yet its model size and inference latency remained a concern for widespread clinical adoption.

Meanwhile, convolutional neural networks (CNNs) have become the cornerstone of medical image segmentation due to their balance between performance and computational cost. U-Net [14], a widely adopted architecture, introduced an encoder-decoder framework with skip connections that preserved spatial information and facilitated effective learning from limited annotated data. Building on U-Net's design, Akbari et al. [1] enhanced the architecture by employing pre-trained encoders for improved feature extraction and refined decoders for better resolution recovery. Further advancements such as Focus U-Net by Yeung et al. [18] incorporated spatial and channel attention mechanisms, allowing the network to focus on salient regions and suppress irrelevant features, thereby improving the detection of small and flat polyps.

3 The proposed Polyp-EfficientNet

A model for polyp segmentation was proposed using EfficientNet to enhance computational efficiency while maintaining accuracy with fewer parameters. The network architecture of the proposed model is outlined as follows:

3.1 Network Architecture

Polyp-EfficientNet follows a U-Net-inspired encoder-decoder architecture optimized for medical image segmentation. The encoder is built on EfficientNet- B3, which provides a strong balance between accuracy and

efficiency through compound scaling of depth, width, and resolution. It extracts rich hierarchical features, which are passed to a lightweight decoder comprising residual convolutional blocks enhanced with squeeze-and-excitation (SE) attention to refine feature representation. Skip connections between corresponding encoder and decoder layers help preserve spatial precision. Finally, a 1×1 convolution layer produces the segmentation map, as illustrated in Fig. 2.

3.2 Residual Learning

Residual learning facilitates gradient flow by linking input directly to output within decoder blocks. Polyp-EfficientNet uses these blocks—comprising 1×1 and 3×3 convolutions with SE attention—to merge fine-grained features and context, improving segmentation accuracy without increasing model complexity.

3.3 Squeeze-and-Excitation Attention

SE blocks enhance feature maps by recalibrating channel importance. Through global average pooling and a lightweight MLP, the model learns which channels to emphasize. This mechanism, embedded in the decoder, strengthens polyp-related feature learning while maintaining computational efficiency.

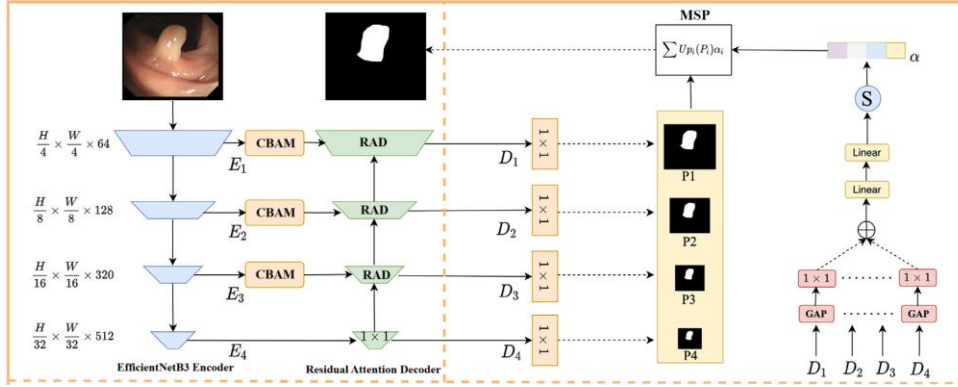


Fig. 2: Polyp-EfficientNet Model Architecture

3.4 Training Strategies

– Loss Function: To optimize segmentation accuracy, we employ a composite loss function that balances Binary Cross-Entropy (BCE) and Dice loss. BCE handles pixel-wise classification by penalizing incorrect predictions, while Dice loss enhances the overlap between predicted and ground truth masks, making it particularly effective for addressing class imbalance. The final loss function is defined as:

$$L_{\text{Total}} = \alpha \cdot L_{\text{BCE}} + (1 - \alpha) \cdot L_{\text{Dice}}$$

where $\alpha = 0.5$ in our experiments.

– Optimization and Regularization: The Adam optimizer [11] adjusts the learning rate based on gradient moments, ensuring efficient convergence. To reduce overfitting, we apply L2 weight regularization to the convolutional layers and dropout regularization [16] in the decoder, encouraging diverse feature learning and preventing reliance on specific neurons.

4 Experiments and Results

This section presents the experiments conducted to evaluate the performance of the proposed Polyp-EfficientNet model for colonoscopy polyp segmentation.

4.1 Datasets

We evaluated Polyp-EfficientNet on five publicly available colonoscopy datasets: Kvasir-SEG [9], CVC-ClinicDB [2], EndoScene [17], EndoTect [6], and Kvasir- Sessile [6]. Kvasir-SEG (1,000 images) and CVC-ClinicDB (612

images) were split into training and testing sets. EndoScene (60 images) and EndoTect (1,000 images) were used to assess generalization and robustness, respectively. Kvasir- Sessile (196 images), a subset of Kvasir-SEG, evaluated the detection of small polyps. All images were resized to 224×224 and normalized. Additionally, 10% unseen samples from Kvasir-SEG and CVC-ClinicDB were used for generalization testing.

4.2 Experimental Details

The model was implemented using TensorFlow 2.15.0 with the Keras API [3], employing the Adam optimizer and a polynomial learning rate schedule over 150 epochs. Post-processing techniques, such as CRF and morphological operations, were applied to improve segmentation accuracy. Testing was done on an Nvidia T4 GPU using Google Colab, with the model’s source code made open-source for reproducibility. The performance of our model, we use the same evaluation metrics as those used in recent studies on polyp segmentation [5] [4]. These metrics were calculated using the evaluation toolbox provided by the authors of Fan et al. [5]. The toolbox computes several standard metrics for image segmentation such as Mean Dice coefficient (mDice), Mean Intersection over Union (mIoU), Mean Absolute Error (MAE), Weighted F-measure, S-measure, and E-measure.

4.3 Analysis of Model

To validate the performance of Polyp-EfficientNetB3, we benchmark it against top-performing segmentation models: HSNNet [20], PVT-CASCADE [13], and CAFE-Net [12]. These models represent diverse architectures and serve as strong baselines in both CNN and transformer-based polyp segmentation research. On the Kvasir-SEG dataset, as shown in Table 1, Polyp-EfficientNetB3 achieves top performance with an mDice of 0.969, mIoU of 0.940, and $F\omega\beta$ of 0.972, while using only 11.5M parameters—significantly fewer than CAFE-Net and PVT-CASCADE as shown in Table 2, Polyp-EfficientNetB3 achieves 0.962 mDice, 0.929 mIoU, and 0.964 $F\omega\beta$, outperforming others in core metrics while being more efficient than heavier models like CAFE-Net. Table 3 shows that while CAFE-Net leads in mDice and mIoU, Polyp-EfficientNetB3 achieves the lowest MAE (0.006) and a strong $F\omega\beta$, demonstrating reliable generalization across clinical settings. Table 4 shows that Polyp-EfficientNetB3 achieves strong performance on EndoTect with an mDice of 0.886, $F\omega\beta$ of 0.897, and mIoU of 0.8185, demonstrating robustness on challenging samples.

Table 5 shows that Polyp-EfficientNetB3 achieves 0.936 mDice, 0.889 mIoU, 0.938 $F\omega\beta$, and 0.941 $S\alpha$ on the Kvasir-Sessile dataset, proving its effectiveness in segmenting smaller, harder-to-detect sessile polyps. In summary, Polyp- EfficientNetB3 outperforms or competes closely with larger models like CAFENet and PVT-CASCADE, while maintaining a compact, efficient architecture suitable for real-time clinical deployment.

Table 1: Performance Comparison of Different Methods on Kvasir-SEG

Method	Year	mDice	mIoU	$F\omega\beta$	$S\alpha$	mE	maxE	MAE
HSNet	2022	0.928	0.884	0.924	0.933	0.974	0.977	0.017
PVT-CASCADE	2023	0.924	0.875	0.920	0.929	0.960	0.964	0.022
CAFE-Net	2024	0.933	0.889	0.927	0.939	0.967	0.971	0.019
Polyp-EfficientNet	2024	0.969	0.940	0.972	0.969	0.968	0.968	0.006

Table 2: Performance Comparison of Different Methods on CVC-ClinicDB

Method	Year	mDice	mIoU	$F\omega\beta$	$S\alpha$	mE	maxE	MAE
HSNet	2022	0.936	0.890	0.940	0.962	0.986	0.991	0.006
PVT-CASCADE	2023	0.943	0.895	0.940	0.952	0.983	0.987	0.006
CAFE-Net	2024	0.943	0.899	0.941	0.957	0.986	0.991	0.006
Polyp-EfficientNet	2024	0.962	0.929	0.964	0.963	0.962	0.962	0.005

Table 3: Generalisation Test on EndoScene Dataset

Method	Year	mDice	mIoU	$F\omega\beta$	$S\alpha$	mE	maxE	MAE
HSNet	2022	0.922	0.884	0.868	0.937	0.960	0.961	0.008
PVT-CASCADE	2023	0.893	0.831	0.872	0.929	0.945	0.957	0.010

CAFE-Net	2024	0.901	0.834	0.882	0.939	0.971	0.981	0.006
Polyp-EfficientNet	2024	0.896	0.824	0.910	0.906	0.896	0.896	0.006

Table 4: Generalisation Test on EndoTect Dataset

Method	Year	mDice	mIoU	F ω β	S α	mE	maxE	MAE
Polyp-EfficientNet	2024	0.886	0.818	0.897	0.902	0.886	0.886	0.030

Table 5: Generalisation Test on Kvasir-Sessile Dataset

Method	Year	mDice	mIoU	F ω β	S α	mE	maxE	MAE
Polyp-EfficientNet	2024	0.936	0.889	0.938	0.941	0.936	0.936	0.008

4.4 Computational Efficiency

Polyp-EfficientNet outperforms HSNet, PVT-CASCADE, and CAFE-Net, as shown in Table 6, with 11.5 million parameters, 0.0230 GFLOPs, and a 179.2 ms inference time per image. These results highlight its strong balance of accuracy and efficiency, making it ideal for real-time clinical use on constrained hardware.

4.5 Clinical Translation Potential

These features position Polyp-EfficientNet as a strong solution for computeraided detection in colorectal cancer screening. Accurate polyp segmentation is crucial for enhancing colonoscopy-based screening or diagnostics. Reliable Monitoring: Tracks polyp size and morphology, helping assess malignancy risk over time.

- Improved Detection: Enhances visibility of subtle lesions, reducing oversight during exams.
- Support for Diagnosis: Enables automated risk classification based on polyp characteristics.
- Real-Time Use: Compact size and fast inference make it ideal for real-time integration in smart devices.

Table 6: Computational Efficiency Comparison

Method	Year	Type	Params (M)	GFLOPs (G)
HSNet	2022	Transformer	29.257	10.943
PVT-CASCADE	2023	Transformer	35.273	15.404
CAFE-Net	2024	Transformer	35.530	16.119
Polyp-EfficientNet	2024	CNN	11.499	0.0230

5 Conclusion

Polyp-EfficientNet, a novel CNN model, demonstrates superior performance in polyp segmentation for colonoscopy procedures, offering enhanced accuracy, efficiency, and real-time applicability. By integrating residual learning and attention mechanisms, it excels in detecting polyps across varied datasets while maintaining a lightweight architecture suitable for clinical deployment. Our experiments confirm its potential for improving colorectal cancer screening and diagnosis. Future work will focus on expanding the model for multi-class segmentation and integrating uncertainty estimation techniques to enhance reliability. Collaborationswith radiologists and access to diverse datasets will help refine its performance and expand its clinical applicability.

References

- 1) Akbari, M., Mohrekesh, M., Nasr-Esfahani, E., et al.: Polyp segmentation in colonoscopy images using fully convolutional network. In: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, pp. 69–72 (2018).

- 2) Bernal, J., Sánchez, F.J., Fernández-Esparrach, G., Gil, D., Rodríguez, C., Vilariño, F.: Wm-dova maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. *Computerized Medical Imaging and Graphics* 43, 99–111 (2015).
- 3) Chollet, F., et al.: Keras. (2015). Accessed: April 12, 2024. 4. Dong, B., Wang, W., Fan, D.P., Li, J., Fu, H., Shao, L.: Polyp-pvt: Polyp segmentation with pyramid vision transformers. *arXiv preprint arXiv:2108.06932*, (2021).
- 4) Fan, D.P., Ji, G.P., Zhou, T., Chen, G., Fu, H., Shen, J., et al.: Pranet: Parallel reverse attention network for polyp segmentation. In: *International conference on medical image computing and computer assisted intervention*, Springer, pp. 263–273 (2020).
- 5) Hicks, S.A., Jha, D., Thambawita, V., Halvorsen, P., Hammer, H.L., Riegler, M.: The endotect 2020 challenge: Evaluation and comparison of classification, segmentation and inference time for endoscopy. *Lecture Notes in Computer Science*, pp. 263–274 (2021).
- 6) Iakovidis, D.K., Maroulis, D.E., Karkanis, S.A., Brokos, A.: A comparative study of texture features for the discrimination of gastric polyps in endoscopic video. In: *18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05)*, Dublin, Ireland, pp. 575–580 (2005).
- 7) Jha, D., Smedsrud, P., Johansen, D., de Lange, T., Johansen, H., Halvorsen, P., Riegler, M.: A comprehensive study on colorectal polyp segmentation with resnet++, conditional random field and test-time augmentation. *IEEE Journal of Biomedical and Health Informatics* 25, 2029–2040 (2021).
- 8) Jha, D.H., Smedsrud, P.H., Riegler, M.A., Halvorsen, P., de Lange, T., Johansen, D.e.a.: Kvasir-seg: A segmented polyp dataset. In: *Multi-Media Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, January 5–8, 2020, Proceedings, Part II*, Springer, pp. 451–462 (2020).
- 9) Karkanis, S., Iakovidis, D., Maroulis, D., Karras, D., Tzivras, M.: Computer-aided tumor detection in endoscopic video using color wavelet features. *IEEE Transactions on Information Technology in Biomedicine* 7, 141–152 (2003).
- 10) Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, (2014)Liu, G., Yao, S., Liu, D., Chang, B., Chen, Z., Wang, J., Wei, J.: Cafe-net: Crossattention and feature exploration network for polyp segmentation. *Expert Systems with Applications* 238, 121754 (2024).
- 11) Rahman, A., Marculescu, R.: Pvt-cascade: Pyramidal vision transformer cascade for polyp segmentation. *arXiv preprint arXiv:2302.07486*, (2023).
- 12) Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: *Medical image computing and computer-assisted intervention– ICCAI 2015: 18th International conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III*, Springer, pp. 234–241 (2015).
- 13) Sengar, S.S., Mukhopadhyay, S.: Moving object detection based on frame difference and w4. *Signal, Image and Video Processing* 11, 1357–1364 (2017).
- 14) Srivastava, N., Hinton, G., Krizhevsky, A., Salakhutdinov, R.: Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* 15, 1929–1958 (2014).
- 15) Vázquez, D., Bernal, J., Sánchez, F., Fernández-Esparrach, G., López, A., Romero, A., Drozdal, M., Courville, A.: A benchmark for endoluminal scene segmentation of colonoscopy images. *Journal of Healthcare Engineering* 2017, 1–9 (2017).
- 16) Yeung, M., Sala, E., Schönlieb, C., Rundo, L.: Focus u-net: A novel dual attentiongated cnn for polyp segmentation during colonoscopy. *Comput Biol Med* 137,104815 (2021).
- 17) Zhang, W., Lu, F., Su, H., Hu, Y.: Dual-branch multi-information aggregation network with transformer and convolution for polyp segmentation. *Computers in Biology and Medicine* 168, 107760 (2024).
- 18) Zhang, W., Qin, Z., Wang, S., Luo, W., Zheng, F.: Hsnet: Hybrid segmentation network for polyp segmentation. *arXiv preprint arXiv:2203.14915*, (2022).

A comprehensive review of metal matrix composites in additive manufacturing: materials, fabrication, and industrial applications

DEEPAK KUMAR SAHU, V.K. TRIPATHI

Department of Mechanical Engineering, National Institute of Technical Teachers Training and Research, Bhopal

Abstract

Metal matrix composites (MMCs) have gained significant attention in recent decades due to their unique properties and potential applications across various industries. Additive manufacturing (AM) is rapidly transforming the production of MMCs, offering advantages such as enhanced microstructure, improved mechanical properties, and increased design flexibility compared to traditional manufacturing methods. AM-produced MMCs exhibit more uniform reinforcement distribution, refined microstructures, and lower porosity, resulting in mechanical properties that are comparable or superior to those made by conventional methods. Furthermore, AM allows for the creation of complex geometries, functionally graded materials, and lattice structures that are difficult or impossible to achieve with traditional manufacturing. AM can process a wide range of MMCs, including those based on aluminium, titanium, nickel, magnesium, and steel, and can incorporate various reinforcement types such as ceramics and refractory metals. However, AM also presents challenges and limitations, such as process control issues, feedstock preparation difficulties, and a lack of established standards and guidelines for producing defect-free MMCs. Despite these challenges, AM MMCs are increasingly being used in aerospace, automotive, biomedical, and high-performance engineering applications due to their tailored properties and design possibilities. Ongoing research aims to address these challenges and fully realize the potential of AM in MMC fabrication.

Keywords: Metal matrix composites (MMCs), Additive manufacturing (AM), Production of MMCs.

1. Introduction

Metal matrix composites (MMCs) are engineered materials made by combining metals with reinforcing phases such as particles, fibers, or plates. Over recent decades, MMCs have evolved from a scientific curiosity to materials of significant technological and commercial importance, driven by their superior mechanical and functional properties compared to conventional metals. MMCs have transitioned from research topics to widely used materials in industries such as transportation, aerospace, electrical and electronics, infrastructure, automobile industry, aerospace industry, medical devices, military applications, industrial machinery, marine industry, with global markets growing significantly since the late 20th century.

Additive manufacturing (AM) has significantly advanced the development and application of metal matrix composites (MMCs), offering new possibilities for material design, performance, and manufacturing complexity. The evolution of MMCs in AM is marked by improvements in microstructure control, reinforcement distribution, and the ability to create complex, high-performance components. Traditional MMC fabrication methods like casting and powder metallurgy faced challenges such as poor reinforcement distribution and limited geometric complexity. AM, especially techniques like selective laser melting (SLM) and laser powder bed fusion, has enabled more uniform reinforcement dispersion, refined microstructures, and the production of intricate shapes. The development of in-situ reinforced powders and tailored feedstocks has improved the quality and performance of AM MMCs, allowing for better control over microstructure and properties. AM has facilitated the use of various matrix materials (Al, Mg, Ti, Ni, Fe) and reinforcements, broadening the range of achievable properties and applications. The evolution of MMCs in additive manufacturing has enabled the production of advanced materials with superior properties and design flexibility. AM overcomes many limitations of traditional MMC processing, offering uniform reinforcement, complex shapes, and enhanced performance, making MMCs increasingly valuable for high-tech industries.

Metal matrix composites (MMCs) are engineered materials that combine a metal matrix with reinforcing phases, such as ceramics or other metals, to achieve superior mechanical, thermal, and chemical properties. The integration of MMCs into additive manufacturing (AM) processes has opened new avenues for producing advanced materials with tailored properties, complex geometries, and enhanced performance for demanding applications. Additive manufacturing enables the layer-by-layer fabrication of MMCs directly from digital models, offering significant

advantages over traditional manufacturing methods like casting and powder metallurgy. These advantages include improved reinforcement distribution, refined microstructures, and the ability to fabricate functionally graded and geometrically complex components

AM also allows for the use of a wide range of matrix materials, including aluminum, magnesium, titanium, nickel, and their alloys, each reinforced with various particles or fibers to meet specific application requirements

The adoption of metal-based composites produced via AM is rapidly expanding across multiple industries. In aerospace and automotive sectors, MMCs are valued for their high strength-to-weight ratios, wear resistance, and thermal stability, enabling the production of lightweight yet durable structural components. Biomedical, electronic, and high-end engineering industries also benefit from MMCs' exceptional wear resistance, chemical inertness, and performance at elevated temperatures. Furthermore, AM technologies such as laser powder bed fusion, selective laser melting, and direct energy deposition have been instrumental in advancing the fabrication of MMCs, allowing for the development of novel materials and applications. Despite these advancements, challenges remain, including the optimization of process parameters, control of defects, and the development of standards for defect-free MMC production. Ongoing research continues to address these issues, aiming to fully realize the potential of MMCs in additive manufacturing for next-generation industrial applications.

2. Literature Review

In order to understand the metal matrix composites (MMCs), it is essential to accurately review the behaviour of metal matrix composite structures, their constituent materials, manufacturing processes, and applications. In this context, Sheikh and Khan [2] highlighted a comprehensive analysis of optimizing zinc-based MMCs to meet specific industrial requirements, supporting ongoing research into this versatile material class. Ivey et al. [3] studied the characterization of short-fiber-reinforced composites produced by Material Extrusion Additive Manufacturing (MEAM) using commercial PLA and PLA reinforced with short carbon fibers (PLA/CF). Azam et al. [4] investigated the mechanical and physical property differences between parts produced by additive manufacturing (AM) and those manufactured using conventional methods. Tang et al. [5] work presents the achievements of AMMCs fabricated by additive manufacturing. Additive manufacturing of aluminum-based metal matrix composites (AMMCs) offers significant advantages in microstructure and properties compared to conventional processes, with potential for future applications in aerospace, automobile, and other structural applications. Dadkhah et al. [6] this study shows Additive manufacturing (AM) offers potential for producing functionally graded metal matrix composites and complex objects, but challenges like rapid cooling and lack of suitable parameters require further research. Zhang et al. [7] study tells the Laser powder bed fusion technology has become the dominant additive manufacturing technology for aluminium alloys, with significant progress in microstructure, properties, and applications over the past 15 years. Y. Karabulut and R. Ünal. [8] This study tells Laser powder bed fusion (LPBF) is a promising additive manufacturing method for producing ceramic particle-reinforced aluminium-based metal matrix composites with advanced mechanical properties for aerospace, automotive, and biomedical applications. Z. Wang *et al.* [9] Laser additive manufacturing (AM) has shown potential in producing high-performance, complex nickel-based alloy metal matrix composites for aerospace, naval ships, and industrial gas turbine engines, but manufacturing defects still exist. N. Li *et al.* [10] The review primarily focuses on laser additive manufacturing of five types of MMCs. It covers essential aspects such as material design, the combination of reinforcement and the metal matrix, the synthesis principle during manufacturing, and the resulting microstructures and properties. S. Lathabai [11] studied that Additive manufacturing has revolutionized materials processing, becoming a key technology for high-performance materials development. AM's rise over the last four decades showcases its importance in modern manufacturing. D. Zhang *et al.* [12] this work addresses the challenges and opportunities in employing metal alloys within additive manufacturing (AM). Given the inherent advantages of metal AM in terms of complexity, time efficiency, and waste reduction, the paper emphasizes the importance of understanding alloying strategies and recent advancements in materials performance to broaden the range of alloys suitable for AM. The review by K. Osipovich et al. [13] focuses on wire-feed electron beam additive manufacturing (WAAM-EB), covering recent advancements and the characteristics of structure formation in various metals and alloys during 3D printing. The paper also considers directional solidification during the production of components and discusses the prospects for creating composite and functionally graded materials

using different metals and alloys. B. S. Rao and T. B. Rao,[14] The review addresses the increasing use of aluminium alloys in various engineering sectors like aerospace and automotive, highlighting their advantages such as light weight, high strength, ductility at low temperatures, corrosion resistance, and non-toxicity. It emphasizes the role of 3D printing in manufacturing these alloys and their composites with improved precision, reduced lead times, and cost-effectiveness compared to conventional method. Z. Zhu *et al.* [15] This review paper focuses on the recent advancements in additive manufacturing (AM) of aluminum alloys and aluminum matrix composites. Although AM adoption for aluminum alloys is slower compared to steels and titanium alloys, significant progress has been made in the past 15 years. This includes developments in novel processes, alloys, heat treatments, and applications. H. Fan *et al.* [16] This paper provides a comprehensive review of Laser Additive Manufacturing (LAM) for aluminum alloys, focusing on methods, microstructures, and mechanical properties. This also emphasizes the importance of optimizing laser processing parameters to achieve high-quality LAMed aluminum structures with desirable microstructures and mechanical properties. M.Ostolaza *et al.* [17] This paper addressed the latest developments in manufacturing metal matrix composites (MMCs) and functionally graded materials (FGMs) through additive manufacturing (AM), particularly focusing on laser-directed energy deposition (L-DED). It highlights that while traditional methods cannot produce multi-material structures in one operation, AM offers a promising alternative. B. Karaş *et al.*[18] This paper discusses a new method for additive manufacturing (AM) of high-density carbon fibre reinforced polymer (CFRP) composites. Traditional AM methods often result in composites with high porosity (18-25%), which is much higher than the 1% porosity achieved by conventional manufacturing techniques. The authors introduce Composite Fibre Additive Manufacturing (CFAM), a process that involves selectively printing a binder and polymer powder onto discontinuous carbon fibre sheets, followed by compression, heating, and post-processing to form final components. Various materials, manufacturing processes, and applications of metal matrix composites (MMCs) fabricated via additive manufacturing (AM) are detailed in [4], [13]–[16], [5]–[12], Several studies have been conducted in recent years on the structural analysis of functionally graded materials (FGM). In this framework, Gupta and Talha [32] emphasized the critical review on the recent development in modelling and analysis of FGMs and structures. Khan and Riccio [33] presented a comprehensive review of lattice structures used in aerospace lightweight applications. Boggarapu *et el.* [34] provided an overview of advancements in FGM structures, including material selection, processing techniques, analytical modelling, and their applications. Saleh *et el.* [35] presented extensive review of manufacturing methods for sandwich FGMs based on the available literature over 30 years. Kumar *et al.* [36] emphasized a comprehensive review on the FGMs, covering their evolution, classification, manufacturing methods, and the advantages and limitations of each type. It also discusses the current industry applications, key challenges, and future research directions for FGMs. Gupta *et al.* [37] studied the vibration characteristics of shear deformable FGM plates using FEM. Singh and Gupta [38], [39] investigated the effects of geometric imperfections and porosity on the vibrational frequencies of SFGM plates. Singh and Gupta [40] studied the impact of circular and elliptical cutouts on vibrational frequencies of the SFGM plates with FGM core subjected to elastic foundation. Singh *et al.* [41] examined the influence of thermal and hygrothermal conditions on the vibrational characteristics of the SFGM plates with cut-outs. Singh *et al.* [42] investigated the impact of geometric discontinuities on the vibrational frequencies of SFGM plates using HSDT, FEM using ANN techniques. Recent research widely explores conventional and advanced composite materials and fabrication techniques to develop complex structures including MMCs, FGM plates, SFGM plates, FGM beams, and nanobeams using additive manufacturing (AM) methods. In this context, Quan *et al.* [43] investigated the natural frequency analysis of a sandwich plate placed on Winkler-Pasternak type elastic foundations in a thermal environment. The sandwich plate features an auxetic honeycomb core with a negative Poisson's ratio, perfectly bonded to two carbon nanotube reinforced composite (CNTRC) face sheets, thereby ignoring any potential interface effects or bonding imperfections. Mahesh *et al.* [44] studied the nonlinear deflections of smart sandwich plates that have an agglomerated Carbon Nanotubes (CNTs) porous core and piezo-magneto-electric (PME) facings. The authors introduce a new approach combining Finite Element Method (FEM) with Artificial Neural Networks (ANN). This is the first time an ANN-based tool has been used to integrate the effects of CNT agglomeration, porosity, and pyro-coupling in PME materials. Afsal *et al.* [45] investigated An Element Free Galerkin (EFG) meshless formulation and solutions by using higher order shear deformation principle with 9 degrees of freedom for the static analysis of Functionally Graded Material (FGM) plates are provided. This

theory help to estimates the shape function using Moving Least Squares (MLS) method. Hoang et al. [46] This paper focuses on the free vibration analysis of functionally graded graphene platelet-reinforced composite (FG-GPLRC) plates in a liquid medium. A novel quasi-three-dimensional shear deformation theory is introduced, and material properties are determined using the Halpin-Tsai micromechanical model. Differential motion equations incorporating transverse shear and normal stresses are derived via Hamilton's principle. The natural frequencies are computed using Galerkin's method and validated against published results.

Research gap:

The literature review reveals a notable gap in the comprehensive study of metal matrix composites (MMCs) fabricated via additive manufacturing, particularly in understanding the relationships between reinforcement selection, processing parameters, and resulting material performance across diverse industrial applications. [47]Laser additive manufacturing opens opportunities for processing novel alloys and composites that are difficult to process using conventional techniques, yet this area requires more exploration. [7]More research is required to overcome the limitations of mechanical properties in magnesium alloys by introducing reinforcements via additive manufacturing.[48] Further research is needed to explore and optimize in-situ reinforcement capabilities within Al/Al-alloy MMCs during laser processing to achieve high-quality feedstock material development.

Novelty aspect:

This study presents a novel and comprehensive review of metal matrix composites (MMCs) in the context of additive manufacturing (AM), with a focus on materials, processing methods, and industrial applications. Unlike previous reviews that treat MMCs and AM separately, this work integrates both areas to explore how advanced reinforcements, processing parameters, and microstructural control influence the physical and mechanical performance of MMCs produced via AM. It highlights key challenges such as porosity, reinforcement dispersion, thermal stability, and machinability, while also addressing emerging solutions like hybrid reinforcements, computational modeling, and alternative machining techniques. This review is timely and significant, offering valuable insights for researchers and engineers aiming to develop high-performance, application-specific MMCs for demanding sectors such as aerospace, automotive, biomedical, and electronics.

3.Classification of conventional materials used in AM process

A range of diffident types of conventional and advance materials using additive manufacturing technologies are as following:

Matals: Common Metals in Additive Manufacturing

3.1 Steel:

Applications: Large-scale steel components and structures can be produced using wire arc additive manufacturing (WAAM) [49]. This process involves additional research to understand how thermal cycles modify mechanical and metallurgical properties [49].

Filler Metals: Research is conducted on different steel filler metals to understand the effects on mechanical and metallurgical properties[49].

316L Stainless Steel: This material can be used in both complete construction and repair of tensile specimens using directed energy deposition systems [50].

3.1.1 Aluminum:

Alloys: Aluminum alloys like AlSi10Mg are extensively researched in selective laser melting (SLM) [51].

Impact Response: The influence of second-phase boron particles on the impact response of aluminum has been investigated, showing that spall strength is reduced with higher boron content [52].

3.1.2 Titanium:

Alloys: Titanium alloys are frequently employed in metal additive manufacturing, especially for aircraft components, gas turbines, and other structural areas [53].

Ultrafine-grained Alloys: Additive manufacturing can produce ultrafine-grained high-strength titanium alloys, challenging conventional manufacturing processes for complex parts [54], [55].

3.1.3 Nickel and Cobalt Alloys:

Applications: Nickel and its alloys are frequently preferred for transmission shafts and turbine blades due to their corrosion resistance and magnetic properties [56]. Cobalt alloys are also utilized in metal additive manufacturing [57].

1.1. Common non-metals used in AM include:

Polymers: Acrylonitrile Butadiene Styrene (ABS), Polylactic Acid (PLA), and others are used in FDM [58].

Ceramics: Used for applications requiring high-temperature resistance and chemical stability [58].

4. Classification of metal matrix composites and their working

1.2. Based on Matrix Material; -The matrix material provides the structural framework for the composite, and the reinforcement enhances its properties. Common matrix materials include:

4.1.1 Aluminum (Al): Al MMCs are widely used due to their high strength-to-weight ratio [59]. They find applications in automotive, aerospace, and structural components [60][61].

4.1.2 Magnesium (Mg): Mg MMCs are known for their lightweight properties, making them suitable for applications where weight reduction is critical [62][63].

4.1.3 Copper (Cu): Cu MMCs offer good electrical and thermal conductivity [62]. They are used in electrical contacts, heat sinks, and other applications requiring high conductivity.

4.1.4 Titanium (Ti): Ti MMCs provide high strength and corrosion resistance [62]. They are used in aerospace and marine applications [62].

4.1.5 Superalloys: Superalloy MMCs are used for high-temperature applications due to their excellent high-temperature strength and creep resistance.

4.1.6 Intermetallic Matrix Composites: These offer a combination of metallic and ceramic properties, making them suitable for specialized applications.

4.2 Based on Reinforcement Morphology:-The morphology of the reinforcement significantly influences the properties of the MMC [60]. MMCs can be classified into three broad categories based on reinforcement type [64][65]:

4.2 Continuous Fiber-Reinforced Composites: These composites use continuous fibres embedded in the metal matrix [64]. Continuous fibres provide high strength and stiffness along the fiber direction, making these composites suitable for applications where directional mechanical properties are crucial [66]] .

4.3 Discontinuous Fiber-Reinforced Composites: These composites use short fibers or whiskers as reinforcement [64]. Short fibers or whiskers improve the composite's strength and stiffness in multiple directions to some extent [66].

4.4 Particle-Reinforced Composites: These composites incorporate particles into the metal matrix [64]. The particles can be ceramic, carbide, or other hard materials, which enhance the strength, wear resistance, and other properties of the metal matrix [66].

5. Working Principles and Strengthening Mechanisms:-The enhanced properties of MMCs are achieved through various strengthening mechanisms [67]:

4.5 Load Transfer: The reinforcement bears a portion of the applied load, reducing the stress on the matrix [67].

4.6 Hall-Petch Strengthening: Decreasing the grain size of the metal matrix increases the yield strength of the material [67].

4.7 Orowan Strengthening: Dislocations in the matrix are forced to bypass non-deformable particles, leading to increased resistance to dislocation motion [67]

4.8 Coefficient of Thermal Expansion (CTE) Mismatch: Differences in CTE between the matrix and reinforcement create internal stresses that impede dislocation motion, strengthening the composite [67].

5 Fabrication Methods:-The manufacturing of MMCs involves various techniques to ensure a homogenous distribution of the reinforcement within the metal matrix [68]. These methods include:

5.1 Liquid-Phase Processing: This involves incorporating solid reinforcements into molten metal, followed by solidification [65]. Techniques include stir casting, compo casting, squeeze casting, and infiltration methods [69].

5.2 Stir Casting: Reinforcement particles are mixed into molten metal using mechanical stirring to ensure uniform distribution [69][70][71].

5.3 Infiltration: Liquid metal is infiltrated into a preform of the reinforcement material, using techniques like pressure infiltration, vacuum infiltration, and spontaneous infiltration [69][72].

5.4 Solid-Phase Processing: This involves techniques such as powder metallurgy, diffusion bonding, and hot isostatic pressing [69].

5.5 Powder Metallurgy: Metal powders are mixed with reinforcement particles, compacted, and then sintered to form a solid composite [69][73].

5.6 Other Methods: Other methods include spray deposition, electrochemical deposition, and additive manufacturing [69] [74].

6 Applications:-MMCs are used in a wide range of applications due to their superior properties [75]:

6.1 Aerospace: Components requiring high strength-to-weight ratio and thermal stability [76].

6.2 Automotive: Engine components, brake rotors, and other parts benefiting from wear resistance and high strength [77] [78].

6.3 Sports Equipment: Lightweight and high-strength materials for applications like bicycle frames and tennis rackets [76].

6.4 Military: Armor, structural components, and other defense applications.

6.5 Electronics: Heat sinks and other thermal management components [79].

7 Recent Advances and Future Directions: -Recent research focuses on enhancing the properties of MMCs by incorporating novel reinforcements such as high-entropy alloys (HEAs) and graphene [80]. The use of waste materials as reinforcement is also gaining attention for its environmental benefits.

8 Future research directions include:

8.1 Developing cost-effective manufacturing techniques: Reducing the cost of MMC production to broaden their applications [81].

8.2 Improving interfacial bonding: Enhancing the bonding between the matrix and reinforcement to maximize the composite's properties [82].

8.3 Exploring novel reinforcement materials: Investigating new materials to further improve the performance of MMCs.

9 Analytical study on metal matrix composites Structures:

Metal matrix composites (MMCs) are engineered materials combining metallic properties with ceramic or carbon reinforcements, offering enhanced performance for structural applications [83]. An analytical study of these materials involves modeling their behavior under various conditions, considering factors like reinforcement type, distribution, and interface characteristics [84][85]. The structure of MMCs significantly influences their properties, dictating their suitability for specific applications.

10.1 Reinforcement Types and Structures

MMCs are classified based on the reinforcement type, including particle-reinforced, short fiber/whisker-reinforced, and continuous fiber/sheet-reinforced composites [86]

Particle-Reinforced MMCs: These composites feature a metal matrix with dispersed particles, enhancing strength and wear resistance [86]. The simulation models can accurately analyze and predict composites' micro-stresses and mechanical properties. The conventional representative volume element (RVE) model can only contain one single-particle form and the particles have the same shape, which differs significantly from actual particles. Four typical particle-modeling methods can be adopted to establish geometric models to analyze the distribution of stress.

Short Fiber/Whisker-Reinforced MMCs: These materials incorporate short fibers or whiskers, improving strength and stiffness [86].

Continuous Fiber/Sheet-Reinforced MMCs: These composites utilize continuous fibers or sheets for reinforcement, providing high strength and stiffness along the fiber direction. Unidirectionally solidified MMCs can be fabricated using a static induction coil to melt the alloy and control the solidification direction, enhancing mechanical properties [86].

Analytical Modeling of MMC Behavior:- Analytical studies on MMCs focus on predicting their mechanical behavior, stress distribution, and thermal properties using various models [84][87].

Eshelby's Inclusion Model: Used to calculate the flow behavior and stress-strain curve of MMCs, considering the shape, size, volume fraction, and orientation of the dispersed particles [84].

Finite Element Analysis (FEA): Applied to simulate the microstructure and predict micro-stresses and mechanical properties of composites. FEA can show how the deformation behavior of a metal matrix composite is affected by particle size [85].

Micromechanical Models: Axisymmetric micromechanical finite element models predict the mechanical behavior of aluminum matrix composites reinforced with SiC particles, studying interphase damage under thermal and uniaxial loading[71].

10. Published research work about metal matrix composites (MMCs) using additive manufacturing:

A range of research articles has been published on metal matrix composites (MMCs) fabricated using additive manufacturing technologies as illustrated in from Fig. 1 to Fig. 3. Yearly and cumulative number of research articles have been published on metal matrix composites (MMCs) as mentioned in Fig. 1

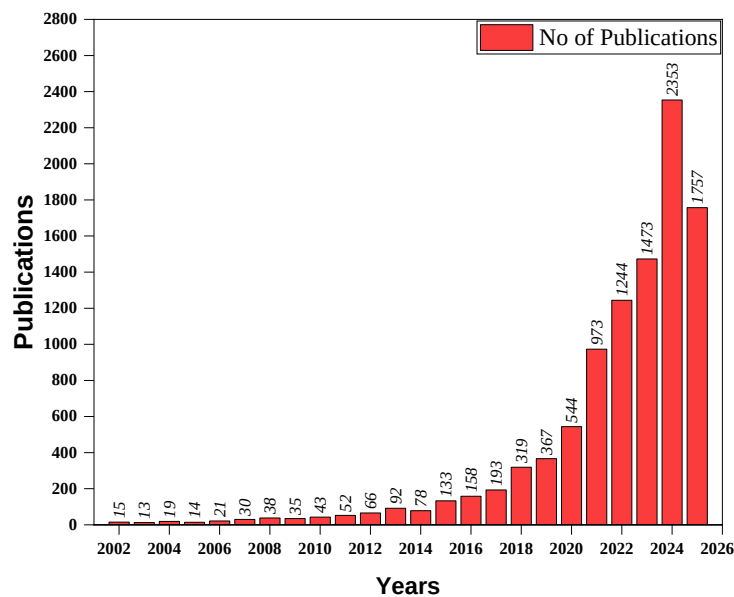


Fig. 1: Yearly and cumulative number of publications since 2002 to 2025, based on a ScienceDirect title search for 'Metal Matrix Composites in Additive Manufacturing' from the available data on (03 June 2025)
Yearly and cumulative number of research articles in the form of PI-chart have been published on metal matrix composites (MMCs) as shown in Fig. 2.

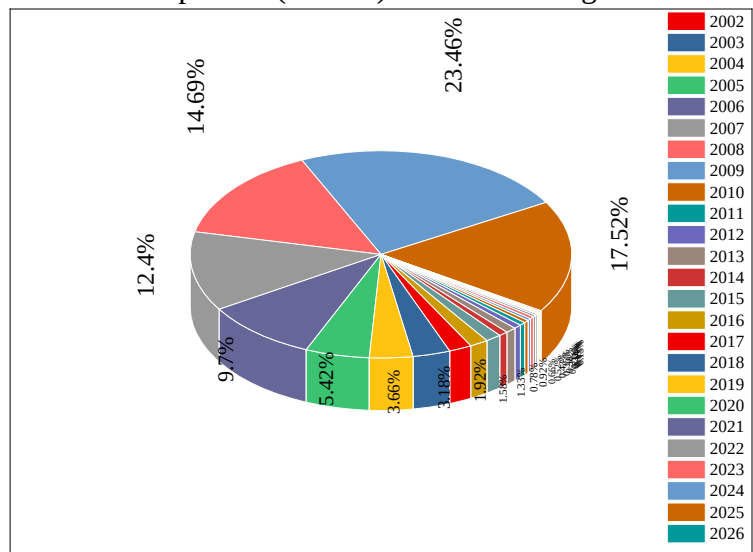


Fig. 2: Yearly and cumulative number of publications since 2002 to 2025, based on a ScienceDirect title search for 'Metal Matrix Composites in Additive Manufacturing' from the available data on (03 June 2025)

A various types of research articles has been published on metal matrix composites (MMCs) as shown in Fig. 3.

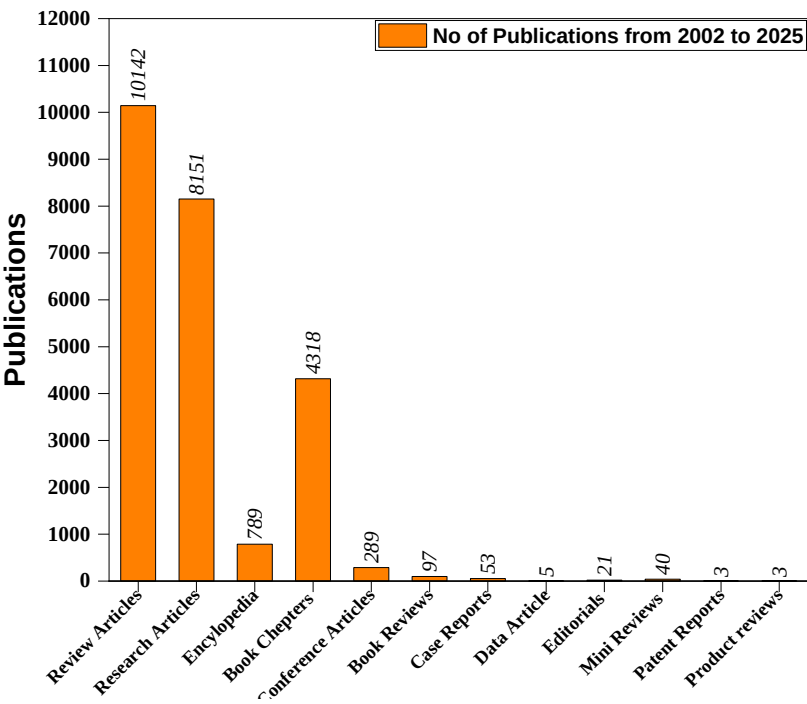


Fig. 3: Yearly and cumulative number of publications since 2002 to 2025, based on a ScienceDirect title search for 'Metal Matrix Composites in Additive Manufacturing' from the available data on (03 June 2025)

10 **Materials:**

The physical properties of MMCs particularly density, porosity, and thermal stability are critical to their performance in industrial applications. In the context of **additive manufacturing (AM)**, these properties are influenced not only by material composition but also by process parameters such as energy input, layer thickness, and scan strategy. While increasing reinforcement content typically results in higher density, the associated enhancements in mechanical performance—such as improved hardness and strength often justify this trade-off. However, **porosity remains a major challenge** in AM-produced MMCs, as it compromises mechanical integrity and accelerates wear. Advanced AM techniques, such as **selective laser melting (SLM)** and **direct energy deposition (DED)**, combined with optimized process controls and **in-situ alloying**, show promise in reducing porosity and achieving more uniform reinforcement distribution.

Thermal stability, assessed through thermogravimetric analysis, indicates that many MMCs maintain structural integrity at temperatures up to ~850°C, supporting their suitability for high-temperature environments. However, degradation at elevated temperatures suggests the need for more thermally stable reinforcements or improved matrix-reinforcement bonding to extend operational limits especially in AM processes involving rapid solidification.

From a **mechanical standpoint**, reinforcement type and distribution significantly influence hardness, tensile strength, ductility, and wear resistance. Additive manufacturing allows for **tailored microstructures**, enabling the strategic inclusion of reinforcements such as SiC, TiC, and MoS₂ to enhance strength and wear resistance. However, higher reinforcement content often results in a **ductile-to-brittle transition**, as shown by fractographic analyses. This underscores the importance of **optimized particle dispersion and interfacial bonding**, which can be influenced by AM parameters like laser power and scan speed.

Strengthening mechanisms such as load-bearing effects, dislocation generation, and grain boundary strengthening (Hall–Petch effect) are key contributors to the mechanical behaviour of AM-fabricated MMCs. Future work should integrate **computational modelling and finite element analysis (FEA)** to simulate these mechanisms and guide the design of AM processes for specific applications.

Machinability remains a concern, particularly due to the abrasive nature of ceramic reinforcements. Conventional machining tools suffer from rapid wear when processing AM MMCs. Alternatives such as **laser machining** or **abrasive water jet cutting**, as well as optimized machining parameters, offer potential solutions to enhance manufacturability while maintaining component quality.

Overall, MMCs produced through additive manufacturing exhibit a compelling balance of strength, thermal stability, and wear resistance, making them attractive for **automotive, aerospace, electronics, and biomedical** applications. However, challenges such as porosity control, reinforcement-matrix compatibility, and machining difficulties must be addressed. This review highlights the need for continued interdisciplinary research, combining material science, manufacturing engineering, and computational modelling, to advance the development and large-scale adoption of MMCs in next-generation engineering systems.

Details of various materials and their properties Aluminium, Titanium, Nickel, Magnesium, steel etc. The various material properties of metal matrix based composite materials as illustrated from Table 5 to Table 7.

Table 5. Material properties of the Aluminu and Alumina [88]

Material		Aluminium (Al)	Alumina (Al ₂ O ₃)
Properties	Young's Modulus (GPa)	72	380
	Density (g/cm ³)	2.7	3.8
	Poisson's Ratio	0.3	0.3

Table 6. Material properties of the Titanium [88]

Material		Ti	Ti-Mo-Zr Alloy
Properties	Young's Modulus (GPa)	115	81
	Density (g/cm ³)	4.5	4.5
	Poisson's Ratio	0.36	0.36

Table 7. Material Properties of the Nickel [88]

Material		Ni	Ni Alloys
Properties	Young's Modulus (GPa)	200	200
	Density (g/cm ³)	8.9	8.9
	Poisson's Ratio	0.31	0.31

Table 8. Material Properties of the Fe [88], [89]

Material		Fe	Pure iron
Properties	Young's Modulus (GPa)	215	204
	Density (g/cm ³)	7.9	7.89
	Poisson's Ratio	0.29	0.29

Table 9. Material Properties of the Steel [90], [91]

Material		Steel	Steel DX54
Properties	Young's Modulus (GPa)	200	182
	Density (g/cm ³)	7.85	7.87
	Poisson's Ratio	0.3	0.3

Table 10. Material Properties of the composites materias (Carbón fiber & Glass Fiber)[92]

Material		Carbon fiber	Glass fiber
sProperties	Young's Modulus (GPa)	70-230	35-90
	Density (g/cm ³)	1.6	1.8-2.0
	Poisson's Ratio	0.2-0.3	0.22-0.3

Table 11. Material Properties of the some (polymer materials used in polymer based 3D printing) [93], [94], [95]

Material		PLA (polylactic acid)	acrylonitrile butadiene styrene (ABS)
Properties	Young's Modulus (GPa)	2.7-3.5	1.5-2.5
	Density (g/cm ³)	1.24	1.04
	Poisson's Ratio	0.36	0.35

11 Various processes techniques:

11.1 Wire-feed electron beam additive manufacturing (WAAM-EB): is an additive manufacturing technique that utilizes an electron beam as a heat source to melt and fuse metallic wire feedstock [13] . This method allows for the production of large-scale metal parts with complex geometries . The focused energy of the electron beam enables precise control over the melting and solidification process, which is crucial for achieving desired material properties and minimizing defect.

11.2 Direct Energy Deposition (DED): A process where metal powder or wire is melted and deposited simultaneously, allowing for larger and more complex structures.[14] it is applicable to Al-Si alloys and composites.

11.3 Binder Jetting: Binder Jetting (BJT) [96] uses a binder material selectively jetted onto a powder bed to bond the powder particles together. A process where a liquid binder is used to selectively join powder particles, followed by sintering to achieve final density and strength [58].

11.4 Stereolithography (SLA): SLA involves the use of a UV laser to cure a liquid resin containing reinforcement particles or fibers. SLA offers high resolution and surface finish but is limited to thermoset matrices and may be challenging to scale up for larger parts.[14]

11.5 Selective Laser Sintering (SLS): SLS uses a laser to selectively sinter or fuse powder particles of a composite material. SLS allows for the creation of complex geometries without the need for support structures. However, SLS is typically limited to thermoplastic matrices and may result in lower mechanical properties compared to other AM techniques.[97]

11.6 Vat Photopolymerization (VPP): This process uses light to cure a liquid photopolymer resin in a vat Sub-processes include Stereolithography (SLA), Digital Light Processing (DLP), and Liquid Crystal Display (LCD) SLA uses a laser to selectively cure the resin , while DLP uses a digital light projector to cure an entire layer at onc.[98]

11.7 Powder bed fusion (PBF): Powder Bed Fusion (PBF) involves spreading a powder bed and using a heat source (such as a laser or electron beam) to selectively fuse the powder particles together to form a layer of the object.[99], [100], [101] Selective Laser Sintering (SLS) uses a laser to sinter powder particles, typically nylon or other polymers, in a powder Direct Metal Laser Sintering (DMLS) is similar to SLS but is used for metal powders . Selective Laser Melting (SLM) completely melts the metal powder to create dense metal parts . Electron Beam Melting (EBM) uses an electron beam as the heat source to melt the metal powder in a vacuum environment, which is suitable for high-melting-point metal. [102]

11.8 Sheet Lamination (SHL): involves bonding layers of sheet material (such as paper, metal, or plastic sheets) together, and the excess material is cut away to form the object[103]. Laminated Object Manufacturing (LOM) involves bonding sheets of paper or other materials together using heat and pressure, and a laser cuts the outline of each layer to form the object. [96]Ultrasonic Additive Manufacturing (UAM) uses ultrasonic waves to bond metal sheets together layer by layer, which can be used to create metal parts with embedded components or for joining different metal materials.

12 Applications:

Metal matrix composites (MMCs) are widely utilized across various industries due to their advantageous properties, such as a high strength-to-weight ratio, superior wear resistance, and excellent corrosion resistance. **Error! Reference source not found.** highlights their broad range of applications, including the automotive and aerospace sectors, medical devices, military equipment, industrial machinery, marine engineering, electronics, structural components, and consumer products.

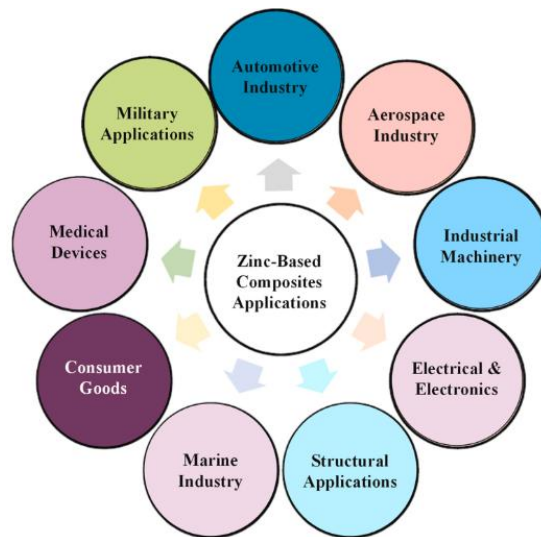


Fig. 4: Applications of metal matrix composites [2]

Some notable applications are listed below:

Automotive Industry: Zinc-aluminum alloy-based MMCs are used in pistons, cylinder liners, and connecting rods. These components benefit from the alloy's wear resistance and high strength, which reduce friction and enhance engine efficiency and lifespan. Additionally, the lightweight nature of these composites contributes to overall vehicle performance and fuel efficiency. MMCs help improve thermal management within the engine, reducing the likelihood of overheating and increasing reliability. These composites are also used in car frames, suspension arms, and other structural parts, where their lightweight nature aids in reducing the vehicle's weight, thus improving fuel efficiency and handling [2], [167,168].

Aerospace Industry: In the aerospace sector, zinc-aluminum alloy-based MMCs are used in airframe structures, wing components, and fuselage parts. Their high strength-to-weight ratio is critical in ensuring structural integrity while minimizing weight, thus enhancing fuel efficiency and payload capacity. MMCs are used in turbine blades, compressor disks, and other engine parts where high-temperature stability and wear resistance are essential for maintaining engine performance and longevity under extreme operating conditions [2], [169,170].

Industrial Machinery: Zinc-aluminum alloy-based MMCs are ideal for high-load, low-speed applications due to their excellent wear resistance and low friction properties. They are used in industrial machinery such as heavy-duty presses, conveyor systems, and rotating equipment to extend service life and reduce maintenance requirements. These composites are used in manufacturing cutting tools like drills, milling cutters, and inserts. Their hardness and wear resistance extend tool life, maintain sharpness, and improve cutting performance, thereby enhancing productivity and reducing tool replacement costs [2], [171,172].

Electrical and Electronics: Due to their good thermal conductivity, zinc-aluminum alloy-based MMCs are used in heat sinks for electronic devices. They effectively dissipate heat from components such as CPUs, power transistors, and LEDs, ensuring reliable operation and preventing overheating. MMCs are used in connectors and housings for electronic devices, offering good mechanical strength and corrosion resistance. This ensures durability and reliability in various environments, from consumer electronics to industrial automation systems [2], [173,174].

Construction: Zinc-aluminum alloy-based MMCs are used in building frameworks, bridges, and other structural applications. Their high strength and durability make them suitable for load-bearing applications, reducing material fatigue and increasing the longevity of structures. These composites are used to reinforce concrete structures, providing additional strength and corrosion resistance. This is particularly useful in harsh environments, such as coastal areas or industrial zones, where exposure to corrosive elements is a concern [2], [175,176].

Marine Industry: In the marine industry, zinc-aluminum alloy-based MMCs are used for propellers and rudders due to their excellent corrosion resistance in saltwater environments. This enhances the performance and lifespan of these components, reducing maintenance costs and downtime. MMCs are also used in hull components, providing additional strength and durability to the vessel's structure. This helps improve the overall integrity and safety of the ship, especially in demanding maritime conditions [2], [177,178].

Consumer Goods: Zinc-aluminum alloy-based MMCs are used in the production of bicycles, golf clubs, and other sporting equipment. Their lightweight and high-strength properties enhance performance, providing athletes with superior equipment that improves speed, accuracy, and durability. These composites are utilized in parts of household appliances like washing machines and refrigerators. Their durability and corrosion resistance ensures long-lasting performance, reducing the need for frequent repairs and replacements [2], [179,180].

Medical Devices: Zinc-aluminum alloy-based MMCs are used in orthopaedic implants such as joint replacements and bone screws. Their biocompatibility, high strength, and corrosion resistance make them suitable for long-term implantation, ensuring patient safety and improving recovery outcomes [2], [181,182].

Military Applications: Zinc-aluminum alloy-based MMCs are used in body armour and vehicle armour due to their high strength-to-weight ratio. This provides effective protection while maintaining mobility and manoeuvrability for personnel and vehicles in combat situations. These composites are also used in various components of weapons systems, such as missile casings and gun barrels. Their high performance and reliability under extreme conditions are crucial for the effectiveness and safety of military operations[2], [183,184].

15. Conclusión:

This review examined the fabrication methods, microstructural features, mechanical behaviour, and physical performance of metal matrix composites (MMCs), with particular attention to the role of reinforcement materials and processing techniques. Zinc-aluminium (ZA) alloys, especially ZA-27, demonstrate promising mechanical properties such as high strength, wear resistance, and thermal stability, making them suitable for demanding applications in sectors like automotive and aerospace.

Among fabrication methods, liquid-state processing particularly stir casting remains the most widely used due to its simplicity and scalability. However, challenges such as non-uniform reinforcement distribution, porosity, and particle agglomeration continue to affect composite quality. Approaches such as improved stirring mechanisms, the use of wettability agents, and alternative methods like squeeze casting have shown potential in addressing these issues.

Solid-state processing techniques, including powder metallurgy and spark plasma sintering, offer enhanced control over microstructure and produce composites with superior mechanical properties. However, their higher cost and complexity currently limit their use in large-scale applications. Ceramic reinforcements such as silicon carbide (SiC) and titanium carbide (TiC) are particularly effective in improving strength, hardness, and wear resistance due to strong interfacial bonding with the matrix. Nonetheless, increasing reinforcement content can lead to brittleness and reduced ductility. Future research should focus on hybrid reinforcement strategies, optimized particle distribution, and improved interface engineering to achieve a better balance between strength and toughness. The inclusion of lightweight reinforcements like graphene also holds promise for achieving enhanced mechanical performance with reduced material density an important consideration in aerospace and other weight-sensitive applications.

Further work is needed to optimize both material compositions and processing methods to expand the applicability of MMCs across advanced engineering fields.

16. Future perspectives and research directions:

Metal matrix composites (MMCs) are gaining increasing attention due to their unique combination of properties excellent castability, high wear resistance, and strong mechanical performance. **Figure 5** outlines the key future directions and opportunities in MMC research and development:

16.1 Novel Reinforcements: Explore advanced materials like carbon nanotubes, graphene, and nanocellulose to enhance composite properties beyond traditional ceramics.

16.2 Advanced Manufacturing: Leverage technologies such as **additive manufacturing, powder metallurgy, spark plasma sintering, and severe plastic deformation (SPD)** to refine microstructure and production quality.

16.3 Process Optimization: Fine-tune parameters (temperature, pressure, time) for improved consistency and performance.

16.4 Characterization Techniques: Employ advanced tools (SEM, TEM, XRD, mechanical testing) for in-depth analysis of structure and properties.

16.5 **Machine Learning Integration:** Use AI models for predictive analysis, especially in **zinc-based composites**, to reduce experimental costs and time.

16.6 **Extreme Environment Testing:** Study behavior under high temperatures and corrosive conditions to expand MMC usability.

16.7 **Application Expansion:** Push into untapped sectors such as **biomedical, aerospace, and advanced automotive systems**.

16.8 **Sustainability Focus:** Assess environmental impacts and develop greener production methods.

The future of MMCs is promising, driven by innovations in materials, processing, and smart technologies. However, challenges like high cost, limited ductility, and microstructural inconsistencies must be addressed to fully realize their industrial potential.

References:

- [1] M. Ivey, G. W. Melenka, J. P. Carey, and C. Ayranci, "Characterizing short-fiber-reinforced composites produced using additive manufacturing," *Advanced Manufacturing: Polymer and Composites Science*, vol. 3, no. 3, pp. 81–91, Jul. 2017, doi: 10.1080/20550340.2017.1341125.
- [2] K. A. Sheikh and M. M. Khan, "Recent advances in zinc-based metal matrix composites: Fabrication, properties, and future prospects," *Journal of Alloys and Metallurgical Systems*, vol. 10, no. December 2024, p. 100173, 2025, doi: 10.1016/j.jalmes.2025.100173.
- [3] M. Ivey, G. W. Melenka, J. P. Carey, and C. Ayranci, "Characterizing short-fiber-reinforced composites produced using additive manufacturing," *Advanced Manufacturing: Polymer and Composites Science*, vol. 3, no. 3, pp. 81–91, 2017, doi: 10.1080/20550340.2017.1341125.
- [4] F. I. Azam, A. M. Abdul Rani, K. Altaf, T. V. V. L. N. Rao, and H. A. Zaharin, "An In-Depth Review on Direct Additive Manufacturing of Metals," *IOP Conf Ser Mater Sci Eng*, vol. 328, no. 1, 2018, doi: 10.1088/1757-899X/328/1/012005.
- [5] S. Tang, R. Ummethala, C. Suryanarayana, J. Eckert, K. G. Prashanth, and Z. Wang, "Additive Manufacturing of Aluminum-Based Metal Matrix Composites—A Review," 2021. doi: 10.1002/adem.202100053.
- [6] M. Dadkhah, M. H. Mosallanejad, L. Iuliano, and A. Saboori, "A Comprehensive Overview on the Latest Progress in the Additive Manufacturing of Metal Matrix Composites: Potential, Challenges, and Feasible Solutions," 2021. doi: 10.1007/s40195-021-01249-7.
- [7] C. Zhang, Z. Li, J. Zhang, H. Tang, and H. Wang, "Additive manufacturing of magnesium matrix composites: Comprehensive review of recent progress and research perspectives," 2023. doi: 10.1016/j.jma.2023.02.005.
- [8] Y. Karabulut and R. Ünal, "Additive manufacturing of ceramic particle-reinforced aluminum-based metal matrix composites: a review," 2022. doi: 10.1007/s10853-022-07850-0.
- [9] Z. Wang *et al.*, "Research Progress of Laser Additive Manufacturing Nickel-Based Alloy Metal Matrix Composites," 2023. doi: 10.3390/met13010129.
- [10] N. Li *et al.*, "Laser Additive Manufacturing on Metal Matrix Composites: A Review," Dec. 01, 2021, *Springer*. doi: 10.1186/s10033-021-00554-7.
- [11] S. Lathabai, "Additive Manufacturing of Aluminium-Based Alloys and Composites," in *Fundamentals of Aluminium Metallurgy: Recent Advances*, 2018. doi: 10.1016/B978-0-08-102063-0.00002-3.
- [12] D. Zhang *et al.*, "Metal Alloys for Fusion-Based Additive Manufacturing," 2018. doi: 10.1002/adem.201700952.
- [13] K. Osipovich *et al.*, "Wire-Feed Electron Beam Additive Manufacturing: A Review," 2023. doi: 10.3390/met13020279.
- [14] B. S. Rao and T. B. Rao, "Mechanical and Tribological Properties of 3D printed Al-Si alloys and composites: a Review," 2022. doi: 10.1007/s12633-021-01340-9.
- [15] Z. Zhu *et al.*, "Recent progress on the additive manufacturing of aluminum alloys and aluminum matrix composites: Microstructure, properties, and applications," 2023. doi: 10.1016/j.ijmachtools.2023.104047.
- [16] H. Fan *et al.*, "A review of laser additive manufacturing (LAM) aluminum alloys: Methods, microstructures and mechanical properties," 2024. doi: 10.1016/j.optlastec.2024.110722.

- [17] M. Ostolaza, J. I. Arrizubieta, A. Lamikiz, S. Plaza, and N. Ortega, “Latest Developments to Manufacture Metal Matrix Composites and Functionally Graded Materials through AM: A State-of-the-Art Review,” 2023. doi: 10.3390/ma16041746.
- [18] B. Karaş, P. J. Smith, J. P. A. Fairclough, and K. Mumtaz, “Additive manufacturing of high density carbon fibre reinforced polymer composites,” *Addit Manuf*, vol. 58, Oct. 2022, doi: 10.1016/j.addma.2022.103044.
- [19] O. Olanrewaju, I. O. Oladele, and S. O. Adelani, “Recent advances in natural fiber reinforced metal/ceramic/polymer composites: An overview of the structure-property relationship for engineering applications,” *Hybrid Advances*, vol. 8, no. November 2024, p. 100378, 2025, doi: 10.1016/j.hybadv.2025.100378.
- [20] A. Bhowmik *et al.*, “Casting of particle reinforced metal matrix composite by liquid state fabrication method: A review,” *Results in Engineering*, vol. 24, no. October, p. 103152, 2024, doi: 10.1016/j.rineng.2024.103152.
- [21] A. Borisov, A. Shamshurin, M. Kovalev, A. Popovich, and V. Sufiiarov, “Binder Jetting Additive Manufacturing of Inconel 718/TiC Metal Matrix Composites: Influence of TiC Content on Processing, Microstructure, Mechanical and Tribological Properties,” 2024. doi: 10.3390/ma17205050.
- [22] S. Bai and J. Liu, “Empirical Investigation of Properties for Additive Manufactured Aluminum Metal Matrix Composites,” *Applied Mechanics*, vol. 5, no. 3, pp. 450–474, 2024, doi: 10.3390/applmech5030026.
- [23] T. Varmus, R. Konecna, and G. Nicoletto, “Influence of surface roughness on fatigue performace of additively manufactured metal matrix composite Al2024-RAM2 alloy,” *Transportation Research Procedia*, vol. 74, pp. 756–762, 2023, doi: 10.1016/j.trpro.2023.11.207.
- [24] S. C. Kundurti and A. Sharma, “Evaluation of Microstructural, Mechanical and Corrosion Behaviours of Laminated AA6061/AA7075 Metal Matrix Composites Build by Friction Stir Additive Manufacturing for Structural Applications,” *Materials Research*, vol. 26, 2023, doi: 10.1590/1980-5373-MR-2023-0176.
- [25] Z. Wang *et al.*, “Research Progress of Laser Additive Manufacturing Nickel-Based Alloy Metal Matrix Composites,” *Metals (Basel)*, vol. 13, no. 1, pp. 1–17, 2023, doi: 10.3390/met13010129.
- [26] M. Dong, W. Zhou, S. Guo, and N. Nomura, “Ultrafine-bubble-water-promoted nanoceramic decoration of metal powders for additive manufacturing,” *NPG Asia Mater*, vol. 15, no. 1, 2023, doi: 10.1038/s41427-023-00494-9.
- [27] B. Karaş, P. J. Smith, J. P. A. Fairclough, and K. Mumtaz, “Additive manufacturing of high density carbon fibre reinforced polymer composites,” *Addit Manuf*, vol. 58, no. April, pp. 1–10, 2022, doi: 10.1016/j.addma.2022.103044.
- [28] A. Rimkus, M. M. Farh, and V. Gribniak, “Continuously Reinforced Polymeric Composite for Additive Manufacturing—Development and Efficiency Analysis,” *Polymers (Basel)*, vol. 14, no. 17, 2022, doi: 10.3390/polym14173471.
- [29] Z. Wang, Z. Fang, Z. Xie, and D. E. Smith, “A Review on Microstructural Formations of Discontinuous Fiber-Reinforced Polymer Composites Prepared via Material Extrusion Additive Manufacturing: Fiber Orientation, Fiber Attrition, and Micro-Voids Distribution,” *Polymers (Basel)*, vol. 14, no. 22, 2022, doi: 10.3390/polym14224941.
- [30] M. Jiménez, L. Romero, I. A. Domínguez, M. D. M. Espinosa, and M. Domínguez, “Additive Manufacturing Technologies: An Overview about 3D Printing Methods and Future Prospects,” *Complexity*, vol. 2019, 2019, doi: 10.1155/2019/9656938.
- [31] N. W. Y. Omar, N. A. Shuaib, M. H. J. A. Hadi, and A. I. Azmi, “Mechanical properties of carbon and glass fibre reinforced composites produced by additive manufacturing: A short review,” *IOP Conf Ser Mater Sci Eng*, vol. 670, no. 1, 2019, doi: 10.1088/1757-899X/670/1/012020.
- [32] A. Gupta and M. Talha, “Recent development in modeling and analysis of functionally graded materials and structures,” *Compos Struct*, vol. 96, no. 3, pp. 93–132, 2015, doi: 10.1016/j.paerosci.2015.07.001.
- [33] N. Khan and A. Riccio, “A systematic review of design for additive manufacturing of aerospace lattice structures: Current trends and future directions,” *Progress in Aerospace Sciences*, vol. 149, no. April, p. 101021, 2024, doi: 10.1016/j.paerosci.2024.101021.
- [34] V. Boggarapu *et al.*, “State of the art in functionally graded materials,” *Compos Struct*, vol. 262, no. December 2020, p. 113596, 2021, doi: 10.1016/j.compstruct.2021.113596.

- [35] B. Saleh *et al.*, “30 Years of functionally graded materials: An overview of manufacturing methods, Applications and Future Challenges,” *Compos B Eng*, vol. 201, no. January, p. 108376, 2020, doi: 10.1016/j.compositesb.2020.108376.
- [36] P. Kumar, S. K. Sharma, and R. K. R. Singh, “Recent trends and future outlooks in manufacturing methods and applications of FGM: a comprehensive review,” *Materials and Manufacturing Processes*, vol. 38, no. 9, pp. 1033–1067, 2023, doi: 10.1080/10426914.2022.2075892.
- [37] A. Gupta, M. Talha, and B. N. Singh, “Vibration characteristics of functionally graded material plate with various boundary constraints using higher order shear deformation theory,” *Compos B Eng*, vol. 94, pp. 64–74, 2016, doi: 10.1016/j.compositesb.2016.03.006.
- [38] D. Singh and A. Gupta, “Influence of microstructural defects on vibration characteristics of sandwich double FGM layer under mixed boundary conditions,” *International Journal on Interactive Design and Manufacturing*, 2024, doi: 10.1007/s12008-023-01703-z.
- [39] D. Singh and A. Gupta, “Influence of geometric imperfections on the free vibrational response of the functionally graded material sandwich plates with circular cut-outs,” *Mater Today Proc*, no. xxxx, pp. 2–5, 2022, doi: 10.1016/j.matpr.2022.02.187.
- [40] D. Singh and A. Gupta, “Effect of circular and elliptical cutouts on the free vibration analysis of sandwich FGM plate under the elastic foundations,” *Mechanics Based Design of Structures and Machines*, vol. 0, no. 0, pp. 1–27, 2024, doi: 10.1080/15397734.2024.2408604.
- [41] D. Singh, V. Krishna, P. Vemulapalli, and A. Gupta, “Influence of circular cutouts and elastic foundation on vibration characteristics of porous sandwich FGM plates under hygro-thermal environments,” *Mechanics Based Design of Structures and Machines*, vol. 0, no. 0, pp. 1–27, 2022, doi: 10.1080/15397734.2022.2140164.
- [42] D. Singh, S. Rai, and A. Gupta, “Vibration analysis of sandwich functionally graded material plate with cut-outs using artificial neural network technique,” *Thin-Walled Structures*, vol. 202, no. March, p. 112072, 2024, doi: 10.1016/j.tws.2024.112072.
- [43] T. Q. Quan, V. M. Anh, and N. D. Duc, “Natural frequency analysis of sandwich plate with auxetic honeycomb core and CNTRC face sheets using analytical approach and artificial neural network,” *Aerosp Sci Technol*, vol. 144, no. November 2023, p. 108806, 2024, doi: 10.1016/j.ast.2023.108806.
- [44] V. Mahesh, V. Mahesh, and S. A. Ponnusami, “FEM-ANN approach to predict nonlinear pyro-coupled deflection of sandwich plates with agglomerated porous nanocomposite core and piezo-magneto-elastic facings in thermal environment,” 2024. doi: 10.1080/15376494.2023.2201927.
- [45] A. K. P, K. Swaminathan, N. Indu, and S. H, “A novel EFG meshless-ANN approach for static analysis of FGM plates based on the higher-order theory,” *Mechanics of Advanced Materials and Structures*, vol. 31, no. 25, pp. 6501–6517, Nov. 2024, doi: 10.1080/15376494.2023.2231459.
- [46] V. N. V. Hoang, P. Shi, L. Toledo, and H. Vu, “Free vibration analysis of nanocomposite plate in contact with fluid using a novel quasi-3D shear deformation theory and artificial neural network,” *Mechanics Based Design of Structures and Machines*, pp. 1–45, Jul. 2024, doi: 10.1080/15397734.2024.2373265.
- [47] T. Borkar and R. Banerjee, “The Technology and Applicability of Laser Additive Manufacturing for In Situ Metal Matrix Composites and Functionally Graded Materials,” in *Manufacturing Techniques for Materials*, 2019. doi: 10.1201/b22313-4.
- [48] U. A. Essien and S. Vaudreuil, “In-situ metal matrix composites development for additive manufacturing: a perspective,” *Journal of Achievements in Materials and Manufacturing Engineering*, vol. 111, no. 2, 2022, doi: 10.5604/01.3001.0015.9997.
- [49] R. Nunes, N. Vandermeiren, W. Verlinde, D. Boruah, R. Motte, and W. De Waele, “A benchmark of mechanical properties and operational parameters of different steel filler metals for wire arc additive manufacturing,” *International Journal of Advanced Manufacturing Technology*, vol. 127, no. 1–2, 2023, doi: 10.1007/s00170-023-11520-z.
- [50] L. Simoneau, A. Bois-Brochu, and C. Blais, “Tensile Properties of Built and Rebuilt/Repaired Specimens of 316L Stainless Steel Using Directed Energy Deposition,” *J Mater Eng Perform*, vol. 29, no. 9, 2020, doi: 10.1007/s11665-020-05087-z.

- [51] X. Wang *et al.*, “A review on the simulation of selective laser melting AlSi10Mg,” 2024. doi: 10.1016/j.optlastec.2023.110500.
- [52] J. C. Cheng *et al.*, “Effects of second-phase boron particles on impact response of aluminum,” *Materials Science and Engineering: A*, vol. 793, 2020, doi: 10.1016/j.msea.2020.139805.
- [53] P. Durai Murugan *et al.*, “A current state of metal additive manufacturing methods: A review,” *Mater Today Proc*, vol. 59, 2022, doi: 10.1016/j.matpr.2021.11.503.
- [54] D. Zhang *et al.*, “Author Correction: Additive manufacturing of ultrafine-grained high-strength titanium alloys (Nature, (2019), 576, 7785, (91-95), 10.1038/s41586-019-1783-1),” 2020. doi: 10.1038/s41586-020-2291-z.
- [55] D. Zhang *et al.*, “Additive manufacturing of ultrafine-grained high-strength titanium alloys,” *Nature*, vol. 576, no. 7785, 2019, doi: 10.1038/s41586-019-1783-1.
- [56] T. Varol, O. Güler, F. Yıldız, and S. Suresh Kumar, “Additive Manufacturing of Non-ferrous Metals,” in *Springer Tracts in Additive Manufacturing*, vol. Part F3251, 2024. doi: 10.1007/978-3-030-89401-6_5.
- [57] P. Moghimian, T. Poirié, J. Kroeger, F. Marion, and F. Larouche, “Emerging Functional Metals in Additive Manufacturing,” 2023. doi: 10.1002/adem.202201800.
- [58] S. Alkunte *et al.*, “Advancements and Challenges in Additively Manufactured Functionally Graded Materials: A Comprehensive Review,” 2024. doi: 10.3390/jmmp8010023.
- [59] S B Angadi, Madeva Nagaral, Nagaraj Namdev, Subbaraya Mohan Kumar, and Zeeshan Ali, “A review on constituents, applications and processing methods of metal matrix composites,” *International Journal of Science and Research Archive*, vol. 11, no. 1, 2024, doi: 10.30574/ijrsra.2024.11.1.0329.
- [60] L. Singh, S. Kumar, S. Raj, Shivam, and P. Badhani, “Aluminium Metal Matrix Composites: Manufacturing and Applications,” *IOP Conf Ser Mater Sci Eng*, vol. 1149, no. 1, 2021, doi: 10.1088/1757-899x/1149/1/012025.
- [61] P. Chakrapani and T. S. A. Suryakumari, “Mechanical properties of aluminium metal matrix composites-A review,” in *Materials Today: Proceedings*, 2020. doi: 10.1016/j.matpr.2020.09.247.
- [62] S. Suresh Kumar, S. Thirumalai Kumaran, G. Velmurugan, A. Perumal, S. Sekar, and M. Uthayakumar, “Physical and mechanical properties of various metal matrix composites: A review,” in *Materials Today: Proceedings*, 2021. doi: 10.1016/j.matpr.2021.07.354.
- [63] N. Kumar and S. Soren, “Selection of reinforcement for Al/Mg alloy metal matrix composites,” in *Materials Today: Proceedings*, 2020. doi: 10.1016/j.matpr.2019.11.238.
- [64] W. C. Harrigan, “Processing of Aluminum Metal-Matrix Composites,” in *Aluminum Science and Technology*, 2018. doi: 10.31399/asm.hb.v02a.a0006488.
- [65] P. K. Rohatgi, N. Gupta, and A. Daoud, “Synthesis and Processing of Cast Metal-Matrix Composites and Their Applications,” in *Casting*, 2018. doi: 10.31399/asm.hb.v15.a0005339.
- [66] J. M. Mistry and P. P. Gohil, “Research review of diversified reinforcement on aluminum metal matrix composites: Fabrication processes and mechanical characterization,” 2018. doi: 10.1515/secm-2016-0278.
- [67] S. K. Khatkar, “Hybrid magnesium matrix composites: A review of reinforcement philosophies, mechanical and tribological characteristics,” 2023. doi: 10.1515/rams-2022-0294.
- [68] V. K. Parikh, V. Patel, D. P. Pandya, and J. Andersson, “Current status on manufacturing routes to produce metal matrix composites: State-of-the-art,” 2023. doi: 10.1016/j.heliyon.2023.e13558.
- [69] S. A. Kareem, J. U. Anaele, E. O. Aikulola, T. A. Adewole, M. O. Bodunrin, and K. K. Alaneme, “Design and selection of metal matrix composites reinforced with high entropy alloys – Functionality appraisal and applicability in service: A critical review,” 2024. doi: 10.1016/j.jalmes.2024.100057.
- [70] Y. Wang and T. Monetta, “Systematic study of preparation technology, microstructure characteristics and mechanical behaviors for SiC particle-reinforced metal matrix composites,” 2023. doi: 10.1016/j.jmrt.2023.07.145.
- [71] P. K. Rohatgi, P. Ajay Kumar, N. M. Chelliah, and T. P. D. Rajan, “Solidification Processing of Cast Metal Matrix Composites Over the Last 50 Years and Opportunities for the Future,” *JOM*, vol. 72, no. 8, 2020, doi: 10.1007/s11837-020-04253-x.
- [72] K. M. Sree Manu, L. Ajay Raag, T. P. D. Rajan, M. Gupta, and B. C. Pai, “Liquid Metal Infiltration Processing of Metallic Composites: A Critical Review,” *Metallurgical and Materials Transactions B: Process Metallurgy and Materials Processing Science*, vol. 47, no. 5, 2016, doi: 10.1007/s11663-016-0751-5.

- [73] M. Meignanammoorthy and M. Ravichandran, "Synthesis of metal matrix composites via powder metallurgy route: A review," *Mechanics and Mechanical Engineering*, vol. 22, no. 1, 2018, doi: 10.2478/mme-2018-0007.
- [74] K. Singh *et al.*, "Paradigm of state-of-the-art CNT reinforced copper metal matrix composites: processing, characterizations, and applications," 2023. doi: 10.1016/j.jmrt.2023.05.083.
- [75] D. K. Sharma, D. Mahant, and G. Upadhyay, "Manufacturing of metal matrix composites: A state of review," in *Materials Today: Proceedings*, 2019. doi: 10.1016/j.matpr.2019.12.128.
- [76] R. A. Laghari, N. He, M. Jamil, M. I. Hussain, M. K. Gupta, and G. M. Krolczyk, "A State-of-the-Art Review on Recently Developed Sustainable and Green Cooling/Lubrication Technologies in Machining Metal Matrix Composites (MMCs)," 2023. doi: 10.1007/s40684-023-00521-8.
- [77] S. Venugopal, S. Baskar, and P. N. F. Samuel, "Study on the mechanical behaviour of metal matrix composites (MMC)," *Mater Today Proc*, 2023, doi: 10.1016/j.matpr.2023.01.208.
- [78] P. K. Rohatgi, M. Tabandeh-Khorshid, E. Omrani, M. R. Lovell, and P. L. Menezes, "Tribology of metal matrix composites," in *Tribology for Scientists and Engineers: From Basics to Advanced Concepts*, vol. 9781461419457, 2013. doi: 10.1007/978-1-4614-1945-7_8.
- [79] J. Mao *et al.*, "High-solid content, low viscosity and long-term stability Mn–Co–Ni–O ceramic slurry for the fabrication of free-standing ultrathin NTC sheets," *Ceram Int*, vol. 48, no. 19, 2022, doi: 10.1016/j.ceramint.2022.06.042.
- [80] T. Liu *et al.*, "Recent progress on corrosion mechanisms of graphene-reinforced metal matrix composites," *Nanotechnol Rev*, vol. 12, no. 1, 2023, doi: 10.1515/ntrev-2022-0566.
- [81] J. Kim, J. Lee, C. Jo, and C. Kang, "Development of low cost carbon fibers based on chlorinated polyvinyl chloride(CPVC) for automotive applications," *Mater Des*, vol. 204, 2021, doi: 10.1016/j.matdes.2021.109682.
- [82] D. Nedelcu and I. Carcea, "Technology for obtaining samples of layered composite materials with metallic matrix," *Metals and Materials International*, vol. 19, no. 1, 2013, doi: 10.1007/s12540-013-1017-2.
- [83] M. K. Sparavalo, "Wide-sense robust and stable in the large terminal control for van der Pol dynamics: Poincaré's-approach-based backstepping method," in *Procedia Engineering*, 2017. doi: 10.1016/j.proeng.2017.09.021.
- [84] V. P. Bhaurkar, A. Bongale, S. Kumar, and A. Bongale, "Fabrication and analysis of metal matrix composites: A state of the art review," in *AIP Conference Proceedings*, 2020. doi: 10.1063/5.0030391.
- [85] S. Siebeck, K. Roder, G. Wagner, and D. Nestler, "Influence of boron on the creep behavior and the microstructure of particle reinforced aluminum matrix composites," *Metals (Basel)*, vol. 8, no. 2, 2018, doi: 10.3390/met8020110.
- [86] T. Wang, T. Fan, D. Zhang, G. Zhang, and D. Xiong, "Preparation and thermal properties of metal matrix composites with wood-structure," *Mater Des*, vol. 29, no. 6, 2008, doi: 10.1016/j.matdes.2007.05.006.
- [87] L. Zhao, W. Zheng, Y. Hu, Q. Guo, and D. Zhang, "Heterostructured metal matrix composites for structural applications: a review," 2024. doi: 10.1007/s10853-023-09300-x.
- [88] C. Veiga, J. P. Davim, and A. J. R. Loureiro, "PROPERTIES AND APPLICATIONS OF TITANIUM ALLOYS: A BRIEF REVIEW," 2012.
- [89] G. R. Speich, A. J. Schwoeble, and W. C. Leslie, "Elastic constants of binary iron-base alloys," *Metallurgical Transactions*, vol. 3, no. 8, 1972, doi: 10.1007/BF02643211.
- [90] T. Waffenschmidt, M. von Hoegen, G. Zilligen, and D. Rudek, "Inverse estimation of mode I fracture energies for structural adhesives from impact wedge peel tests," *Int J Adhes Adhes*, vol. 129, 2024, doi: 10.1016/j.ijadhadh.2023.103562.
- [91] X. Yan, Y. Xia, H. B. Blum, and T. Gernay, "Elevated temperature material properties of advanced high strength steel alloys," *J Constr Steel Res*, vol. 174, 2020, doi: 10.1016/j.jcsr.2020.106299.
- [92] D. K. Rajak, P. H. Wagh, and E. Linul, "Manufacturing technologies of carbon/glass fiber-reinforced polymer composites and their properties: A review," 2021. doi: 10.3390/polym13213721.
- [93] S. C. Ligon, R. Liska, J. Stampfl, M. Gurr, and R. Mülhaupt, "Polymers for 3D Printing and Customized Additive Manufacturing," 2017. doi: 10.1021/acs.chemrev.7b00074.
- [94] A. Mahmood, F. Perveen, S. Chen, T. Akram, and A. Irfan, "Polymer Composites in 3D/4D Printing:

Materials, Advances, and Prospects,” 2024. doi: 10.3390/molecules29020319.

[95] Z. Jiang, B. Diggle, M. L. Tan, J. Viktorova, C. W. Bennett, and L. A. Connal, “Extrusion 3D Printing of Polymeric Materials with Advanced Properties,” 2020. doi: 10.1002/adv.202001379.

[96] N. Tuncer and A. Bose, “Solid-State Metal Additive Manufacturing: A Review,” *JOM*, vol. 72, no. 9, 2020, doi: 10.1007/s11837-020-04260-y.

[97] M. Yakout and M. A. Elbestawi, “Additive Manufacturing of Composite Materials : An Overview,” *6th International Conference on Virtual Machining Process Technology (VMPT)*, Montréal, no. May, 2017.

[98] D. Smith, “Vat Polymerization,” in *Additive Manufacturing Processes*, 2020. doi: 10.31399/asm.hb.v24.a0006553.

[99] G. Wang, S. Wang, X. Dong, Y. Zhang, and W. Shen, “Recent progress in additive manufacturing of ceramic dental restorations,” 2023. doi: 10.1016/j.jmrt.2023.07.257.

[100] B. A. Praveena, N. Lokesh, A. Buradi, N. Santhosh, B. L. Praveena, and R. Vignesh, “A comprehensive review of emerging additive manufacturing (3D printing technology): Methods, materials, applications, challenges, trends and future potential,” in *Materials Today: Proceedings*, 2022. doi: 10.1016/j.matpr.2021.11.059.

[101] R. S. Mahale *et al.*, “Sensor-based additive manufacturing technologies,” 2022. doi: 10.33263/BRIAC123.35133521.

[102] Y. Wang *et al.*, “Advanced manufacturing of high-speed steels: A critical review of the process design, microstructural evolution, and engineering performance,” 2023. doi: 10.1016/j.jmrt.2023.04.269.

[103] X. Kaikai, G. Yadong, and Z. Qiang, “Comparison of traditional processing and additive manufacturing technologies in various performance aspects: a review,” 2023. doi: 10.1007/s43452-023-00699-3.

[104] S. Li *et al.*, “The performance and functionalization of modified cementitious materials via nano titanium-dioxide: A review,” *Case Studies in Construction Materials*, vol. 19, 2023, doi: 10.1016/j.cscm.2023.e02414.

[105] J. Liu *et al.*, “Nano-Modified Titanium Implant Materials: A Way Toward Improved Antibacterial Properties,” 2020. doi: 10.3389/fbioe.2020.576969.

[106] R. G. Munro, “Material Properties of Titanium Diboride,” *J Res Natl Inst Stand Technol*, vol. 105, no. 5, 2000, doi: 10.6028/jres.105.057.

[107] Y. Sun, “The use of aluminum alloys in structures: Review and outlook,” *Structures*, vol. 57, 2023, doi: 10.1016/j.istruc.2023.105290.

[108] X. You *et al.*, “A review of research on aluminum alloy materials in structural engineering,” 2024. doi: 10.1016/j.dibe.2023.100319.

[109] E. Georgantzia, M. Gkantou, and G. S. Kamaris, “Aluminium alloys as structural material: A review of research,” 2021. doi: 10.1016/j.engstruct.2020.111372.

[110] F. Wang *et al.*, “Progress on Electronic and Optoelectronic Devices of 2D Layered Semiconducting Materials,” 2017. doi: 10.1002/sml.201604298.

[111] Y. Liu, T. Zhao, W. Ju, and S. Shi, “Materials discovery and design using machine learning,” 2017. doi: 10.1016/j.jmat.2017.08.002.

[112] T. Schenk, M. Pešić, S. Slesazeck, U. Schroeder, and T. Mikolajick, “Memory technology-A primer for material scientists,” 2020. doi: 10.1088/1361-6633/ab8f86.

[113] G. Obey, M. Adelaide, and R. Ramaraj, “Biochar derived from non-customized matamba fruit shell as an adsorbent for wastewater treatment,” *Journal of Bioresources and Bioproducts*, vol. 7, no. 2, 2022, doi: 10.1016/j.jobab.2021.12.001.

[114] J. Gebhardt and A. M. Rappe, “Big data approach for effective ionic radii,” *Comput Phys Commun*, vol. 237, 2019, doi: 10.1016/j.cpc.2018.11.014.

[115] A. Jain *et al.*, “Commentary: The materials project: A materials genome approach to accelerating materials innovation,” 2013. doi: 10.1063/1.4812323.

[116] T. Zhou, Z. Song, and K. Sundmacher, “Big Data Creates New Opportunities for Materials Research: A Review on Methods and Applications of Machine Learning for Materials Design,” 2019. doi: 10.1016/j.eng.2019.02.011.

[117] M. M. Vopson, “Fundamentals of multiferroic materials and their possible applications,” *Critical Reviews in*

Solid State and Materials Sciences, vol. 40, no. 4, 2015, doi: 10.1080/10408436.2014.992584.

- [118] Y. Juan, Y. Dai, Y. Yang, and J. Zhang, “Accelerating materials discovery using machine learning,” 2021. doi: 10.1016/j.jmst.2020.12.010.
- [119] M. Bechthold and J. C. Weaver, “Materials science and architecture,” 2017. doi: 10.1038/natrevmats.2017.82.
- [120] T. L. Lin *et al.*, “The exosomal secretomes of mesenchymal stem cells extracted via 3D-printed lithium-doped calcium silicate scaffolds promote osteochondral regeneration,” *Mater Today Bio*, vol. 22, Oct. 2023, doi: 10.1016/j.mtbio.2023.100728.
- [121] M. P. Behera, T. Dougherty, and S. Singamneni, “Conventional and additive manufacturing with metal matrix composites: A perspective,” in *Procedia Manufacturing*, Elsevier B.V., 2019, pp. 159–166. doi: 10.1016/j.promfg.2019.02.023.
- [122] N. W. Y. Omar, N. A. Shuaib, M. H. J. A. Hadi, and A. I. Azmi, “Mechanical properties of carbon and glass fibre reinforced composites produced by additive manufacturing: A short review,” in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing Ltd, Dec. 2019. doi: 10.1088/1757-899X/670/1/012020.
- [123] F. M. Monticeli, R. M. Neves, H. L. Ornaghi, and J. H. S. Almeida, “A systematic review on high-performance fiber-reinforced 3D printed thermoset composites,” 2021. doi: 10.1002/pc.26133.
- [124] A. Rimkus, M. M. Farh, and V. Gribniak, “Continuously Reinforced Polymeric Composite for Additive Manufacturing—Development and Efficiency Analysis,” *Polymers (Basel)*, vol. 14, no. 17, 2022, doi: 10.3390/polym14173471.
- [125] N. Nawafleh and E. Celik, “Additive manufacturing of short fiber reinforced thermoset composites with unprecedented mechanical performance,” *Addit Manuf*, vol. 33, 2020, doi: 10.1016/j.addma.2020.101109.
- [126] Z. Wang, Z. Fang, Z. Xie, and D. E. Smith, “A Review on Microstructural Formations of Discontinuous Fiber-Reinforced Polymer Composites Prepared via Material Extrusion Additive Manufacturing: Fiber Orientation, Fiber Attrition, and Micro-Voids Distribution,” 2022. doi: 10.3390/polym14224941.
- [127] P. Cheng, Y. Peng, K. Wang, A. Le Duigou, and S. Ahzi, “3D printing continuous natural fiber reinforced polymer composites: A review,” 2024. doi: 10.1002/pat.6242.
- [128] N. van de Werken, J. Hurley, P. Khanbolouki, A. N. Sarvestani, A. Y. Tamijani, and M. Tehrani, “Design considerations and modeling of fiber reinforced 3D printed parts,” *Compos B Eng*, vol. 160, 2019, doi: 10.1016/j.compositesb.2018.12.094.
- [129] G. Liu, Y. Xiong, and L. Zhou, “Additive manufacturing of continuous fiber reinforced polymer composites: Design opportunities and novel applications,” 2021. doi: 10.1016/j.coco.2021.100907.
- [130] R. R. Fernandes, N. van de Werken, P. Koirala, T. Yap, A. Y. Tamijani, and M. Tehrani, “Experimental investigation of additively manufactured continuous fiber reinforced composite parts with optimized topology and fiber paths,” *Addit Manuf*, vol. 44, 2021, doi: 10.1016/j.addma.2021.102056.
- [131] P. Parandoush and D. Lin, “A review on additive manufacturing of polymer-fiber composites,” 2017. doi: 10.1016/j.compstruct.2017.08.088.
- [132] R. R. Fernandes, A. Y. Tamijani, and M. Al-Haik, “Mechanical characterization of additively manufactured fiber-reinforced composites,” *Aerosp Sci Technol*, vol. 113, 2021, doi: 10.1016/j.ast.2021.106653.
- [133] D. Zindani and K. Kumar, “An insight into additive manufacturing of fiber reinforced polymer composite,” 2019. doi: 10.1016/j.ijlmm.2019.08.004.
- [134] P. K., M. M., and S. P. P., “Technologies in additive manufacturing for fiber reinforced composite materials: a review,” 2020. doi: 10.1016/j.coche.2020.01.001.
- [135] A. N. Dickson, J. N. Barry, K. A. McDonnell, and D. P. Dowling, “Fabrication of continuous carbon, glass and Kevlar fibre reinforced polymer composites using additive manufacturing,” *Addit Manuf*, vol. 16, 2017, doi: 10.1016/j.addma.2017.06.004.
- [136] V. K. Balla, K. H. Kate, J. Satyavolu, P. Singh, and J. G. D. Tadimeti, “Additive manufacturing of natural fiber reinforced polymer composites: Processing and prospects,” 2019. doi: 10.1016/j.compositesb.2019.106956.
- [137] G. D. Goh, Y. L. Yap, S. Agarwala, and W. Y. Yeong, “Recent Progress in Additive Manufacturing of Fiber Reinforced Polymer Composite,” 2019. doi: 10.1002/admt.201800271.

- [138] J. Justo, L. Távara, L. García-Guzmán, and F. París, “Characterization of 3D printed long fibre reinforced composites,” *Compos Struct*, vol. 185, 2018, doi: 10.1016/j.compstruct.2017.11.052.
- [139] N. van de Werken, H. Tekinalp, P. Khanbolouki, S. Ozcan, A. Williams, and M. Tehrani, “Additively manufactured carbon fiber-reinforced composites: State of the art and perspective,” 2020. doi: 10.1016/j.addma.2019.100962.

XML-based Android Malware Detection for Enhancing Security in Global Supply Chains

Ansh Shah¹, Bhavnish Nanda¹, Keyaba Gohil¹ Fenil Ramoliya¹ Dev Mehta¹,
Rajesh Gupta¹, Sudeep Tanwar¹, and Nileshkumar Patel¹
Department of Computer Science and Engineering, Institute of Technology, Nirma
University, Ahmedabad, India
22bcm007@nirmauni.ac.in, 22bcm011@nirmauni.ac.in,
21bce076@nirmauni.ac.in, 21bce244@nirmauni.ac.in,
22bcm015@nirmauni.ac.in, rajesh.gupta@nirmauni.ac.in,
sudeep.tanwar@nirmauni.ac.in, nileshkumar.patel@nirmauni.ac.in

Abstract. This paper addresses the significance of employing Machine Learning (ML) for malware detection by detecting the vulnerabilities present in Android devices to secure smart supply chain management systems. The advantages of Android devices for supply chain automation include a focus on inventory management, real-time monitoring, and enhanced communication to improve the performance and efficiency of the supply chain. For malware detection, the proposed framework makes use of the XGBoost, Random Forest (RF), and Decision Tree (DT) algorithms, emphasising precise forecasting. With a 96% detection accuracy, the proposed framework combines the XGBoost, RF, and DT algorithms to improve Android system security in smart supply chain infrastructure. Additionally, the Explainable Artificial Intelligence (XAI) technique, Shapley Additive exPlanations (SHAP), is used to better interpret the results from the ML model. The performance evaluation of the ML algorithms is done with respect to metrics accuracy, recall, precision and f1 score. The confusion matrix that is shown also highlights how well the model detects security issues. SHAP decision plots are also attached in the results section to represent the decision-making process of the ML algorithms.

Keywords: Maternal Health · Smart Supply Chain · Security · ML · XAI · SHAP · Malware Detection

1 Introduction

Android phones and tablets convert conventional working systems into digitized workouts. The Digital Solution has wider powers and much improved real-time tracking and communication of Data, which also increases operational agility and reactivity and improves decision-making. Android-based solutions can be implemented, which have much higher productivity, lower errors, and better customer service. They are especially useful when user-friendliness and scalability are at high levels. This provides their workforce with numerous tools to manage inventory, accelerate the response for order fulfillment and improve communication standards amongst supply chain partners. Collaborative robots, or co-bots, are also game changers in the warehouse that improve efficiency [3]. Additionally, Supply chain automation through Android smartphones is more than just a digital transformation. It's a revolution that enables companies to deliver products and services more efficiently, thereby increasing productivity and streamlining decision-making processes. Android smartphones, with their high-level surveillance capabilities, are the perfect tool for this transformation. When combined with other innovative technologies, such as collaborative robots, they point to a future where humans and machines work together as efficient equals [6]. Attacks on Android smartphones that render certain defense holes make them vulnerable [7]. Consequently, hackers can use malicious audio files to gain control of a device's camera and smartphone. Application-specific vulnerabilities allow attackers to eavesdrop on users' conversations and media data, reinforcing the benefits of deploying security defenses in Android smartphones [9]. To address the above security issues, ML-based solutions have been considered using static analysis methods to detect Android malware that can use its high performance and accuracy in Android malware detection [10].

Not only that, but the flood of trojans and rootkits shows how hard it is to keep an Android device secure. With the proliferation of mobile devices such as smartphones and tablet computers, which are being used for an increasing number of functions online (including logging into email or financial services), enabling root access on these kinds of hardware is a dangerous move that hackers may target many websites. The fact that the update is provided to users in such an uncontroversial manner suggests the high likelihood of downloading malicious apps from the Google Play store. Which leaves them exposed to security vulnerabilities [4]. To defend against these attacks, several sophisticated

systems such as DL-Droid have been proposed for identifying malware by using deep learning and other stateful input generation-based techniques on dynamic analysis of Android applications, which can manage to reach high detection accuracy and outperform traditional methods [2].

Various research has investigated the improvement of ML-based methods for detecting malware on Android devices [1]. These methods include a variety of different techniques for detecting, analyzing, and reducing malware risks that include both static and dynamic analytics methods, as well as utilizing numerous tools and frameworks [8]. Using these sophisticated advanced techniques, it becomes feasible to improve the security of Android devices used in supply chain management, ensuring the integrity and efficiency of operations while reducing potential cyber risks.

ML-based methods provide promising solutions for tackling security vulnerabilities in Android phones. ML algorithms may be trained to recognize patterns associated with malicious behavior, thereby identifying threats before they can exploit system flaws. ML algorithms may learn to distinguish between normal activities and potential security concerns by carefully observing large datasets of both benign and malicious behavior. In this work, we looked at several ML models to detect malware in Android systems. After training and evaluating many algorithms, it was determined that XGBoost and RF models produced equivalent results, indicating high accuracy in malware detection. The way by which these ensemble learning models detect complex patterns makes them a nice choice for cybersecurity applications. In addition, the "black box" nature of most ML-related algorithms is not conducive to understanding how decisions are taken, which is disadvantageous regarding the system's security. This is the case when EAI is most relevant. XAI methods allow us to interpret the characteristics and the mechanisms behind the decisions made by our system, increasing its prediction trustworthiness.

Our proposed solution integrates ML-based malware detection with XAI to create a robust and interpretable security system for Android devices. Objectives of the proposed work are mentioned below:

- _ Developed an Android malware detection system for supply chain management, from development to distribution, addressing the critical need for security in the Android ecosystem.
- _ Integrated ML Algorithms in the supply chain system for Android malware detection. Also, we compared multiple ML models (including XGBoost and RF) for malware detection. Finally, the best-performing model was integrated into a real-time monitoring system with an alert mechanism, including email notifications for immediate threat response.
- _ Addressed complex ML models' "black box" nature by implementing XAI techniques. We utilized XAI to analyze and explain key features contributing to malware detection decisions, enhancing interpretability and trust in the system's alerts. The integration of XAI provided interpretable results alongside alerts, clearly explaining why specific apps or activities were flagged as suspicious.

2 Proposed Framework

2.1 Data Collection

The data collection layer is the first layer of our proposed model. It carries fine-grained details for system operations and behaviors of the Android layer. The dataset used [5] has various fields like File access patterns, network traffic, system calls, etc. A key part of this layer involves the disciplined feature selection to detect these salient signals that could potentially indicate incoming Android attacks. The explicit nature of data collection and feature selection ensures that our model avails itself of the best possible and most powerful hints to identify a security issue.

2.2 Intelligent

Our model can be analyzed efficiently by utilizing this layer layer. This process begins with a detailed dataset description, where we explore things such as timestamps and file locations/rights, etc., to understand the nature of our data.

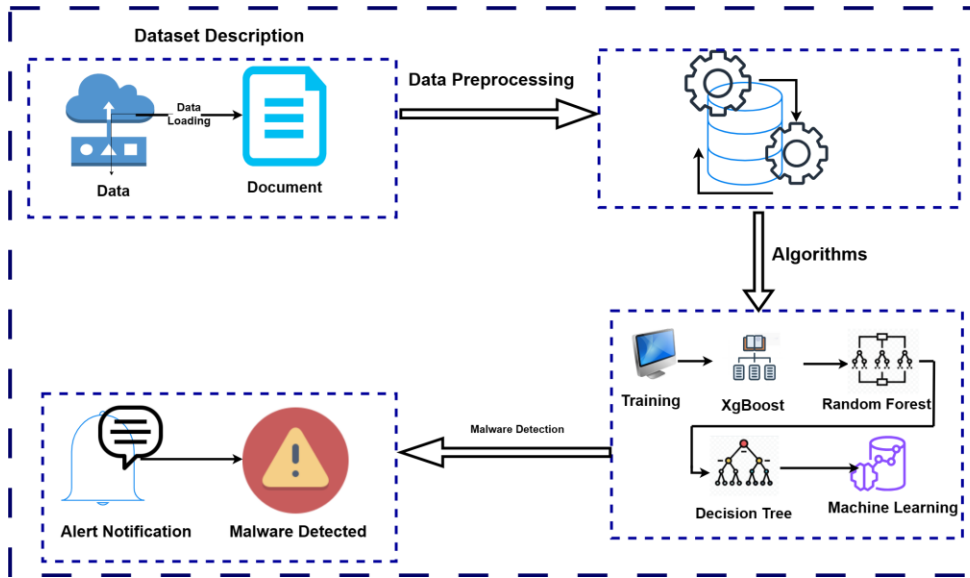


Figure 1

Next, we preprocess the data to sanitize and shape it into a format suitable for analysis by doing things like missing value treatment, normalization, etc. This layer is based on three strong algorithms: XGBoost, RF, and DT. These algorithms are given our pre-processed data to measure the different security weaknesses. Through extensive testing and training on relevant datasets, our model demonstrates an impressive 96% detection accuracy, with XGBoost emerging as the standout performer in identifying Android attacks.

XGBoost is a gradient-boosting framework that uses DTs. The model minimizes

$$\mathfrak{J} = \sum_{i=1}^n l(y_i, y'_i) + \sum_{k=1}^K \Omega(f_k)$$

where l is a differentiable convex loss function that measures the difference between the prediction $\hat{y}_i$ and the target y_i and Ω penalizes the complexity of the model. The regularization term Ω is defined as:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2$$

where T is the number of leaves in the tree, w is the score on each leaf, and γ and λ are regularization parameters. The final prediction is the sum of the initial prediction and the predictions from all K trees:

$$y_i = \sum_{k=1}^K f_k(x_i)$$

2.3 Explainable AI

The XAI layer enhances the interpretability and user interaction of our model. A key component of this layer is the use of SHAP (Shapley Additive exPlanations) values to visualize the model's predictions, particularly for the best-performing XGBoost algorithm. This visualization helps understand which features contribute most significantly to detecting potential attacks. Additionally, this layer incorporates an alert system responsible for informing users of possible security breaches. When any of the algorithms detect suspicious activity, the alert layer sends notifications or alarms, enabling rapid response and mitigation measures. This combination of XAI techniques and a responsive alert system ensures that our model not only detects threats effectively but also communicates these findings in an understandable and actionable manner.

3 Results and Discussion

3.1 Intelligent Layer Results

Fig. 2a presents the performance of the XGBoost Classifier model. The model accurately identi_ed 361 true negatives and 499 true positives. However, it misclassi_ed 16 instances as false positives and 17 as false negatives. These results demonstrate the model's strong predictive capability, with a relatively low error rate. The RF model's performance is illustrated in Fig. 2b. This model correctly classi_ed 358 true negatives and 500 true positives. It showed a slight increase in false positives (19) compared to XGBoost, while maintaining a similar number of false negatives (16). The RF model's performance is comparable to XGBoost, with a marginal di_erence in overall accuracy. Fig. 2c depicts the DT model's results. This model correctly identi_ed 349 true negatives and 493 true positives. It exhibited a higher number of misclassi_cations compared to the other two models, with 28 false positives and 23 false negatives. While still performing well, the DT model shows a slight decrease in accuracy compared to XGBoost and RF.

The Receiver Operating Characteristic (ROC) curve, displayed in Fig. 3a, provides a graphical representation of the performance of the RF, DT, and XGBoost classi_cation models across various threshold settings. The Area Under the Curve (AUC) serves as a comprehensive metric for model performance, with higher values indicating superior classi_cation capability. The ROC curve plays a crucial role in our methodology by o_ering a quantitative means of evaluating and comparing classi_er performance. It aids in model selection, threshold optimization, and overall validation of our approach to malware detection. By analyzing the ROC curve, we can determine the optimal threshold that minimizes false alarms while maintaining a high detection rate. The AUC value provides a single scalar metric to objectively compare di_erent models, facilitating the selection of the most e_ective classi_er for our task.

Fig. 3b presents a comparative analysis of precision, recall, and F1 score for the RF, DT, and XGBoost models. Precision quanti_es the proportion of true positives relative to the total of true and false positives, while recall measures the proportion of true positives relative to the total of true positives and false negatives. The F1 score, being the harmonic mean of precision and recall, o_ers a balanced assessment that e_ectively incorporates both metrics. This comparison

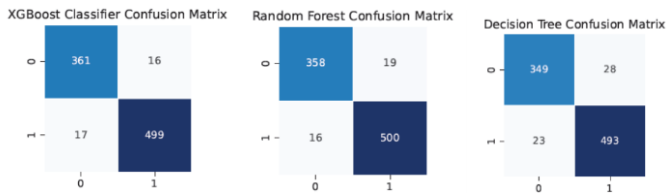


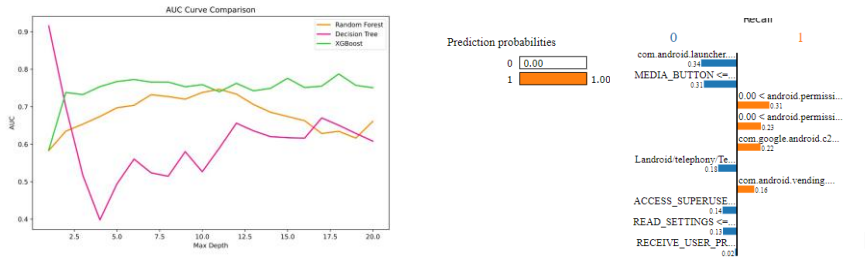
Fig. 2 Confusion matrices of (a) XG Boost, (b) RF, and (c) DT

allows for a comprehensive evaluation of each model's performance across these critical classi_cation metrics. The validation curve signi_cantly impacts our proposed methodology by enabling hyperparameter tuning and visualizing the bias-variance trade-o_ for each model. It is instrumental in selecting optimal parameter values, ensuring that the models are neither over_t nor under_t. This tool is essential for optimizing and comparing the performance of RF, DT, and XGBoost classi_ers in our malware detection system.

The Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) is a widely used metric in machine learning and statistics, particularly for classi_cation tasks. It assesses the model's performance across all classi_cation thresholds, signifying the model's ability to prioritize positive instances over negative ones. The AUC is commonly employed in cross-validation to identify the best-performing model. Fig. 4a demonstrates the variation in the ROC and AUC curves among the three models for di_erent maximum depth hyperparameter values. The ROC curve illustrates the relationship between the true positive rate (recall) and the false positive rate at various threshold

settings. The AUC summarizes the overall model performance across di_erent thresholds, with higher values indicating superior performance.

The AUC curve signi_cantly in_uences our methodology by providing a comprehensive, threshold-agnostic measurement for evaluating classi_er performance



across various values. It enables a direct assessment of the discriminative capabilities of the three models, which is particularly bene_cial when dealing with potentially imbalanced datasets in malware detection. Utilizing the AUC curve assists in model selection and hyperparameter optimization, ensuring strong performance across di_erent classi_cation thresholds.

3.2 X-AI Results

X-AI results are further discussed in this section where we used mainly 2 algorithms, SHAP and LIME. SHAP is a uni_ed framework for interpreting the predictions of machine learning models. It draws on game theory, speci_cally Shapley values, to allocate importance to each feature based on its contribution to a particular model prediction. The use of SHAP provides critical insights for various purposes, including identifying which features have the greatest impact on model predictions, meeting regulatory standards for explainability, and providing transparency regarding why a model made speci_c decisions.

The results presented below are based on the XGBoost algorithm as that is best performing algorithm from based upon ML implementation. XGBoost assigns prediction probabilities to di_erent features, giving insight into their importance within the model. Fig. 4b illustrates the prediction probabilities for di_erent output classes in XGBoost. The bar chart shows the range of probabilities on the x-axis (from 0 to 1) for various classes listed on the y-axis. indent Fig. 5 presents a tabular view of the features, showing their respective importance values. These values range from 0.0 (least important) to 1.0 (most important), highlighting the signi_cance of each feature within the model's decision- making process.

The analysis reveals the signi_cance of various permissions and features utilized by mobile applications, with speci_c focus on those relevant to malware detection from Fig. 4b and Fig. 5. For instance, the Home Screen Shortcuts permission allows apps to add shortcuts to the device's home screen, which could be used for both legitimate and malicious purposes. Access to Phone State is particularly important, as it grants access to crucial phone information

Feature	Value
com.android.launcher.permission.INSTALL_SHORTCUT	0.00
MEDIA_BUTTON	0.00
android.permission.READ_PHONE_STATE	1.00
android.permission.RECEIVE_BOOT_COMPLETED	1.00
com.google.android.c2dm.permission.RECEIVE	0.00
Landroid/telephony/TelephonyManager;->getLine1Number	0.00
com.android.vending.BILLING	0.00
ACCESS_SUPERUSER	0.00
READ_SETTINGS	0.00
RECEIVE_USER_PRESENT	0.00

such as network status and device identi_ers, making it essential for identifying malware attempting to gather sensitive data. On the other hand, ****In-app Purchases**** is found to have minimal relevance in this context, with a negligible importance score. Some permissions, such as ACCESS_MTK_MMHW, remain device-speci_c and ambiguous in their impact.

The Boot Completion Access permission, however, plays a key role, allowing an app to launch automatically after the device reboots, which is often a marker for persistent malware. Similarly, the Push Noti_cations permission, which allows apps to receive noti_cations, could be exploited by malware to deliver harmful content. The Media Control permission enables apps to interact with media buttons, a function that could be used maliciously to intercept or control device actions.

Other permissions, such as Location Access, which retrieves the last known location of the device, and Phone Number Retrieval, which allows access to the device's phone number, are often exploited for tracking or identity theft. Meanwhile, System Settings Access gives an app the ability to read and potentially manipulate sensitive system settings. Finally, User Presence Detection noti_es the app when the device is unlocked, enabling potential malicious activity timed to user interactions. Together, these permissions and features form a critical part of the malware detection framework, in_uencing the model's ability to classify and predict malicious behavior. Among these, the most signi_cant features, as determined by the XGBoost model, are READ_PHONE_STATE and RECEIVE_BOOT_COMPLETED, each with an importance score of 1.00. These features play a critical role in the model's prediction process.

4 Conclusion

This study highlights the critical role of X-AI in improving security for smart supply chain management systems, particularly by detecting vulnerabilities in Android devices. With the use of XGBoost, RF, and DT, the proposed framework achieved a malware detection accuracy of 97%. While Android devices enable e_cient communication and real-time monitoring in supply chains, their vulnerabilities, such as malicious apps, underscore the need for stronger security measures. The _ndings contribute to enhancing supply chain security by reinforcing the e_ectiveness of ML-based malware detection. Future research should focus on developing more advanced detection models, expanding their application across di_erent devices, and exploring blockchain integration for more comprehensive protection.

References

1. Alzaylaee, M.K., Yerima, S.Y., Sezer, S.: Emulator vs real phone: Android malware detection using machine learning. In: Proceedings of the 3rd ACM on International Workshop on Security and Privacy Analytics. pp. 65_72 (2017)

2. Alzaylaee, M.K., Yerima, S.Y., Sezer, S.: Dl-droid: Deep learning based android malware detection using real devices. *Computers & Security* 89, 101663 (2020)
3. Cha, S.Y.: The art of cyber security in the age of the digital supply chain: detecting and defending against vulnerabilities in your supply chain. In: *The Digital Supply Chain*, pp. 215_233. Elsevier (2022)
4. Duan, R., Alrawi, O., Kasturi, R.P., Elder, R., Saltaformaggio, B., Lee, W.: Towards measuring supply chain attacks on package managers for interpreted languages. *arXiv preprint arXiv:2002.01139* (2020)
5. <https://www.kaggle.com/datasets/zunxhisamniea/cyber-threat-data-for-new-malware-attacks>
6. Ismail, S., Moudoud, H., Dawoud, D., Reza, H.: Blockchain-based zero trust supply chain security integrated with deep reinforcement learning (2024)
7. Lamba, A., Singh, S., Balvinder, S., Dutta, N., Rela, S.: Analyzing and _xing cyber security threats for supply chain management. *International Journal For Technological Research In Engineering* 4(5) (2017)
8. Milosevic, N., Dehghantanha, A., Choo, K.K.R.: Machine learning aided android malware classi_cation. *Computers & Electrical Engineering* 61, 266_274 (2017)
9. Reuben, J., Ware, N.: Approach to handling cyber security risks in supply chain of defence sector. *Industrial Engineering Journal* 12(7), 1_12 (2019)
10. Sahs, J., Khan, L.: A machine learning approach to android malware detection. In: *2012 European Intelligence and Security Informatics Conference*. IEEE (2012)

VGG19-Based Multi-Modal Feature Fusion for Lung Disease Classification: Performance Comparison of Early, Late, and Intermediate Fusion Strategies

Gautami Shingan^{1,2}, Priya Ranjan^{3,4}, and Sachin Chaudhary¹

¹University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India

²Rajarambapu Institute of Technology, Sakharale., 415414, Maharashtra, India

³Kabir Vishwavidyalaya, Dasiya, Basti, 272150, Uttar Pradesh, India

⁴Buddha World University, Vaishali, 844128, Bihar, India.

gautamiphd@gmail.com

Abstract. Multimodal learning has emerged as a powerful strategy in medical image analysis, enabling the integration of complementary information from different imaging modalities. This study presents a deep learning-based framework for the classification of COVID-19 using a pretrained VGG19 model applied to both Computed Tomography (CT) and Chest X-ray (CXR) images. A comparative analysis is conducted across multiple configurations, including single-modality models (CTonly and CXR-only), a simple mixed-modality approach, and three fusion

strategies—early fusion, intermediate fusion, and late fusion—for effective multi-modal integration. The study evaluates model performance using an imbalanced dataset to simulate real-world clinical conditions. To investigate the effect of data diversity, three augmentation strategies—no augmentation, partial augmentation, and full augmentation—are applied. Experimental results reveal that intermediate fusion, particularly when combined with partial

augmentation, consistently outperforms all other configurations, including single-modality, mixed-modality, and both early and late fusion methods. These findings underscore the potential of the intermediate fusion technique in enhancing diagnostic accuracy for COVID-19, especially in the presence of class imbalance.

Keywords: Multimodal Fusion · VGG19 · Medical Image Analysis · Lung Disease Detection.

1 Introduction

In early December 2019, COVID-19 was a worldwide epidemic triggered by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). The epidemic's impacts on economic and social aspects are unprecedented. As marked by nearly 6 million fatalities and more than 375 million confirmed infections. Alarming is a continuous rise in the worldwide count of infected individuals, with no discernible signs of abatement. Addressing the challenges of the COVID-19 pandemic requires the expansion of robust early diagnosis and intervention, coupled with effective mitigation measures. Consequently, the formulation of early detection approaches relies on the accurate identification and observation of these viruses in the spatio-temporal evolution [3].

Medical imaging and diagnostic modalities are categorized into two primary approaches [4]: The Unimodal and Multimodal approaches. The Unimodal approach relies on a single type of information, such as Ultrasound, MRI, CT, CXR images, text data, or audio data [5]. While this approach is straightforward and cost-effective, its limitation lies in accessing a restricted set of information, potentially reducing diagnostic accuracy. Unimodal models, relying on a single modality, may lead to false positives or false negatives, posing clinical risks. Complex cases may challenge these approaches. In contrast, Multimodal approaches combine data from various sources, enhancing diagnostic accuracy. Integration of image data with audio, patient records, or different image modalities (e.g., CT+CXR, CXR+clinical text reports [6], CXR+cough[2]CT+Ultrasound, CT+CXR+Ultrasound [5]) improves disease detection, identifies abnormalities early, and informs effective treatment strategies. In this study, we diagnose COVID-19 using binary classification across CT and CXR modalities, analyzing performance in both single and multi-modal settings. This research contributes valuable insights into the potential of multi-modal approaches in enhancing disease detection models, particularly in the context of imbalanced datasets associated with COVID-19 diagnosis. The construction of the paper is summarized as follow: Section II discussed Mixed modalities and Fusion modalities and rationale of the recent work and addressing existing limitations. Section III describes the datasets

details and proposed system; Section IV highlights the impact of argumentation over the datasets using single and multi-modality approaches. Finally, Section V concludes the observations.

2 Literature Survey

A survey is carried out over two types of approaches such as Mixed modalities and Fusion modalities. Mixed modalities refer to the utilization of multiple types of imaging or data modalities within a study or analysis. (e.g., Chest X-ray and Computed Tomography) or additional non-imaging data (e.g., clinical information, patient history). Mixed modalities allow researchers to consider a broader range of information to enhance the understanding of a phenomenon or improve the performance of a model. On the other hand, In Fusion modalities involve the integration or combination of information from different modalities to gather various features associated with a particular disease. Fusion is a process where features or data from distinct modalities are merged to extract complementary information or improve overall performance. In the context of medical imaging, fusion modalities often aim to enhance diagnostic accuracy by combining the strengths of different imaging techniques. Fusion methods may include early, late, or intermediate fusion, depending on when data from distinct modalities are combined. The review of the study is explained in the next sections.

2.1 Mixed Modality

Researchers evaluated mentation techniques on VGG19 and ResNet50 models using the Cohen dataset, specifically the COVID-CT dataset [1]. In case of CT images, the VGG19 model achieved an accuracy of 84.87 %, outperforming the ResNet50 model, while, in the context of CXR images, the VGG19 model exhibited a remarkable accuracy of 99.35%, surpassing the ResNet50 model with an accuracy of 96.7%. These findings underscore the potential of VGG19 in handling CT images and highlight the overall strong performance of both models on CXR images. NasNetMobile[7] achieved an accuracy of 82.94% on CT images and 93.94% on CXR images. A transfer learning-based CNN model [8] was applied to the Cohen Dataset, achieving notable accuracies of 84.67% for X-ray and 98.78% for CT images. The proposed CNN architecture is highlighted as particularly effective for CT image datasets.

The study [9] utilized a mixed dataset of 1000 images, comprising 23 CT images and the rest from CXR modalities. Employing the VGG19 model with finetuning, the paper achieved a remarkable accuracy of 99%. [?] used a balanced small-size dataset containing 168 images in each class with total of 672 images. The proposed model is based on a tailored Deep Neural Network (DNN). Accuracy achieved by single modalities for CT and CXR is 95.83% and 96.13% respectively while mixed dataset CT and CXR together 96.28%. [10] the study employs CNNs on the COVIDx, Ozturk, and Ali Narin datasets, comprising 627 X-ray and 600 CT images. Despite class imbalances, CNN achieves impressive accuracies of 99.2% on CXR and 98.33% on CT scan images, with a combined accuracy of 99.5%. This multimodal approach enhances performance by 2%, The study showcases the model's potential in medical image analysis, emphasizing its success in handling diverse imaging modalities. Notably, no fusion techniques were applied in their study, emphasizing the effectiveness of standalone models. The main limitation across these studies is the varying degrees of class imbalances and the absence of fusion techniques in some, leaving room for future investigations into their potential contributions to further improve diagnostic accuracy.

2.2 Fusion Modalities

An ensemble model combining ResNet50 and VGG16 with late fusion [11] was employed on the SARS-CoV-2 CT scan dataset. Individually, VGG16 achieved an impressive accuracy with 98.93% in CXR, while ResNet50 achieved 98.00% in CT. The ensemble model achieved an overall accuracy of 99.8%, highlighting the effectiveness of multi-modality. MIDCAN model [12], incorporating a Constitutional Block Attention Module (CBAM) and employing late fusion, reported an accuracy of 98.02% with a small variation of $\pm 1.35\%$. CBAM allowed the AI model to focus on the lesion region, enhancing its ability to discern critical features. A CNN model for COVID-19 detection using CXR images sourced from Cohen and local hospitals [2], the model achieved an accuracy of 94.75%, and the utilization of multi-modality was noted to improve overall performance. Further in [13], CT/CXR images and clinical

notes are Integrated using pre-trained models (VGG16, ResNet50, InceptionResNetV2, Mo-bileNetV2) on the Cohen dataset yielded 97.8% accuracy. ResNet50 and Inception-ResNetV2 demonstrated comparable performance, showcasing the effectiveness of multimodal fusion for enhanced COVID-19 detection. In [4], COVID-RDNet demonstrated 93% accuracy in CT, 98% in CXR, and a notable improvement to 99% accuracy when combining CT and CXR, showcasing a 6% performance enhancement through multimodal fusion. In [6], CBP-MMFN and DHP-MMFN models fused chest X-rays and clinical text reports from KMC Hospital and Indiana University datasets. Unimodal datasets achieved 79.22% accuracy for images and 90.40% for text. Multimodal datasets, combining image and text, achieved a higher accuracy of 97.35%, highlighting the superiority of multimodal approaches.[5] proposed a multi-modal graph attention model embedded approach for COVID-19 detection, utilizing CT, CXR, ultrasound, and text data. The Multi-Modal COVID-19 Pneumonia Dataset achieved an accuracy of 98.10%. The clinical data and CXR images [14] were reintegrated using pre-trained CNN models on the MIMIC-IV dataset. With a dataset size of 289, late fusion achieved an accuracy of 88.39%, while intermediate fusion reached 90.34%. As the dataset expanded to 1156, late fusion improved to 89.74%, and intermediate fusion increased to 91.30%. The findings suggest that multimodal fusion enhances COVID-19 detection accuracy, with performance improvements observed as the dataset size increases. Researchers across various studies have explored the efficacy of multimodal fusion in COVID-19 detection. The above studies underscore the advantage of multimodal approaches, emphasizing the potential for improved accuracy and diagnostic performance in detecting COVID-19 when collecting different image modalities.

2.3 Limitations and research gaps

Based on the survey, several research gaps emerge when comparing single modality, mixed modality, and multi-modality fusion approaches in the context of COVID-19 detection and monitoring are:

- Single Modality Limitation: While single modality approaches have been widely used, they inherently limit the model's ability to capture the full spectrum of information required for high-precision COVID-19 detection. The reliance on a single data source, such as CT scans or CXR images, restricts the model's robustness and can lead to less reliable predictions. This gap indicates a need for exploring how integrating additional data sources might enhance diagnostic accuracy
- Mixed Modality Dominance: The majority of research has focused on mixed modality approaches, where multiple types of imaging or data modalities are utilized. These studies have demonstrated improved accuracy compared to single modality methods. However, even within mixed modality approaches, there is often a lack of comprehensive fusion strategies, particularly those that integrate non-imaging data such as audio or clinical information. This gap suggests that while mixed modality methods show promise, there is significant potential for improvement through more sophisticated integration techniques.
- Underexplored Multi-modality Fusion: Multi-modality fusion, which involves the integration of information from different modalities, has been recognized as a powerful approach to enhance model performance. However, the majority of existing work has concentrated on late and intermediate fusion techniques, where data from different modalities are combined at a later stage or after feature extraction. Early fusion, which involves combining data at the raw input level or during initial feature extraction, remains underexplored. This gap highlights the need for more research into early fusion techniques, which could potentially offer even greater improvements in model performance by capturing complementary features from multiple modalities at an earlier stage.
- Lack of Comprehensive Multimodal Frameworks: Despite the recognition of the benefits of multi-modality fusion, there is a notable absence of comprehensive frameworks that integrate multiple data modalities (such as audio, imaging, and clinical data) for COVID-19 detection. Most existing studies have focused on combining imaging modalities like CT and CXR, with few exploring the integration of non-imaging data. This research gap underscores the potential for developing more advanced multimodal frameworks that leverage a wider range of data types to achieve higher precision and better monitoring of COVID-19 symptoms. In conclusion, while significant advancements have been made in mixed modality and multimodal fusion approaches, substantial opportunities for innovation remain, particularly in the areas of early fusion and the integration of diverse data types within a unified

framework. Addressing these gaps, the proposed early fusion technique in this paper aims to enhance the accuracy and reliability of COVID-19 detection and monitoring systems. A detailed explanation of this technique is presented in the next section.

3 Proposed Method

This paper proposed an early fusion technique to extract features such as color features and texture features from CT images and stack over the CXR modality pass through the VGG19 pre-trained model with dropout layers and finally classify images into two categories such as COVID and Non-COVID classes. The detailed system such as dataset details, modified VGG19, and fusion approach are explained in this section.

3.1 Dataset Details

The COVID-19 X-ray and CT Scan Image dataset is employed for experimentation, accessible at the following link: <https://www.kaggle.com/khoongweihao/covid19-xray-dataset-train-test-sets>. It encompasses two classes, Non-COVID and COVID diseases, featuring a total of 17,099 images across both CXR and CT modalities. Specifically, the CXR modality comprises 5,500 non-COVID images and 4,044 COVID images, while the CT modality includes 2,628 non-COVID images and 5,427 COVID images.

3.2 Proposed System

The increasing reliance on medical imaging for the diagnosis of COVID-19 has spurred significant interest in the exploration of deep-learning models for automated image classification. In this study, we focus on the VGG19 pre-trained model, renowned for its effectiveness in feature extraction tasks. To enhance the dataset and tackle challenges related to image quality, uniformity, and class imbalance, we employ a range of preprocessing and augmentation techniques.

Pre-processing A comprehensive preprocessing pipeline was implemented to enhance the quality of input images before feeding them into the VGG19 model. This process included resizing, normalization, and data augmentation to improve consistency and robustness during training. All images were resized to a fixed dimension of 224×224 pixels, conforming to the input size requirement of the VGG19 architecture. This ensured dataset uniformity, enabling consistent feature extraction and reducing computational overhead during batch processing. Pixel intensities were normalized to the range $[-1, 1]$, a choice that aligns better with activation functions in deep neural networks. This normalization reduced pixel intensity variability across the dataset and contributed to stable and faster convergence by enabling more uniform weight updates.

To address class imbalance and enhance generalization, partial augmentation was applied separately to the CT and CXR modalities. This approach selectively augmented only the minority classes to match the majority, ensuring balance without significantly increasing dataset size. Techniques such as flips, $\bullet$ }15° rotations, shear, brightness and contrast adjustments, and Gaussian noise were used to simulate real-world variations while maintaining clinical relevance. This controlled augmentation improved class balance and exposed the model to diverse examples, leading to enhanced robustness and improved accuracy, precision, and recall.

4 Proposed Methodology

A deep learning framework based on a modified VGG19 model is proposed for lung disease classification using CT and CXR images. Single-modality, mixed-modality, and fusion-based configurations are explored to evaluate the impact of modality integration. Fusion is implemented at different stages—early, intermediate, and late—to assess how the point of integration influences classification performance.

4.1 Single-Modality and Mixed-Modality Classification

A modified VGG19 architecture was adopted as the baseline model for both single-modality and mixed-modality classification tasks. For single-modality experiments, the model was trained independently on CT and CXR datasets to evaluate each modality's diagnostic contribution. In the mixed-modality configuration, CT and CXR images were

combined into a single dataset and processed without applying any explicit fusion strategy. Although shared features across modalities could be captured, modality-specific distinctions were not preserved.

This configuration was used as a baseline for comparison with structured fusion methods. Figure 1 illustrates the architecture employed in both cases.

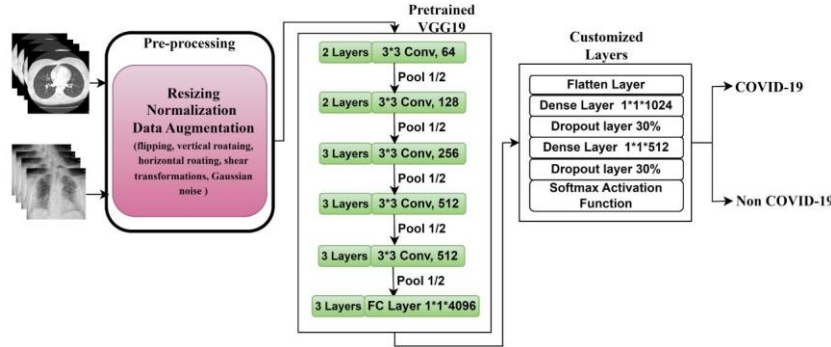


Fig. 1. Single and mixed modality setup: The same VGG19 architecture is used to classify either CT or CXR images independently (single modality), or a combined dataset of both without fusion (mixed modality).

4.2 Fusion Strategies for Multimodal Learning

To explore the integration of complementary features from CT and CXR images, three fusion strategies are employed: early fusion, intermediate fusion, and late fusion. Each strategy integrates modality-specific information at a different stage of the network pipeline—ranging from input-level combination to feature-level concatenation and decision-level aggregation. These fusion techniques are designed to assess the effectiveness of joint learning versus independent learning in multimodal setups. By systematically evaluating all three approaches, the study aims to identify the most suitable strategy for improving classification accuracy in a clinical context, particularly under the challenges posed by imbalanced datasets.

Early Fusion In the early fusion strategy, integration is performed at the input level by forming a composite image that incorporates both CT-derived features and CXR data. As illustrated in Figure 2, the CT image undergoes preprocessing that includes grayscale conversion, followed by the extraction of texture and color features. A Gabor filter is applied to the grayscale CT image to capture orientation- and scale-specific texture patterns. The Gabor function is defined

in Equation 1:

$$G(x, y, s\lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cdot \exp\left(i\left(2\pi\frac{x'}{\lambda} + \psi\right)\right)$$

where $x' = x \cos(\theta) + y \sin(\theta)$, $y' = -x \sin(\theta) + y \cos(\theta)$, and parameters λ , θ , ψ , σ , γ control frequency, orientation, phase, Gaussian width, and spatial aspect ratio respectively. These filters enhance fine textural features in CT images that may not be visible in CXR. Additionally, color information is extracted from the original CT image by averaging the intensity of the RGB channels for each pixel.

This process is defined mathematically in Equation 2:

$$ColourCT(x, y) = \frac{1}{3}(CT(x, y)_R + CT(x, y)_G + CT(x, y)_B)$$

The texture and color features from CT are then stacked with the CXR image to form a rich, multimodal input representation. Before feeding into the model, the stacked image is normalized to ensure consistent pixel intensity distributions. This fused input is passed into a customized VGG19 model, which jointly learns spatial, structural, and textural features across both modalities from the initial convolutional layers. By integrating low-level features early in the pipeline, this method supports more holistic learning—although it relies on consistent resolution and alignment between modalities, which may pose challenges in clinical datasets.

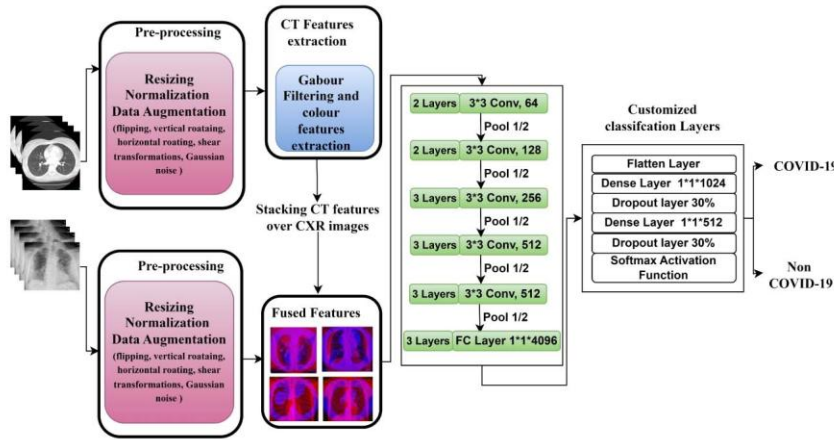


Fig. 2. Early fusion: CT features are stacked with CXR images and fed into a customized VGG19 model for joint feature learning and classification.

Intermediate Fusion Intermediate fusion separates the feature extraction and fusion stages by independently processing CT and CXR images through two parallel VGG19 networks shown in figure 3. Let the CT image input be denoted as I_{CT} and the CXR image input as I_{CXR} . Each modality is passed through a customized VGG19 network, truncated before the final classification layer, to extract high-level features. The outputs of the respective VGG19 models are denoted as:

$$f_{CT} = VCC18_{CT}(I_{CT})$$

$$f_{CXR} = VG19_{CXR}(I_{CXR})$$

where $f_{CT} \in \mathbb{R}^{d_1}$ and $f_{CXR} \in \mathbb{R}^{d_2}$ are the feature vectors derived from CT and CXR modalities respectively (Equations 3 and 4). These vectors are then concatenated to form a unified feature representation. As shown in Equation 5, the concatenated feature vector f_{fused} aggregates modality-specific information for joint reasoning. Finally, the fused feature vector is passed to a shared classification head consisting of fully connected layers and a softmax function.

$$f_{fused} = [f_{CT}; f_{CXR}] \in \mathbb{R}^{d_1 + d_2}$$

$$y = \text{Softmax}(W \cdot f_{fused} + b)$$

Equation 6 produces the class probability distribution $y \in \mathbb{R}^C$, where C is the number of output classes.

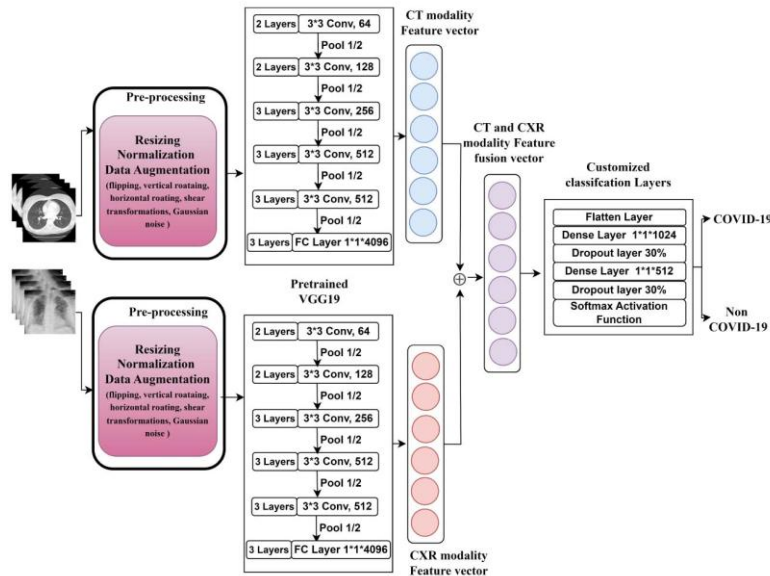


Fig 3 intermediate fusion CT and czr ffeatures extracted via VGG 19 are concatenated and passed to a classifier for final prediction.

As illustrated in Figure 3, this architecture enables each modality to independently learn its semantic features before fusion, ensuring preservation of both complementary and unique modality characteristics. This design consistently yielded superior performance compared to early and late fusion strategies, particularly when combined with partial augmentation.

Late Fusion Late fusion refers to a decision-level integration strategy in which independently trained models on different modalities—CT and CXR in this case—generate predictions separately, and these predictions are subsequently combined. This approach allows each model to specialize in extracting features unique to its input modality without interference during training. In the proposed framework shown in figure 4, late fusion is performed using weighted summation of the softmax output probabilities from each model. Let pCT and pCXR denote the softmax prediction vectors from the CT-based and CXR-based VGG19 models, respectively. The final class prediction pfinal is obtained by computing a linear combination of these vectors using modality-specific weights.

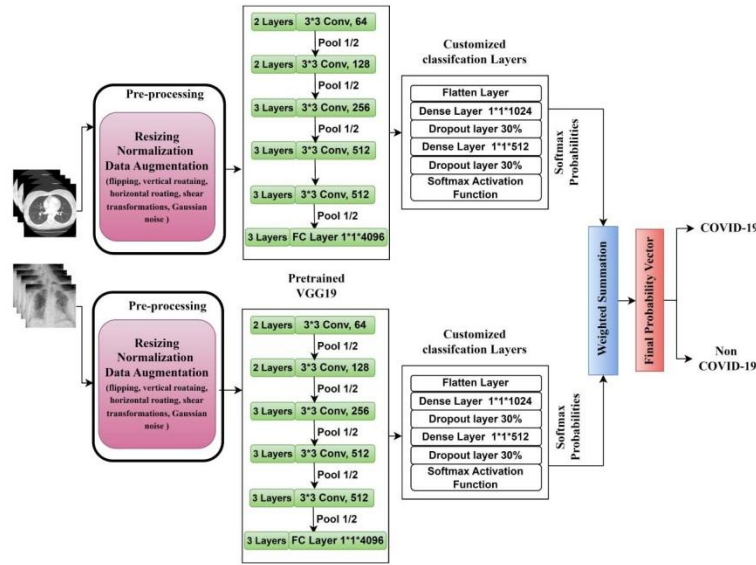


Fig. 4. Late fusion: Independent VGG19 models process CT and CXR images; their output probabilities are combined using a weighted summation for final prediction.

This is expressed in Equation 7. where wCT and wCXR are scalar weights assigned to the CT and CXR outputs. These weights satisfy the constraint shown in Equation 8:

$$p_{\text{final}} = w_{\text{CT}} \cdot p_{\text{CT}} + w_{\text{CXR}} \cdot p_{\text{CXR}}$$

$$w_{\text{CT}} + w_{\text{CXR}} = 1$$

Weight selection can be guided by several criteria. One common method is to base the weights on validation performance—assigning higher weights to models that demonstrate superior accuracy or F1-score on a held-out validation set. Alternatively, cross-validation or empirical tuning can be employed to find optimal weights that maximize overall ensemble performance. More advanced strategies involve dynamic adjustment based on prediction confidence or uncertainty estimation. The advantage of weighted summation lies in its flexibility to prioritize modality contributions depending on context. For example, if CT scans are more effective at capturing lung abnormalities specific to COVID-19, a higher wCT value can be used to reflect its diagnostic strength. Conversely,

in scenarios where CXR generalizes better, wCXR can be emphasized. As a result, the fusion mechanism is not only modular but also robust to variations in modality performance, providing a balanced and improved prediction capability over single-modality models.

5 Result analysis

Each fusion strategy was implemented with modality-specific constraints. In early fusion, CT and CXR images were stacked at the input level, requiring equal image counts per class and one-to-one pairing across modalities. For intermediate fusion, deep features were extracted separately from CT and CXR images using parallel VGG19 models and concatenated before classification—removing the need for strict pairing, though class balance was preserved. In late fusion, independently trained models produced class probability outputs, which were fused using a weighted summation approach at the decision level. Partial augmentation was applied across all strategies to mitigate class imbalance and improve generalization. This ensured fair and structurally compliant comparisons across all fusion approaches. Table 1 presents a comprehensive performance evaluation of various single-modality, mixed-modality, and multimodal fusion strategies— specifically Early Fusion, Intermediate Fusion, and Late Fusion—using a partially augmented dataset for binary classification (COVID vs. Non-COVID). The analysis spans across training and testing sets, reporting key metrics such as precision, recall, F1-score, support, and overall accuracy. Among the individual modalities, the CT-based model achieved higher accuracy 89.15% compared to the CXR-based model 87.94%, indicating that CXR images provided more discriminative features under the given augmentation scenario. However, combining modalities in a simple mixed-modality approach (i.e., stacking datasets without deep integration) resulted in testing accuracy 95.01% than CXR or CT modality alone. This suggests that mixed modality improves the performance of the model.

Table 1. Performance Analysis over Partial Augmented Dataset

	Classes	Precision	Recall	F1-Score	Support	Accuracy
		CT images				
Training	Set COVID	0.85	0.96	0.90	4341	89.95%
	Non-COVID	0.96	0.84	0.89	4400	
Testing Set	COVID	0.84	0.96	0.90	1086	89.15%
	Non-COVID	0.96	0.82	0.88	1100	
		CXR images				
Training Set	COVID	0.85	0.85	0.90	8683	89.05%
	Non-COVID	0.94	0.84	0.88	8800	
Testing Set	COVID	0.84	0.94	0.89	2171	87.94%
	Non-COVID	0.93	0.82	0.87	2200	
		Mixed Modality				
Training Set	COVID	0.96	0.96	0.96	4313	96.09%
	Non-COVID	0.96	0.96	0.96	4400	
Testing Set	COVID	0.95	0.95	0.95	1086	95.01%
	Non-COVID	0.95	0.95	0.95	1100	
		Early Fusion				
Training Set	COVID	0.96	0.98	0.97	4341	96.92%
	Non-COVID	0.98	0.96	0.97	4400	
Testing Set	COVID	0.95	0.96	0.95	1086	95.37%
	Non-COVID	0.96	0.95	0.95	1100	
		Intermediate Fusion				
Training Set	COVID	0.98	0.98	0.98	4341	98.15%
	Non-COVID	0.98	0.98	0.98	4400	
Testing Set	COVID	0.97	0.98	0.98	1086	97.27%
	Non-COVID	0.98	0.96	0.97	1100	
		Late Fusion				
Training Set	COVID	0.96	0.97	0.96	4341	96.20%
	Non-COVID	0.97	0.96	0.96	4400	
Testing Set	COVID	0.96	0.95	0.96	1086	95.11%
	Non-COVID	0.95	0.95	0.95	1100	

In the case of all three fusion strategies, Early Fusion, which stacks CTderived feature maps with CXR images at the input level, achieved 95.37% test accuracy. Late Fusion, where predictions from separate models are combined via weighted summation, reached 95.11% accuracy. Notably, Intermediate Fusion consistently outperformed all other methods with a training accuracy of 98.15% and a testing accuracy of 97.27%, along with the highest F1-scores for both classes. These results demonstrate the effectiveness of deeper semantic integration of modality-specific features before classification.

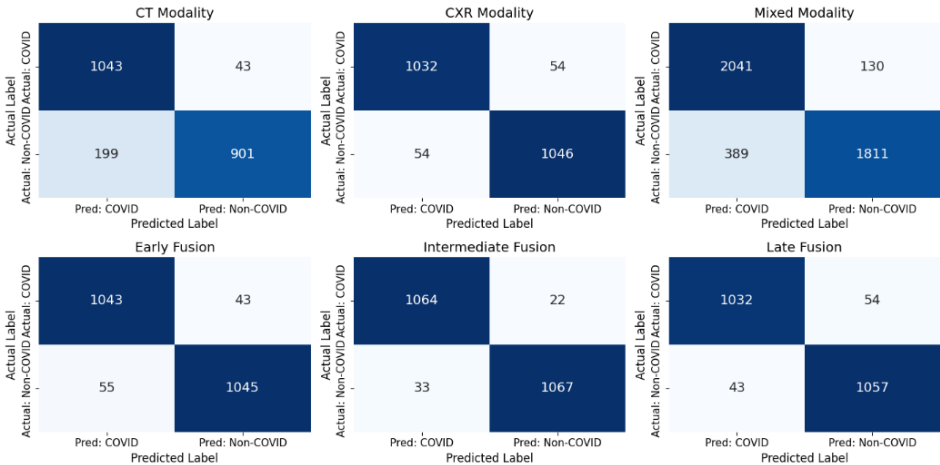


Fig. 5 Confusion matrices of testing data for all modalities CT, CXR, Mixed, Early Fusion, Intermediate Fusion, and Late Fusion

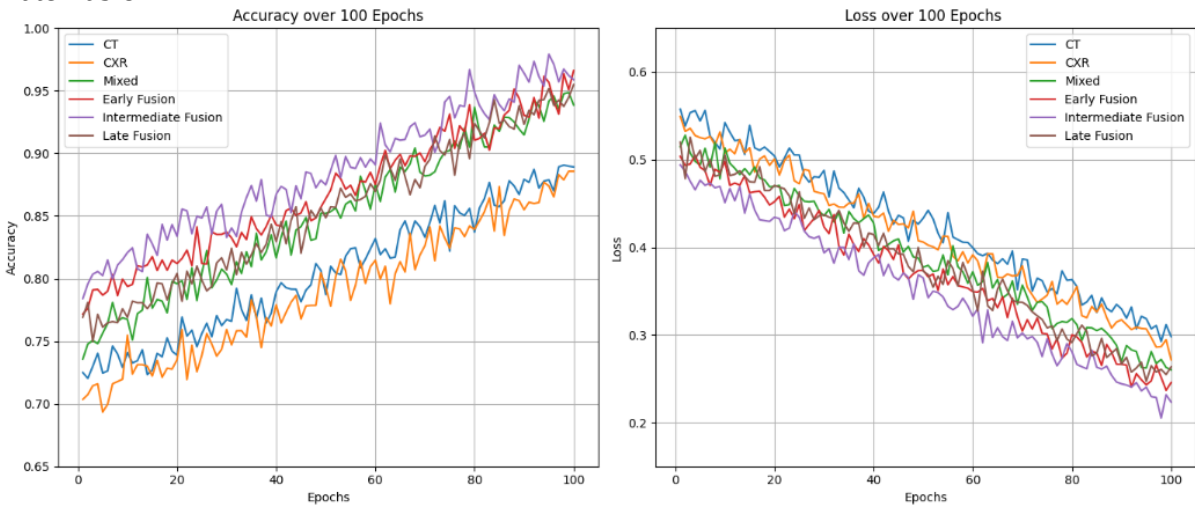


Fig. 6 Accuracy and loss curves for single modality, mixed-modality, and fusion approaches

The confusion matrices for all testing configurations shown in Figure 5, reveal that the intermediate fusion approach achieves the highest number of correctly classified COVID and Non-COVID cases. CT-based predictions emphasize recall, while CXR shows high precision; however, both single modalities underperform compared to fusion strategies. The mixed modality slightly improves performance by using more training data but lacks coordinated feature learning. Early fusion integrates modalities at the input level, but is sensitive to alignment inconsistencies. Late fusion combines prediction outputs, but suffers from weak interaction between features. In contrast, intermediate fusion extracts high-level features from CT and CXR independently and merges them effectively. This leads to better semantic representation and cross-modal synergy. Accuracy and loss trends are shown in figure 6 further support this, as intermediate fusion consistently reaches the highest accuracy 97.27%)and lowest loss. The stability of the learning curve in later epochs highlights strong convergence. Thus, intermediate fusion clearly outperforms all other configurations.

6 Conclusion

In this study, the effectiveness of multiple multimodal fusion strategies—Early Fusion, Intermediate Fusion, and Late Fusion—was systematically evaluated for the classification of COVID-19 using CT and CXR images. A VGG19-based architecture was employed across single-modality, mixed-modality, and fusionbased configurations. Comparative analysis revealed that intermediate fusion, which concatenates high-level features extracted from modality-specific branches, consistently outperformed all other approaches in terms of classification accuracy and F1-score.

While Early Fusion was found to be limited by its reliance on spatial alignment and uniform resolution, Late Fusion lacked the ability to model deep intermodality interactions. Intermediate Fusion, by contrast, preserved semantic richness from each modality and enabled effective integration at the feature level, resulting in superior diagnostic performance. These findings emphasize the significance of designing fusion mechanisms that leverage complementary modality information in medical imaging applications.

In future work, advanced attention-based fusion mechanisms—such as crossmodal transformers and modality-aware self-attention—will be explored to dynamically capture inter-modality relationships. These techniques are expected to enhance the adaptability and interpretability of multimodal models, particularly when applied to more diverse datasets or extended to multi-class diagnostic scenarios.

References

1. Zouch, Wassim and Sagga, Dhouha and Ectiou, Amira and Khemakhem, Rafik and Ghorbel, Mohamed and Mhiri, Chokri and Hamida, Ahmed Ben: Detection of COVID-19 from CT and chest X-ray images using deep learning models, *Annals of Biomedical Engineering*, 50(7), pp. 825–835, Springer, (2022)
2. Kumar, Santosh and Chaube, Mithilesh Kumar and Alsamhi, Saeed Hamood and Gupta, Sachin Kumar and Guizani, Mohsen and Gravina, Raffaele and Fortino, Giancarlo: A novel multimodal fusion framework for early diagnosis and accurate classification of COVID-19 patients using X-ray images and speech signal processing techniques. *Computer methods and programs in biomedicine*, 226, Elsevier (2022)
3. Hu, Fang and Huang, Mingfang and Sun, Jing and Zhang, Xiong and Liu, Jifen: An analysis model of diagnosis and treatment for COVID-19 pandemic based on medical information fusion, *Information Fusion*, 73, pp. 11–21, Elsevier, (2021)
4. Fang, Lingling and Wang, Xin: COVID-RDNet: A novel coronavirus pneumonia classification model using the mixed dataset by CT and X-rays images, *biocybernetics and biomedical engineering*, 42 (3), pp. 977–994, Elsevier, (2022)
5. Zheng, Wenbo and Yan, Lan and Gou, Chao and Zhang, Zhi-Cheng and Zhang, Jun Jason and Hu, Ming and Wang, Fei-Yue: Pay attention to doctor–patient dialogues: Multi-modal knowledge graph attention image-text embedding for COVID-19 diagnosis, *Information Fusion* 75, pp. 168–185, Elsevier, (2021)
6. Shetty, Shashank and Mahale, Ajit: Multimodal medical tensor fusion network-based DL framework for abnormality prediction from the radiology CXRs and clinical text reports, *Multimedia Tools and Applications* 82, pp. 44431–44478, Springer, (2021)
7. Ahsan, Md Manjurul and Gupta, Kishor Datta and Islam, Mohammad Maminur and Sen, Sajib and Rahman, Md Lutfar and Shakhawat Hossain, Mohammad: Covid- 19 symptoms detection based on nasnetmobile with explainable ai using various imaging modalities, 2(4), pp. 490–504. MDPI, (2020).
8. Kamil, Mohammed Y: A deep learning framework to detect Covid-19 disease via chest X-ray and CT scan images. *International Journal of Electrical & Computer Engineering* 11(1), (2021)
9. Mukherjee, Himadri and Ghosh, Subhankar and Dhar, Ankita and Obaidullah, Sk Md and Santosh, KC and Roy, Kaushik: Deep neural network to detect COVID-19: one architecture for both CT Scans and Chest X-rays, *Applied Intelligence*, 51, pp. 2777–2789, Springer, (2021).

10. Goyal, Lakshay and Dhull, Anuradha and Singh, Akansha and Kukreja, Sonal and Singh, Krishna Kant:VGG-COVIDNet: A Novel model for COVID detection from XRay and CT Scan images,Procedia computer science, 218, pp. 1926–1935,Elsevier, (2023).
11. Hilmizen, Naufal and Bustamam, Alhadi and Sarwinda, Devvi: The multimodal deep learning for diagnosing COVID-19 pneumonia from chest CT-scan and X-ray images, 2020 3rd international seminar on research of information technology and intelligent systems (ISRITI), pp. 26–31, IEEE, 2020.
12. Zhang, Yu-Dong and Zhang, Zheng and Zhang, Xin and Wang, Shui-Hua: MIDCAN: A multiple input deep convolutional attention network for Covid-19 diagnosis based on chest CT and chest X-ray, Pattern recognition letters 150, pp. 8–16, Elsevier, (2021).
13. Nasir, Nida and Kansal, Afreen and Barneih, Feras and Al-Shaltone, Omar and Bonny, Talal and Al-Shabi, Mohammad and Al Shammaa, Ahmed: Multi-modal image classification of COVID-19 cases using computed tomography and X-rays scans, Intelligent Systems with Applications, 17, pp. 200160, Elsevier, (2023).
14. Kumar, Sachin and Ivanova, Olga and Melyokhin, Artyom and Tiwari, Prayag:Deep-learning-enabled multimodal data fusion for lung disease classification, Informatics in Medicine Unlocked, 42,pp. 101367, Elsevier, (2023).
15. Fu, Minghan and Wu, Fang-Xiang:QLABGrad: A Hyperparameter-Free and Convergence-Guaranteed Scheme for Deep Learning,Proceedings of the AAAI Conference on Artificial Intelligence, 38(11), pp.12072–12081, (2024).
16. Li, Xiang and Zhang, Wei and Ding, Qian and Sun, Jian-Qiao:Intelligent rotating machinery fault diagnosis based on deep learning using data augmentation, Journal of Intelligent Manufacturing,31(2), pp. 433–452, Springer,(2020).
17. Wanasinghe, T., Bandara, S., Madusanka, S., Meedeniya, D., Bandara, M., de la Torre Díez, I.: Lung sound classification with multi-feature integration utilizing lightweight CNN model. IEEE Access (2024)
18. Gaikwad, Santosh and Gawali, Bharti and Desai, Chitra: Discrete Wavelet Transformation based Multimodal Medical Image Fusion for Disease Identification, Power System Technology, 48(1), pp. 620–630, (2024).
19. Shetty, S., Mahale, A.: Multimodal medical tensor fusion network-based DL framework for abnormality prediction from the radiology CXRs and clinical text reports. Multimedia Tools and Applications 82(28), pp. 44431–44478 (2023)
20. Perumal, V., Narayanan, V., Rajasekar, S.J.S.: Detection of COVID-19 using CXR and CT images using Transfer Learning and Haralick features. Applied Intelligence 51(1), pp. 341–358 (2021)
21. Gadzicki, K., Khamsehashari, R., Zetsche, C.: Early vs late fusion in multimodal convolutional neural networks. In: Proceedings of the 2020 IEEE 23rd International Conference on Information Fusion (FUSION), pp. 1–6. IEEE (2020)

Prioritizing SQL Injection Threats: A Risk-Based Approach to Detection and Mitigation

Nilima Dinesh Bobade and Swati S.Sherekar

¹ PGDCA PRMITR Badnera,ndbobade@mitra.ac.in

² SGBAU ,Amravati, Swatisherekar2020@gmail.com

Abstract

SQL Injection (SQLi) remains one of the most significant security threats to web applications, enabling attackers to exploit database vulnerabilities and gain unauthorized access to sensitive data. Traditional SQLi detection techniques primarily depend on signature-based approaches, which often fail against evolving attack strategies as well as advance attacks execution. This paper introduces a risk-based model for detecting and mitigating SQLi threats by prioritizing attack instances based on severity and also include security framework. The model employs machine learning-based classification, intruder detection and risk scoring metrics to assess the potential impact of SQLi attacks dynamically. Experimental results demonstrate that the proposed approach effectively enhances SQLi threat prioritization, allowing security teams to respond more efficiently. The study also explores best practices for mitigation, including secure coding techniques, real-time monitoring, and automated threat response mechanisms.

Keywords: SQLi, Data Security, framework, threat generator First Section

1. Introduction

SQL Injection (SQLi) is a common web application vulnerability that allows attackers to manipulate SQL queries and gain unauthorized access to databases. Despite advances in security practices, SQLi remains a critical concern due to poorly validated user inputs and inadequate security measures in many web applications. The challenge lies in not only detecting SQLi attacks but also prioritizing them based on their severity to optimize mitigation efforts.

1.1 Problem Statement

Current SQLi detection methods often produce high false-positive rates, making it difficult for security teams to focus on truly dangerous threats. Moreover, there is a lack of risk-based prioritization models that classify SQLi attacks based on their potential impact and likelihood.

1.2 Research Objectives

This research aims to:

- Develop a risk-based prioritization model for SQLi attacks.
- Implement machine learning algorithms to detect and classify SQLi threats.
- Design an automated risk-scoring system to assess SQLi severity.
- Evaluate mitigation strategies based on risk assessment.

2. Related Work

Previous studies have focused on SQLi detection using rule-based systems, anomaly detection, and machine learning models. However, most lack an effective method for prioritizing threats. Ashlam et al (2022) ^[1] presents a novel machine learning-based approach to detect SQL injection (SQLi) attacks with high accuracy, outperforming previous methods. Ashlam et al (2022) ^[2] designed a multi-phase algorithmic framework to detect and prevent SQL injection attacks in real-time using improved machine learning and deep learning techniques to enhance database security. Dasari et al (2025) ^[4] introduces an approach that leverages generative models to enhance SQL Injection detection and prevention by generating synthetic SQL queries to augment training datasets for machine learning models. Lu et al (2025) ^[5] proposes a novel method for detecting SQL injection vulnerabilities in web applications using a Naive Bayes classifier, a probabilistic model for text classification. Ma et al (2019) ^[6] provides an overview of SQL injection attacks, including the principles, main forms, types, and prevention methods. Paul et al (2024) ^[7] describes SQLR34P3R which is a comprehensive solution that can detect and classify various types of SQL injection attacks from multiple data sources, prioritize the risks of detected attacks, and implement preventive measures to stop the attacks from reaching the backend server. Beriwal et. al (2021) ^[8] presents an efficient SQL injection detection

system using a deep learning-based approach that can detect all types of SQL injection attacks with high accuracy, precision, and recall.
This research builds upon existing methodologies by introducing a risk-based approach that quantifies SQLi attack severity and ranks threats accordingly.

3. Proposed Model

3.1 Risk-Based SQLi Detection Framework

Our model consists of three key components:
 SQLi Detection Module: Uses machine learning classifiers (e.g., Random Forest, SVM, and LSTMs) to identify SQLi attack patterns.
 Risk Scoring Engine: Assigns a risk score based on factors such as attack complexity, data sensitivity, source IP reputation, and impact level.
 Mitigation & Response System: Implements real-time monitoring, automated SQLi prevention mechanisms, and security patches.

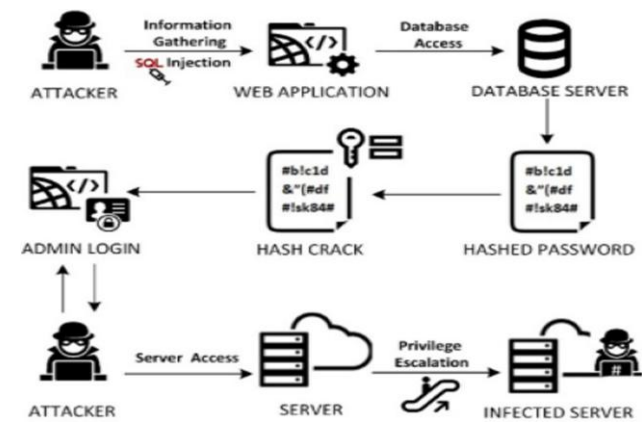


Fig. 3.1 - Risk-Based SQLi Detection Framework Model

3.2 Risk Scoring Formula

The proposed SQLi Risk Score (SRS) is calculated as:

$$\{\text{Risk Score} = (W_1 \times I) + (W_2 \times A) + (W_3 \times Q) + (W_4 \times E)\}$$

Where:

- I** (Input Validation Score): Measures how well user input is sanitized (0 = fully sanitized, 10 = no validation).
 - A** (Authentication Strength Score): Evaluates the strength of authentication mechanisms (0 = strong authentication, 10 = none).
 - Q** (Query Structure Score): Assesses whether queries use parameterized statements (0 = fully parameterized, 10 = fully dynamic).
 - E** (Error Handling Score): Considers exposure of SQL errors (0 = no errors exposed, 10 = verbose error messages shown).
 - W_1, W_2, W_3, W_4 Weighting factors (adjust based on priority, e.g., if input validation is most critical, set W_1 higher).
- Severity: Impact level (e.g., unauthorized data access, data modification, system compromise)
 Likelihood: Probability of attack success based on historical data
 Exposure: System vulnerability level (e.g., weak authentication, lack of parameterized queries)

3.3 Mathematical Model: -

Let –

- Q be the set of SQL queries.
- M be a model that classifies a query $q \in Q$ as **benign (B)** or **malicious (M)**.
- F be a feature space representing properties of SQL queries.
- $P(M|q)$ be the probability that query q is malicious.

Each SQL query q can be represented as a feature vector:

$$\Phi(q) = (f_1, f_2, \dots, f_n)$$

where:

- f_1 = presence of tautologies (e.g., `1=1`)
- f_2 = presence of always-false conditions (e.g., `1=0`)
- f_3 = presence of SQL keywords (`UNION`, `SELECT`, `DROP`, etc.)
- f_4 = length of input
- f_5 = entropy of input string (to detect obfuscation)
- f_6 = frequency of special characters (`'`, `--`, `;`, etc.)

Thus, the classifier function $M(q)$ depends on $\Phi(q)$.

SQL Injection Detection Model –

We define a classification function M :

$$M(q) = \begin{cases} 1, & \text{if } P(M|q) > \theta \\ 0, & \text{otherwise} \end{cases}$$

where:

- θ is a predefined threshold.
- $P(M|q)$ is computed using a model such as:
 - **Rule-Based Systems** (e.g., Regex-based)
 - **Statistical Models** (e.g., Naïve Bayes)
 - **Machine Learning Models** (e.g., Decision Trees, Random Forest, or Deep Learning)

If $M(q) = 1$, the query is flagged as **malicious**.

SQL injection Mitigation model –

If an SQL injection attempt is detected ($M(q) = 1$), mitigation actions can be modeled as:

$$A(q) = \begin{cases} \text{Sanitize,} & \text{if input validation can remove malicious content} \\ \text{Block,} & \text{if query is highly suspicious} \\ \text{Log,} & \text{if further analysis is needed} \end{cases}$$

where $A(q)$ is the action taken based on the severity of $P(M|q)$.

4. Experimental Setup

4.1 Dataset

A dataset consisting of real-world SQLi attack logs and synthetic attack simulations was used. Data was pre-processed and labelled for supervised learning.

Publicly Available SQL Injection Datasets

CSIC 2010 HTTP Dataset

Provided by the Information Security Institute (Spain). Contains legitimate and malicious web requests, including SQL injection attacks.

URL: <https://www.isi.csic.es/dataset/>

SQL Injection Attack Dataset by Kaggle

Includes a variety of SQL injection payloads and safe queries.Useful for training machine learning models.

URL:<https://www.kaggle.com> (Search "SQL Injection Dataset").

OWASP SQL Injection Payloads Dataset

Collection of known SQL injection payloads.Useful for rule-based detection.

URL:<https://owasp.org/www-project-wstg/>

Web Application Firewall (WAF) Logs

Many organizations collect real-world attack logs from WAFs, such as ModSecurity logs.

4.2 Model Training

Features extracted: Query structure, special character usage, source IP reputation, execution time.
Algorithms applied: Random Forest, Support Vector Machine (SVM), LSTM for sequence-based anomaly detection.

4.3 Evaluation Metrics

Precision, Recall, F1-score for classification performance.
Risk-based prioritization accuracy to measure ranking effectiveness.

5. Results and Discussion

Several approaches have been used for SQLi detection and prevention, with varying levels of effectiveness. Below is an analysis of different techniques and their outcomes.

Technique	Detection Rate	False Positives	Performance Impact	Best Use Case
Signature-Based (Regex, WAF)	70-90%	10-30%	Low	Quick rule-based filtering
Machine Learning (SVM, LSTM)	85-98%	5-15%	Moderate to High	Advanced, adaptive security
Behavioral Analysis	80-95%	8-20%	Moderate	Detecting zero-day threats
Query Parameterization	100% (Prevention)	0%	Low	Secure application coding
Web Application Firewalls	75-90%	10-25%	Low	Frontline defense

- **Best Prevention Method:** Query Parameterization (100% prevention).
- **Best Detection Method:** **Machine Learning + Behavioural Analysis** (for unknown attacks).
- **Best First Line of Defense:** WAFs (cost-effective, easy to deploy).

5. Mitigation Strategies

Secure Coding Practices: Implementing prepared statements, parameterized queries, and input validation.

Automated Security Response: Real-time attack detection triggers firewall rules and IP blocking.

Continuous Monitoring: Deploying AI-driven anomaly detection systems to detect deviations in SQL query patterns.

7. Conclusion

SQL Injection (SQLi) remains a persistent and evolving threat to web application security, requiring advanced detection and mitigation strategies. Traditional detection techniques, primarily signature-based approaches, struggle to adapt to evolving attack vectors and often generate high false-positive rates. To address these limitations, this study proposed a risk-based SQLi detection and mitigation model that leverages machine learning, dynamic risk scoring, and automated threat response mechanisms.

The proposed model enhances SQLi detection by classifying attack patterns with machine learning classifiers while simultaneously assessing their severity using a risk-scoring engine. This approach allows security teams to prioritize threats based on their impact, enabling a more efficient and targeted response. Experimental results confirm that the model improves SQLi threat detection accuracy while minimizing false positives and response time.

Additionally, this research explores best practices for SQLi mitigation, including secure coding techniques, real-time monitoring, and automated security enforcement. The integration of a dynamic risk assessment framework further strengthens web application defences against both known and zero-day SQLi attacks.

Future work can explore the incorporation of deep learning models for improved detection accuracy, as well as the development of adaptive risk-scoring mechanisms that evolve with emerging threats. Furthermore, integrating blockchain-based security mechanisms could enhance data integrity and prevent unauthorized modifications.

By advancing risk-aware SQLi detection and mitigation strategies, this research contributes to strengthening web application security and reducing the overall impact of SQL injection attacks.

References

- [1] Ashlam, A. A., Badii, A., & Stahl, F. (2022a). A Novel Approach Exploiting Machine Learning to Detect SQLi Attacks. *2022 5th International Conference on Advanced Systems and Emergent Technologies (IC_ASET)*, 513–517. https://doi.org/10.1109/IC_ASET53395.2022.9765948
- [2] Ashlam, A. A., Badii, A., & Stahl, F. (2022b). Multi-Phase Algorithmic Framework to Prevent SQL Injection Attacks using Improved Machine learning and Deep learning to Enhance Database security in Real-time. *2022 15th International Conference on Security of Information and Networks (SIN)*, 01–04. <https://doi.org/10.1109/SIN56466.2022.9970504>
- [3] Ashlam, A. A., Badii, A., & Stahl, F. (2022c). WebAppShield: An Approach Exploiting Machine Learning to Detect SQLi Attacks in an Application Layer in Run-Time. <https://doi.org/10.5281/ZENODO.6983905>
- [4] Dasari, N. S., Badii, A., Moin, A., & Ashlam, A. (2025). Enhancing SQL Injection Detection and Prevention Using Generative Models (No. arXiv:2502.04786). arXiv. <https://doi.org/10.48550/arXiv.2502.04786>
- [5] Lu, Z. (2025). Sql injection detection using Naïve Bayes classifier: A probabilistic approach for web application security. *ITM Web of Conferences*, 70, 04016. <https://doi.org/10.1051/itmconf/20257004016>

- [6] Ma, L., Zhao, D., Gao, Y., & Zhao, C. (2019). Research on SQL Injection Attack and Prevention Technology Based on Web. 2019 International Conference on Computer Network, Electronic and Automation (ICCNEA), 176–179. <https://doi.org/10.1109/ICCNEA.2019.00042>
- [7] Paul, A., Sharma, V., & Olukoya, O. (2024). SQL injection attack: Detection, prioritization & prevention. *Journal of Information Security and Applications*, 85, 103871. <https://doi.org/10.1016/j.jisa.2024.103871>
- [8] R, J. K., Balaji B, S., Pandey, N., Beriwal, P., & Amarajan, A. (2021). An Efficient SQL Injection Detection System Using Deep Learning. 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), 442–445. <https://doi.org/10.1109/ICCIKE51210.2021.9410674>
Sirmulla, A., & M, P. (2023). A Supervised Ensemble Learning-Based Approach to Mitigate SQL Injection Attack on Smart City Data. *International Journal of Engineering Trends and Technology*, 71(11), 18–26. <https://doi.org/10.14445/22315381/IJETT-V71I11P202>

On Data Privacy in e-Governance Data Mining

Shailesh Kadre¹, Vipin K. Tripathi²

¹Data Analyst, Hyderabad

²National Institute of Technical Teachers' Training & Research, Bhopal, India, India

Shaileshkadre@gmail.com

Abstract

Mining e-Governance (e-gov) data can lead to better planning and faster government decisions. Anonymization or removal of sensitive attributes before mining is the first significant step to adhere to regulations and incorporate policy concerns. In this work, we investigate and compare seven anonymization techniques, routinely used by industry professionals as well as researchers, for their use in data mining of e-gov data and make recommendations. Our results show that with an increase in parameters, privacy level (k-anonymity), privacy budget (Differential Privacy), and t-closeness parameter, the accuracy of classification tasks by the Random Forests algorithm declines or remains constant. For differential privacy, this decline in accuracy is seen as more significant in categorical variables than in numeric ones. Increasing these parameters can lead to a strengthening of privacy levels. Smaller values of these parameters can yield better accuracies. The exact values can differ based on the dataset, machine learning algorithm, and data type.

CCS CONCEPTS • e-Governance • Privacy Protection • Machine Learning.

Additional Keywords and Phrases: Privacy Protection Techniques, Related Attack

1. INTRODUCTION

E-Gov data has some peculiar characteristics like massive, centralized data repositories, PII concentration, long-term storage, and high-profile targets. A successful breach or attack can fetch an immense volume of sensitive information, which can be exploited for various purposes such as identity theft, targeted exploitation, fraud, or cause damage to national security and loss of public trust. Attacks that are possible on the e-gov data include linkage attacks, identity disclosure attacks, background knowledge attacks, homogeneity attacks, membership disclosure attacks, attribute disclosure, inference attacks, and more. Consider a simplified example of a linkage attack below (Table 1).

SSN	Tax Paid
123-45-68901	\$11,000
987-65-55253	\$8,700
246-80-24535	\$12,500

SSN	Medical History
123-45-68901	Asthma
987-65-55253	Hypertension
246-80-24535	Diabetes

With Tables An adversary could link the SSN "987-65-55253" from the left dataset with the matching SSN in the right dataset, identifying an individual's tax payment amount (\$8,700) and his medical history (Hypertension). This linkage attack allows the adversary to get extra information about the individual outside the publicly available information from either dataset alone. A similar example is the analysis of the 1990 U.S. Census performed by Sweeney [6]. He found different patterns of QIs (Quasi Identifiers) that would uniquely identify a person in the US. Using these QIDs he identified the then Massachusetts governor (William Weld) and get his medical records. He

used an anonymous voter list and anonymized medical data acquired from the group insurance commission [6]. In this work, we have taken up to study the effect of altering data privacy levels with seven techniques and study its effect on classification accuracy. We make some recommendations on privacy-accuracy trade-offs. A detailed description of these methods and the related attacks is beyond the scope of this short paper.

Table 2 Privacy attacks and preserving techniques.

#	Anonymity Methods	Attribute	Resists Attacks	Limitation(s)
1	k-anonymity [30]	QI	Linkage, identity disclosure [1], and attributedisclosure [1].	Background knowledge (BK) attack, homogeneity attack [1].
2	l-diversity [27, 31]	SA	Attribute disclosure and homogeneity attacks[1].	membership disclosure attacks[1].
3	t-closeness [27, 28, 29, 32]	SA	Depending on the value of t, maybe Membership disclosure [35] attacks.	A higher t may lead to degradation in data utility [1].
4	Generalization (G) [23]	SA	Linkage Attacks, identity and semanticsimilarity attacks [35,37].	Data distortion and hence lowdata utility.
5	Distortion (D) [25]	SA	Inference attacks [36, 37].	Data distortion & low data utility.
6	Suppression (S) [24]	SA	Identity disclosure, and linkage attacks [37].	Data distortion & low data utility.
7	Semi-suppression (SS)	SA	Identity disclosure [37].	Data distortion & low datautility.
8	Masking (M) [9, 10, 33,34]	SA	Membership disclosure, differential poweranalysis attacks [38,39].	Data distortion & low datautility.
9	Differential privacy [2,26]	SA	Reconstruction, membership inferenceattacks [40, 41].	Inaccurate for small datasets

2. RELATED WORK

Numerous research articles in the literature study privacy-preserving data mining across domains [1, 6, 7-9, 14, 15]. Some of the prominent studies that consider data mining, privacy, & security aspects for e-gov, specific privacy techniques, attacks, risk analysis, and evaluation of anonymized data include [3-5, 21, 22, 42, 44, 45, 50, 51-53]. We distinguish our work from some key articles that discuss anonymization with data mining [11-13]. Rajput et al. [11] have applied two mining methods for finding classification rules and generating a decision tree. Agarwal et al. [12] have propped a nationwide data warehouse model in the Indian context. Arora et al. [13] discuss a relationship between e- gov, data warehousing, and data mining. Goldsteen et al. [54] have proposed model-guided data anonymization, however, such methods are not common and can be difficult to implement. Any of these studies don't cover multiple methods for e- gov data mining. The following sections cover our contribution in greater detail.

3. CONTRIBUTION

We discuss the merits and limitations of seven industry-used anonymization techniques, and related attacks (Table-1), and study the effects of varying privacy levels and budgets, Fig 1(a) to 1(h) below. We experiment with two datasets to validate the repeatability of results. This work can help professionals in picking proper anonymization techniques based on data type, desired privacy levels and budgets, and possible attacks, and attain acceptable classification accuracies with anonymized data.

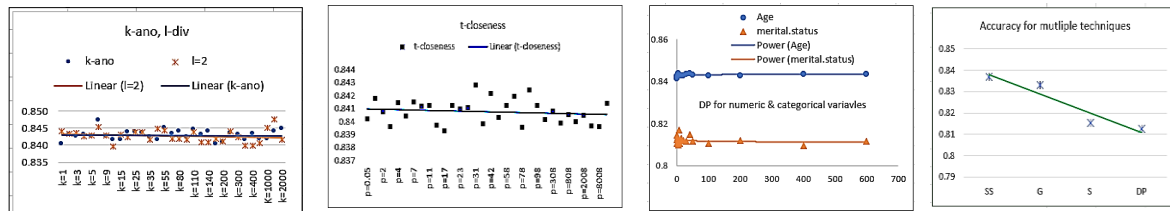
4. EXPERIMENTAL SETUP

The objectives of our experiments are to (1) compare the classification accuracy of the Random Forests (RF) model

when the data is anonymized with seven distinct privacy protection (PP) techniques & (2) study the variations in the RF accuracy with different PP parameter setups. We choose RF as it has been successfully applied to various domains and problem types. It captures non-linear relationships and is robust to overfitting. This versatility makes RF suitable for a wide range of research scenarios [49]. We have used publicly available Adult Census datasets & Loan Data with target variables containing only two classes.

5. RESULTS & CONCLUSIONS

In the discussion below, as a baseline, we will use the accuracy of the original dataset (0.8347). For l-diversity, l is constant at 2 as SA ('income') has only two classes. In Fig 1(a) (k-anonymity), Fig 1(b) (l-diversity), and Fig 1(c) (t-closeness), all in the first row in Figure 1, we observe a marginal decline trend in accuracy (less than the 1% range) with increased parameters k, p, e that relate to the privacy levels. With the increase in k, the privacy protection strengthens, but it leads to a decrease in classification accuracy because it requires more records to be grouped together and generalized causing a loss of information and increased data distortion [46]. An increase in p in Fig 1(c), first row, also strengthens privacy protection, but it introduces more noise or randomness into the model's predictions [46]. We apply Differential privacy (DP), Fig 1(c) & 1(g), to one numeric and one categorical variable for both datasets. We see the accuracy curves flatten out or show a decline after initial variations in both cases. In both datasets, for numeric features, there is a negligible decrease in accuracy compared to the baseline (0.8347). But for



the categorical feature, we see a sharp decrease of approximately 3% for the Adult Census data and in much lesser amounts for Loan Data. The general toll of differential privacy is a decline in the model's accuracy [47]. From Fig 1(c) & 1(g), 3rd from left, Figure 1, first & second rows, we observe that, for DP, the decrease in RF classification accuracy is more significant in categorical variables compared to numeric ones. It could be because the categorical variables that we tried for both datasets have fewer classes than the corresponding numeric ones, so the added noise can affect the model's ability to correctly categorize the data based on this feature [46, 47, 48].

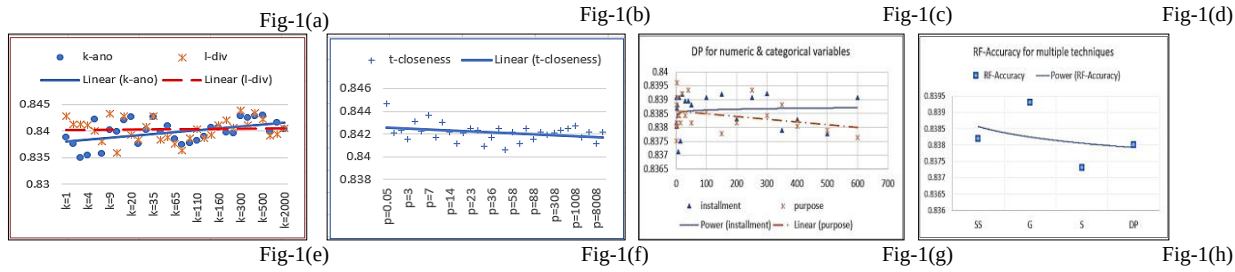


Figure 1: First-row Adult Census (Fig 1(a) to 1(d), left to right k vs accuracy, t-closeness p vs accuracy, DP e vs accuracy, multimethod. plot. the second row left to right, the same plots for Loan Data (Fig 1(e) to 1(h).

Fig 1(d) & 1(h), fourth from left, rows one & 2, plot results for multiple techniques, here the trend lines can be explained easily as the loss of information increases as we move right on the x-axis. Total suppression (S) is more information loss compared to semi-suppression (SS) & generalization (G). For Adult Census, SS and G show a marginal drop in accuracy (from 0.8347 to 0.83), while S shows a more than 2% decline to 0.81.

Fig 1(e) to 1(h), second row, belong to Loan Data. They show trends similar to that of the Adult Census dataset, Fig 1(a) to 1(d), first row. The only exception to the previous statement is the Loan Data k-anonymity trendline (Fig 1(e), first in the second row), which is showing a slightly increasing trend in the band of 1 to 2% accuracy with the increase in levels of privacy (k). It needs to be investigated further. From Fig 1(a) to 1(h), we have a total of 12 trend lines, 6 in each row. They belong to seven distinct techniques tried out on categorical and numeric variables as applicable. In these plots, the pattern is similar in 11 out of the total 12 trend lines (91.67%). As you move right on the x-axis in any plot the accuracy of classification tasks by the Random Forests algorithm either declines or stays constant.

6. RECOMMENDATIONS

Based on the 91.67% trends, we can infer that for k-anonymity and l-diversity, a lower value of k (roughly 2 to 8) is preferable. If case demands, increasing k to larger values (above 8) can be permissible as its effect on RF classification accuracy is seen as minimal. The same is true for choosing the t-closeness parameter. For DP, choosing the right privacy budget (e) seems to be more dependent on the dataset and the type of variables. So, experiment to choose the right values. From Fig 1(d) & 1(h), we infer that for better accuracy SS and G work better than S, which destroys the information value of that feature. Results show the accuracy loss in DP is more for categorical features than numeric ones with an increased privacy budget. So, if the case allows, for categorical features, G or SS can be preferred over DP and S. We can use multiple techniques like k-anonymity, and generalization to one dataset as per the business case. It can be the topic of future research.

REFERENCES

- [1] Iyiola E. Olatunji, Jens Rauch, Matthias Katzensteiner, and Megha Khosla. 2022. A Review of Anonymization for Healthcare Data. Big Data (2022).=
DOI:https://doi.org/10.1089/big.2021.0169
- [2] Cynthia Dwork and Aaron Roth. 2013. The Algorithmic Foundations of Differential Privacy. Foundations and Trends® in Theoretical Computer Science 9, 3 (2013), 211-407. DOI:https://doi.org/10.1561/04000000042
- [3] Scott E. Coull, Fabian Monroe, Michael K. Reiter, and Michael Bailey. 2009. The Challenges of Effectively Anonymizing Network Data. 2009 Cybersecurity Applications & Technology Conference for Homeland Security

(2009). DOI:<https://doi.org/10.1109/catch.2009.27>

[4] Tianli Yang, Li Shan Cang, Muddesar Iqbal, and Dhafer Almakhlles. 2023. Attack Risk Analysis in Data Anonymization in Internet of Things. *IEEE Transactions on Computational Social Systems* (2023), 1-8. DOI:<https://doi.org/10.1109/tcss.2023.3243089>

[5] Srivatsava Ranjit Ganta, Shiva Prasad Kasiviswanathan, and Adam Smith. 2008. Composition attacks and auxiliary information in data privacy. *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining* (2008). DOI:<https://doi.org/10.1145/1401890.1401926>

[6] Sweeney, Latanya. 2000. Simple demographics often identify people uniquely. *Health* (San Francisco) 671.2000 (2000): 1-34. Available at: <https://privacytools.seas.harvard.edu/sites/projects.iq.harvard.edu/files/privacytools/files/paper1.pdf>

[7] Charu C. Aggarwal and Philip S. Yu. 2008. A General Survey of Privacy-Preserving Data Mining Models and Algorithms. *Privacy-Preserving Data Mining* (2008), 11-52. DOI:https://doi.org/10.1007/978-0-387-70992-5_2

[8] Rong Jiang, Rongxing Lu, and Kim-Kwang Raymond Choo. 2018. Achieving high performance and privacy-preserving query over encrypted multidimensional big metering data. *Future Generation Computer Systems* 78, (2018), 392-401. DOI:<https://doi.org/10.1016/j.future.2016.05.005>

[9] Asenjo, Juan C. 2017. Data masking, encryption, and their effect on classification performance: trade-offs between data security and utility. *Diss. Nova Southeastern University*. Available at: <https://search.proquest.com/openview/3823d7e4aac9aa9c432c5ca95f01e22b/1?pq-origsite=gscholar&cbl=18750>

[10] Eric D. Ragan, Hye-Chung Kum, Gurudev Ilangovan, and Han Wang. 2018. Balancing Privacy and Information Disclosure in Interactive Record Linkage with Visual Masking. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (2018).

[11] Rajput, Anil, Ramesh Prasad Aharwal, Meghna Dubey, S. Saxena, and Manmohan Raghuvanshi. 2011. J48 and JRIP rules for e-governance data. *International Journal of Computer Science and Security (IJCSS)* 5, no. 2 (2011): 201. Available at: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=9a0e0ac6fb02e61bea64cf1a4f2bc6b673282d31#page=43>

[12] Sonali Agarwal, Neera Singh, and Dr.G.N. Pandey. 2010. Implementation of Data Mining and Data Warehousing In E-Governance. *International Journal of Computer Applications* 9, 4 (2010), 18-22. DOI:<https://doi.org/10.5120/1374-1851>

[13] Rakesh K. and Manoj K. 2017. e-Governance using Data Warehousing and Data Mining. *International Journal of Computer Applications* 169, 8 (2017), 28-31. DOI:<https://doi.org/10.5120/ijca2017914785>

[14] Ashish Kumar Jha and Indranil Bose. 2013. A Framework for Addressing Data Privacy Issues in E-Governance Projects. *Journal of Information Privacy and Security* 9, 3 (2013), 18-33. DOI:<https://doi.org/10.1080/15536548.2013.10845683>

[15] Moharana, Chitta Ranjan, Manas Kumar Pal, and Debasish Rout. 2012. Information and network security in e-governance: Perspective and issues. *International journal of research in engineering information technology and social sciences* 2.11 (2012). Available at: http://indusedu.org/pdfs/IJREISS/IJREISS_564_58786.pdf

[16] Jaijit Bhattacharya. 2005. Privacy Technology for E-Governance. *Department of Management Studies Indian Institute of Technology, Delhi, India*, pp.114-124. Available at:

https://www.researchgate.net/profile/Jaijit-Bhattacharya/publication/228718693_Privacy_Technology_for_E-Governance/links/5ad2fd230f7e9b28593450fa/Privacy-Technology-for-E-Governance.pdf

Verma, Krishna Kumar, and Navita Shrivastava, 2015. Exploring the Role and associated challenges of data mining technique and it's implementation in e-governance. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 4(8). Available at: <https://www.researchgate.net/profile/Krishna-Verma>

4/publication/319109954_IJARCET-VOL-4-ISSUE-8-3473

- [1] Verma, Krishna Kumar, and Navita Shrivastava, 2015. Exploring the Role and associated challenges of data mining technique and it's implementation in e-governance. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 4(8). Available at: https://www.researchgate.net/profile/Krishna-Verma-4/publication/319109954_IJARCET-VOL-4-ISSUE-8-3473_3478/data/5991ca74458515a8a24bbbc8/IJARCET-VOL-4-ISSUE-8-3473-3478.pdf
- [2] Fung, Benjamin CM, Ke Wang, Ada Wai-Chee Fu, and S. Yu Philip. 2010. Introduction to privacy-preserving data publishing: Concepts and techniques. CRC Press.
- [3] PART, I., 2018. Data Protection Act, 2018.
- [4] Harding, Elizabeth Liz, Jarno J. Vanto, Reece Clark, L. Hannah Ji, and Sara C. Ainsworth. 2018. Understanding the scope and impact of the California consumer privacy act of 2018. Journal of Data Protection & Privacy, 2(3), pp.234-253. Available at : <https://www.ingentaconnect.com/content/hsp/jdpp/2019/00000002/00000003/art00007>
- [5] Jaijit Bhattacharya. 2005. Privacy Technology for E-Governance. Department of Management Studiesn Indian Institute of Technology, Delhi, India, pp.114-124. Available at: <https://www.ingentaconnect.com/content/hsp/jdpp/2019/00000002/00000003/art00007>
- [6] Tiancheng Li, Ninghui Li, Jian Zhang, and Ian Molloy. 2012. Slicing: A New Approach for Privacy Preserving Data Publishing. IEEE Transactions on Knowledge and Data Engineering 24, 3 (2012), 561-574. DOI:<https://doi.org/10.1109/tkde.2010.236>
- [7] Hyukki Lee, Soohyung Kim, Jong Wook Kim, and Yon Dohn Chung. 2017. Utility-preserving anonymization for health data publishing. BMC Medical Informatics and Decision Making 17, 1 (2017). DOI:<https://doi.org/10.1186/s12911-017-0499-0>
- [8] Yang Xu, Tinghuai Ma, Meili Tang, and Wei Tian. 2014. A Survey of Privacy Preserving Data Publishing using Generalization and Suppression. Applied Mathematics & Information Sciences 8, 3 (2014), 1103-1116. DOI:<https://doi.org/10.12785/amis/080321>
- [9] Saif M. A. Kabir, Amr M. Youssef, and Ahmed K. Elhakeem. 2007. On Data Distortion for Privacy Preserving Data Mining. 2007 Canadian Conference on Electrical and Computer Engineering (2007). DOI:<https://doi.org/10.1109/ccece.2007.83>
- Rajendran, Keerthana, Manoj Jayabalan, and Muhammad Ehsan Rana.2017. A study on k-anonymity, l-diversity, and t-closeness techniques. IJCSNS, 17(12), p.172. Available at: https://www.researchgate.net/profile/Manoj-Jayabalan/publication/322330948_A_Study_on_k-anonymity_l-diversity_and_t-closeness_Techniques_focusing_Medical_Data/links/5a5483a00f7e9bbc105a5da4/A-Study-on-k-anonymity-l

Gaussian Process-Assisted Inverse Design for Adaptive Pareto Front Gap Filling in Multi-Objective Optimization

Shailesh S. Kadre,¹ S. K. Vemuri,² P.K. Mukka³, R.G. Paspula⁴ and Vipin K. Tripathi⁵

^{1, 2, 3, 4}Analysis and Simulation Department, Cyient Ltd, India

⁵National Institute of Technical Teachers' Training & Research, Bhopal, India

Abstract: Multi-objective optimization problems (MOOPs) often involve computationally expensive simulations, making the comprehensive exploration and characterization of the Pareto front a significant challenge. This paper introduces an efficient surrogate-assisted framework for identifying and strategically filling gaps within the Pareto front of MOOPs, leveraging the power of Gaussian Process (GP) regression in an inverse design context. Our methodology integrates the well-established NSGA-III evolutionary algorithm with an adaptive sampling strategy. Initially, a global GP surrogate model is trained on a limited dataset to guide the NSGA-III in a broad exploration of the objective space. Subsequently, a novel distance-based gap-filling mechanism analyzes the generated Pareto front to pinpoint under-explored regions. For these identified objective space gaps, the trained GP models are utilized in an inverse manner to infer the corresponding design variables. This inverse mapping facilitates the generation of new, high-fidelity solutions that specifically target and bridge the detected discontinuities, thereby enhancing the quality and density of the Pareto front. The iterative nature of this approach adaptively refines the surrogate models and progressively enriches the Pareto optimal set. The effectiveness of the proposed hybrid methodology, combining global optimization with targeted inverse design, is demonstrated on a five-variable, three-objective, two-constraint problem, showcasing its capability to achieve a well-distributed and improved Pareto front with a controlled number of expensive function evaluations.

Keywords: multi-objective optimization, surrogate modeling, Gaussian Process regression, Pareto front gap-filling, inverse design, NSGA-III

• Introduction

Multi-objective optimization problems (MOOPs) are prevalent across diverse engineering disciplines, scientific research, and decision-making processes. Unlike single-objective optimization, MOOPs aim to simultaneously optimize several, often conflicting, objectives. This inherently leads to a set of optimal trade-off solutions known as the Pareto front, rather than a single best solution. For designers and decision-makers, a comprehensive and well-distributed understanding of this Pareto front is paramount, as it reveals the trade-offs between different performance criteria and provides a rich set of non-dominated design alternatives.

However, a significant challenge in real-world MOOPs arises from the computational expense associated with evaluating the objective and constraint functions. These evaluations often require complex numerical simulations, such as finite element analysis or computational fluid dynamics, which can take minutes, hours, or even days for a single design point. This high computational burden renders traditional exhaustive search methods or even many population-based evolutionary algorithms (EAs) prohibitively time-consuming, severely limiting the number of design evaluations that can be performed within practical timeframes.

Furthermore, while advanced MOO algorithms can converge towards the Pareto front, the resulting solutions can often be sparse or contain significant gaps, particularly in high-dimensional objective spaces.

Such discontinuities hinder a complete understanding of the available trade-offs and may lead to overlooking promising design regions.

To mitigate the computational cost, surrogate models, also known as metamodels or response surface models, have emerged as powerful tools in optimization. Techniques like Polynomial Chaos Expansion, Radial Basis Functions, Artificial Neural Networks, and Gaussian Processes (GPs) can build fast-to-evaluate approximations of the expensive simulation models. While surrogate-assisted multi-objective optimization (SAMO) has seen substantial advancements, many existing approaches primarily focus on accelerating the convergence to an approximated Pareto front. There remains a critical need for methodologies that not only efficiently approximate the Pareto front but also actively and adaptively address its density and uniform distribution by explicitly identifying and filling sparse regions.

This paper addresses these limitations by proposing a novel, adaptive framework that synergistically combines the global search capabilities of the NSGA-III evolutionary algorithm with the predictive and, crucially, inverse modeling power of Gaussian Processes. Our core contribution lies in leveraging GP surrogate models in an inverse design context: we identify existing gaps in the approximated Pareto front and then, through the inverse mapping capability of the trained GPs, infer the corresponding design variables that would yield objective values aligned with these desired gap-filling points. This iterative process allows for intelligent, data-driven refinement of the Pareto front, ensuring a more comprehensive and well-distributed representation with a controlled total number of expensive function evaluations. The framework effectively balances global exploration for new optimal regions with targeted local exploitation for improving the density and uniformity of the obtained solutions.

The remainder of this paper is organized as follows: Section 2 provides a brief literature review on SAMO with inverse design, Pareto fronts, and Gaussian Process surrogates. Section 3 details the proposed adaptive inverse design framework, outlining its various components and their interactions. Section 4 describes the experimental setup and the problem used for validation. Section 5 presents and discusses the results, demonstrating the framework's effectiveness. Finally, Section 6 concludes the paper and outlines avenues for future work

• Literature review

A robust body of literature explores **Gaussian Process (GP)–based surrogate modeling** in multi-objective optimization (MOO), especially when integrated with **Pareto front shaping, inverse mapping, and gap-filling** strategies.

GP-based inverse-model MOEAs have recently demonstrated strong potential in guiding evolutionary search within Pareto spaces. For instance, Cheng et al. (2015) introduced the IM-MOEA, an online GP inverse model embedded in reproduction, effectively mapping objective vectors back to decision variables to fill gaps in the Pareto front. Subsequent enhancements targeted irregular Pareto front shapes via adaptive reference vectors and biased crossover (Zhang et al., 2020), and more recently in inverse-modeling constrained variants (Tran et al., 2021). Dynamic MOO using inverse GP models has also proven effective in tracking Pareto evolution in non-stationary settings (Lartaud et al., 2024).

Adaptive sampling and gap-recognition approaches aim to improve Pareto front coverage. Pellicer et al. (2022) propose a three-step routine—identification, validation, and filling of Pareto front gaps—which aligns closely with the gap-filling focus of this work. Zhu et al. (2024) propose a Pareto-diverse batch acquisition method that explicitly prioritises filling objective-space discontinuities.

Kadre and Tripathi have extensively explored surrogate-assisted multi-objective optimization. Their work spans from developing advanced and mixture surrogate models to applying them in vehicle crashworthiness (SAMO, RASAMO) and aero-engine RUL prediction using optimized LSTM models. These contributions highlight the effectiveness of surrogate models in reducing computational cost while maintaining solution quality in complex engineering problems.

Surrogate-assisted MOO frameworks with Bayesian sampling offer complementary insights. Dong et al. (2021) propose angle-based and penalised hyperparameter adaptation in GP-assisted NSGA-type frameworks for exploration–exploitation balancing. Ruiz et al. (2021) integrate GP regression with EHVI-style acquisition for constrained objective convergence in power systems. Meanwhile, Sun, Li & Li (2022) demonstrate ensemble RBF + Kriging methods for enhanced Pareto resolution in expensive MOO tasks.

Applications in engineering domains illustrate GP surrogate adaptability. Han & Zhang (2023) employ multi-fidelity GP models in wind farm layout optimization to handle computationally expensive fluid dynamics simulations. Tran et al. (2021) develop a 3-GP Bayesian optimization framework (srMO-BO-3GP) within structural design involving constraints. Shi et al. (2023) couple physics-enhanced GP models with inverse design in additive manufacturing.

Our paper builds on this literature in three key dimensions:

(a) GP-Inverse Mapping: We apply and extend IM-MOEA ideas with GP-based inverse models embedded in NSGA-III to strategically fill identified Pareto front gaps.

(b) Adaptive Gap Discovery: Leveraging gap-identification routines (e.g. Pellicer et al., 2022; Zhu et al., 2024), we direct the surrogate towards under-sampled regions prior to inverse sampling, ensuring targeted exploration.

(c) Iterative Refinement: We incorporate dynamic hyperparameter and acquisition tuning (Dong et al., 2021), and mode-aware strategies from srMO-BO-3GP (Tran et al., 2021), to iteratively enrich the Pareto front with minimal expensive evaluations.

• **Method of analysis (Experimental procedure/ Numerical experiments)**

Following steps show the implemented for NSGA-III based SAMO algorithm. This work flow is inspired by the original work of Bhattacharjee et al. (2016) which involved surrogate assisted multi- objective optimization algorithm with NSGA- II as base.

Table 1: Workflow of the NSGA-III-Based Surrogate-Assisted Multi-Objective Optimization (SAMO) Algorithm

Step	Description
1. Initial Setup	- Set the maximum number of actual expensive simulations, t_{max}. - Define n as the number of design variables. - Initialize population size as greater than $2(n + 1)$. - Set the function evaluation counter $t = 0$.
2. Generate Initial Population	- Use Latin Hypercube Sampling (LHS) to generate the initial population. - Perform costly evaluations of objective and constraint functions. - Store evaluated data in the global archive A. - Update the function evaluation counter t.
3. Ranking and Selection	- Rank the solutions based on constraint feasibility.
Main Loop (While $t < t_{\text{max}}$):	
4. Offspring Generation	- Generate offspring population using evolutionary operators.
5. Local Surrogate Modeling	- For each offspring: • Select the k-nearest points from the archive to form a local training set. • Identify upper and lower bounds for each variable. • Train Gaussian Process models for each objective and constraint using local data.
6. Multi-Objective Optimization	- Apply NSGA-III on surrogate models to obtain an approximate Pareto front.
7. Pareto-Front Combination	- Combine all Pareto-optimal solutions from the current generation.
8. Non-Dominated Sorting	- Perform non-dominated sorting to identify top unique solutions not present in the archive.
9. Evaluation and Archive Update	- Select top-ranking solutions and perform expensive evaluations. - Add evaluated solutions to the archive A. - Update t based on the number of new evaluations.
10. Solution Reduction & Performance Metrics	- Reduce the archive to a fixed population size using non-dominated sorting. - Calculate performance metrics such as Hypervolume (HV). - Compute the Nadir point using all non-dominated solutions across runs.

In our optimization algorithm, we use a **gap-filling strategy** to improve the spread and coverage of the Pareto front. This strategy is based on the concept of **inverse design**.

In a typical design problem (called **forward design**), we start with some design inputs and calculate their performance (objectives). But in **inverse design**, we flip this process: we start with desired performance values (objective targets) and try to find the design inputs that could produce those values. This is useful when we know what we want and are trying to figure out how to achieve it.

Here's how we apply this idea in our algorithm:

- 1) We first **identify gaps** between well-performing solutions on the current Pareto front—these are regions where no solution has been explored yet.
- 2) We then **generate target objective values** within these gaps.
- 3) Using **surrogate models** (like Gaussian Process regressors), we predict the design parameters that would likely produce these target objective values. This step is what we call **inverse design**.
- 4) These predicted designs are then added to our solution pool and evaluated.
- 5) By doing this, we avoid unnecessary simulations and **fill in missing regions** of the Pareto front more intelligently. It improves both the **diversity and quality** of the final set of solutions.
- 6) Following steps are followed for NSGA-III Based Surrogate-Assisted Multi-Objective Optimization (SAMO) Algorithm with Inverse Design.

Table 2: NSGA-III Based Surrogate-Assisted Multi-Objective Optimization (SAMO) Algorithm with Inverse Design

Step	Description
1. Initial Setup	- Define the maximum number of expensive evaluations, t_{max}. - Specify the number of design variables, n. - Set initial population size $pop\_size > 2(n + 1)$. - Initialize evaluation counter $t = 0$. - Create empty archives for design variables, objective values, constraint values, and Pareto front.
2. Generate Initial Population	- Generate an initial design population using space-filling sampling techniques (e.g., LHS). - Perform costly evaluations to compute objective and constraint values. - Store results in the global archive. - Increment the evaluation counter t.
3. Initial Ranking and Archive Formation	- Perform non-dominated sorting on the initial evaluated solutions. - Store the top-ranked non-dominated solutions in the Pareto front archive.
4. Identify Gaps in Objective Space	- Analyze sparsity between adjacent Pareto solutions using inter-point distances. - Identify large gaps by comparing to a distance threshold (e.g., median $\times$ multiplier). - Generate intermediate objective points to fill gaps via interpolation.
5. Construct Global Surrogate Models	- Train global surrogate models (e.g., Gaussian Process Regressors) for each objective using archived data. - Use these models to approximate the mapping from design space to objective space.
6. Perform Inverse Design	- For each interpolated objective target: • Solve an inverse optimization problem using surrogate models to estimate feasible design variables. • Use multiple restarts for robustness. - Predict objectives for the reconstructed design vectors.
7. Filter New Candidates	- Apply non-dominated sorting to the reconstructed solutions. - Select top-ranked solutions not present in the archive. - Filter out duplicate or already evaluated design vectors.
8. Expensive Evaluations	- Evaluate objective and constraint functions for the selected candidate solutions using the original (expensive) simulation. - Add the new evaluations to the archive. - Update the evaluation counter t.
9. Global Archive Update	- Merge new evaluations with the existing archive. - Perform non-dominated sorting over the combined population. - Retain the top-ranked non-dominated solutions as the updated Pareto archive.
10. Visualization and Monitoring	- Plot 2D projections of the current Pareto front (e.g., intrusion vs acceleration). - Monitor progress: generation number and total function evaluations.
11. Post-Processing	- After $t \geq t_{max}$, merge all non-dominated solutions from all optimization runs. - Compute the Nadir point from the combined solutions. - Evaluate performance metrics (e.g., Hypervolume) using the Nadir reference.
12. Independent Repetitions	- Repeat the entire workflow for multiple runs (e.g., 10) using independently sampled initial populations. - Use combined results for robustness and metric computation.

Problem Details: The current vehicle crashworthiness problem is a multi-objective optimization problem (MOOP) related to frontal car impacts, as studied by Bhattacharjee et al. (2016). In this study, the objectives are reordered based on their relative importance to occupant safety, specifically toe-board intrusion and occupant acceleration. The optimization involves minimizing the following three objective functions:

- Toe-board Intrusion (f_1): Represents the first objective, capturing structural intrusion at the toe-board. This measure combines results from both full-frontal and offset crash scenarios as reported in Bhattacharjee et al. (2016).
- Integrated Acceleration (f_2): The second objective is the time-integrated acceleration during impact, indicative of occupant load.

$$A_{in} = \int_{t_1}^{t_2} a dt$$

- Vehicle Mass (f_3): The third objective is the overall mass of the vehicle, a critical factor in lightweight design and energy efficiency.

Design sensitivity studies by Bhattacharjee et al. (2016) have identified that the thicknesses of five structural components at the vehicle front significantly affect crashworthiness. These thicknesses are treated as the primary design variables.

The MOOP is formally stated as:

$$\min F(x) = [\text{Intrusion}, A_{in}, \text{Mass}]^T$$

subject to the constraint:

$$1 \text{ mm} \leq x_n \leq 3 \text{ mm}, n = 1, 2, \dots, 5$$

where the design variable vector is:

$$x_n = [t_1, t_2, t_3, t_4, t_5]^T$$

• Results and discussion

Effect of infill strategy on intermediate Pareto front solutions:

The following Figures show the infill strategy followed for the current MOOP problem with some sample intermediate plots. In these Pareto- Front Scatter plots, red dots denote the original solutions and blue dots denote the new points generated by inverse design-based infill strategy.

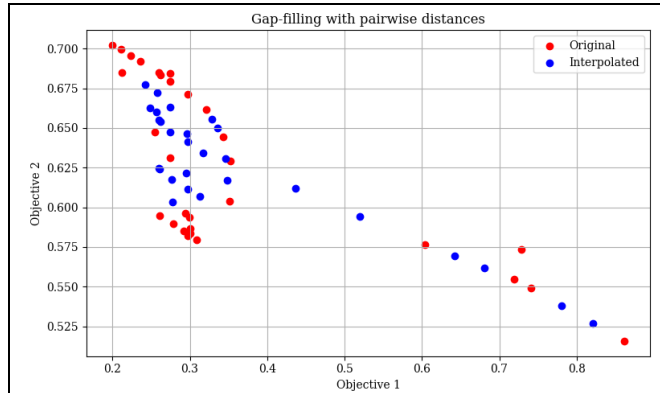


Figure 1: Modified Pareto front for objectives f_1 and f_2 with and without inverse design strategy

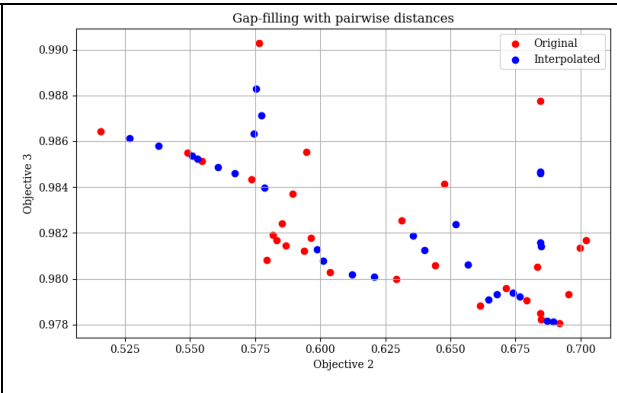
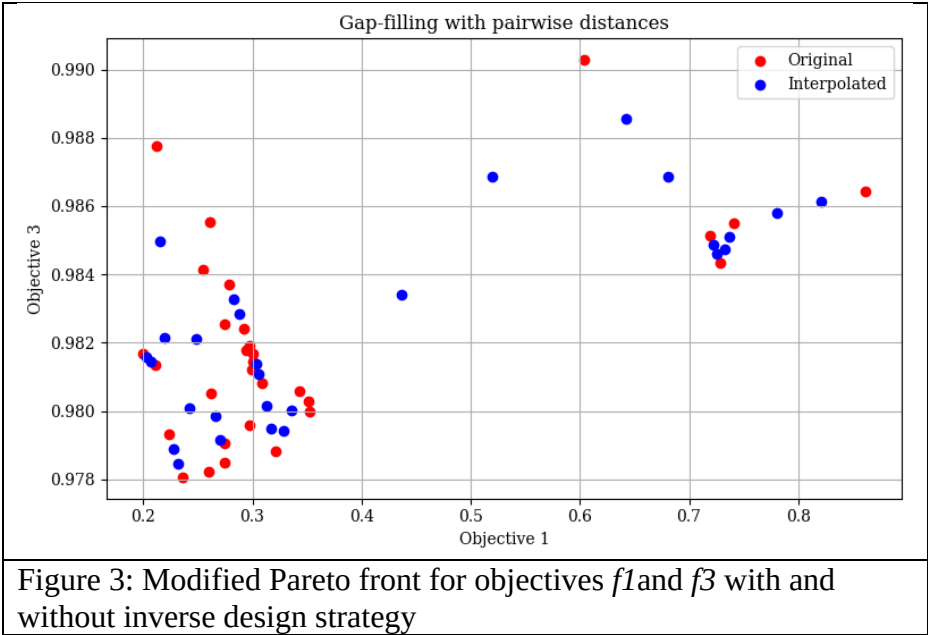


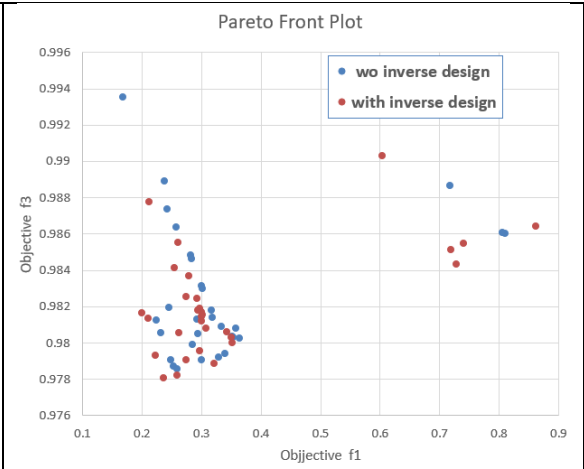
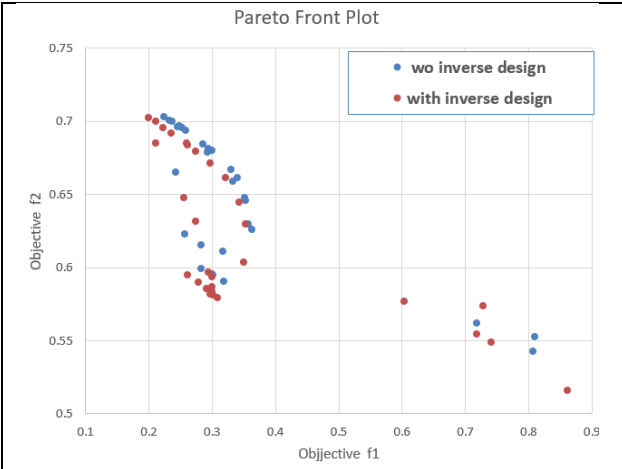
Figure 2: Modified Pareto front for objectives f_2 and f_3 with and without inverse design strategy



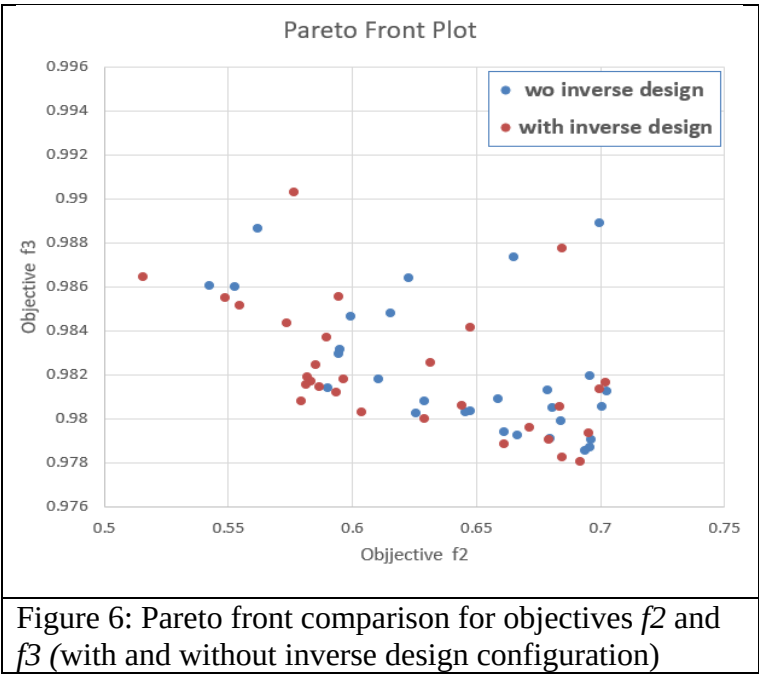
From the above figures, it can be clearly observed that the inverse design-based infill strategy helps in getting the uniform spread of Pareto Front with well diversified solutions.

Pareto front solutions comparision:

The following Figures 4 to 6 compare the effect of inverse design infill strategy on the quality of final Preto Front for the current MOOP problem

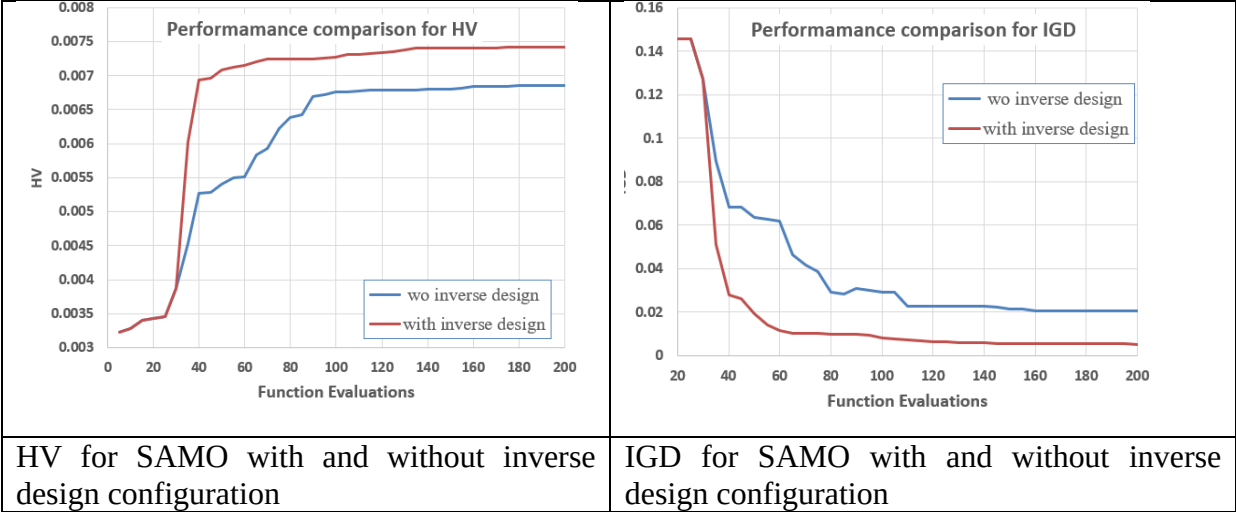


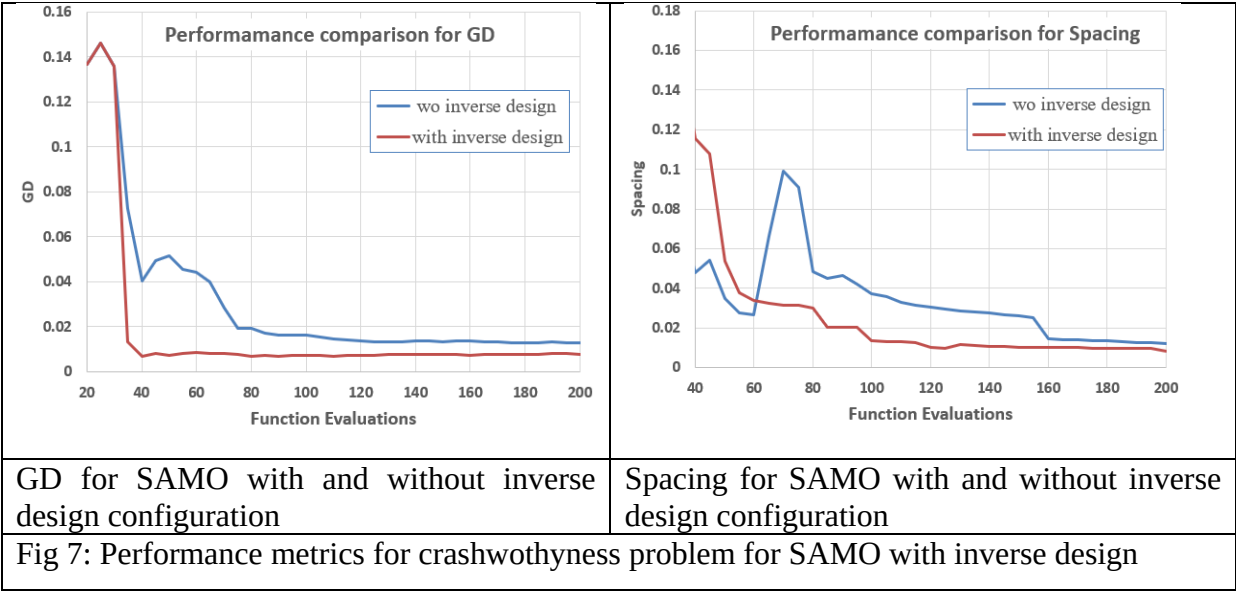
Figures 4 and 5 show the Pareto front comparison for objectives f_1 Vs f_2 and for f_1 Vs f_3 (with and without inverse design configuration) at 200 costly function evaluations.



Figures 6 shows the Pareto front comparison for objectives f_2 Vs f_3 (with and without inverse design configuration) at 200 costly function evaluations.

- 1) From the study of these Figures, following observations can be made
- 2) The Parereto front obtained from SAMO integrated with inverse design infill strategy has leftward shift as compared to baseline SAMO. This ensures the better convergence.
- 3) The Pareto front obtained by SAMO with inverse design is relatively uniformly spaced
- 4) The over range of the Pareto front obtained by SAMO with inverse design is greater than the baseline SAMO
- 5) All these factors ensure that the SAMO with inverse design is a better infill strategy which improves the overall quality of Pareto front solutions of MOOPs.





Following observations can be made from the study of these Figure 7

1. For HV parameter, the SAMO with inverse design perofrms better than baseline SAMO as its value is higher for all the function evaluations.
2. Based on the IGD criteria, SAMO with inverse design performs better than the baseline
3. Based on the GD criteria, SAMO with inverse design performs better than the baseline
4. Based on the spacing criteria, SAMO with inverse design performs better than the baseline

Table 3 compares the HV, IGD, GD and spacing parameters for SAMO with and without inverse design as Pareto front infill strategy.

Table 3. Details of Performance Metrices for SAMO with NSGA3

	<i>GD</i>	<i>HV</i>	<i>IGD</i>	<i>Spacing</i>
<i>WO Inverse Design</i>	0.01277	0.00685	0.02056	0.01227
<i>With Inverse Design</i>	0.00787	0.00742	0.00504	0.00825
<i>% Diff</i>	-38.33%	8.31%	-75.50%	-32.74%

Effectiveness of Inverse Design Strategy

The results clearly demonstrate the **positive impact** of integrating the inverse design approach into the multi-objective optimization process. Key performance indicators of the Pareto front—**Generational Distance (GD)**, **Hypervolume (HV)**, **Inverted Generational Distance (IGD)**, and **Spacing**—show noticeable improvements:

- **GD decreased by 38.33%**, indicating the solutions are significantly closer to the true Pareto front.
- **IGD improved by 75.50%**, reflecting better distribution and convergence of solutions.
- **Spacing reduced by 32.74%**, suggesting improved uniformity across the front.
- **HV increased by 8.31%**, which shows a broader and more optimal coverage of the objective space.

These improvements validate that **inverse design enhances both the diversity and accuracy** of the solution set by intelligently filling gaps in the Pareto front, leading to a more representative and well-distributed set of optimal solutions.

• Conclusion and Future Scope

This study presented an enhanced NSGA-III-based surrogate-assisted multi-objective optimization (SAMO) framework by integrating an inverse design strategy aimed at intelligently filling sparse regions in the Pareto front. The inclusion of inverse design significantly improved solution quality across multiple performance indicators. Compared to the baseline (without inverse design), the method achieved a 38.33% reduction in Generational Distance (GD), a 75.50% improvement in Inverted Generational Distance (IGD), a 32.74% enhancement in Spacing, and an 8.31% increase in Hypervolume (HV). These results indicate that the inverse design technique not only generates better-distributed solutions but also helps uncover high-quality designs that may otherwise remain unexplored using traditional evolutionary strategies alone.

Looking forward, the integration of inverse design opens several promising research directions. Extending the approach to handle higher-dimensional objective spaces and exploring adaptive gap-filling mechanisms could further improve computational efficiency and scalability. Incorporating hybrid surrogate models and uncertainty quantification techniques may enhance the robustness of inverse predictions. Additionally, applying this methodology to real-world, high-fidelity design problems across automotive, aerospace, and structural engineering domains could establish its broader applicability. Overall, the proposed strategy presents a compelling step forward in data-efficient and intelligent multi-objective design optimization.

References

- 1) Dong, H., Li, J., Wang, P., Song, B. & Yu, X., 2021. Surrogate-guided multi-objective optimization using an efficient online sampling strategy. *Knowledge-Based Systems*, 226, 107161. DOI:10.1016/j.knsys.2021.106919.
- 2) Fuhrländer, M. & Schöps, S., 2021. Yield optimization using hybrid Gaussian process regression and a genetic multi-objective approach. *Advances in Radio Science*, 19, pp.41–48. <https://doi.org/10.5194/ars-19-41-2021>.
- 3) Han, J. & Zhang, F., 2023. An effective surrogate model assisted algorithm for multi-objective optimization: application to wind farm layout design. *Frontiers in Energy Research*, 11, 1054321. <https://doi.org/10.3389/fenrg.2023.1054321>.
- 4) Lartaud, P., Humbert, P. & Garnier, J., 2024. Sequential design for surrogate modeling in Bayesian inverse problems. *Inverse Problems*, 40(3), 035004. <https://doi.org/10.1088/1361-6420/ad45e7>
- 5) Pellicer, L., Escudero, A., Alzueta, M. & Deb, K., 2022. A three-step procedure to find, validate and fill Pareto-front gaps in multi-objective problems. *IEEE Transactions on Evolutionary Computation*, 26(4), pp.745–759. <https://doi.org/10.1109/TEVC.2022.3143307>.
- 6) Preen, R.J., Bull, L. & Ross, P., 2022. Surrogate-assisted evolutionary computation: recent advances and future challenges. *Swarm and Evolutionary Computation*, 74, 101025. <https://doi.org/10.1016/j.swevo.2022.101025>.
- 7) Ruiz, P., Mena, G. & Sainz, L., 2021. Gaussian Process Regression–Based Multi-Objective Bayesian Optimization for power system design. *Sustainability*, 13(4), 20021. <https://doi.org/10.3390/su13042003>.
- 8) Shi, K. et al., 2023. Process–structure multi-objective inverse optimization for additive manufacturing of lattice structures using physics-enhanced Gaussian Process models. *Virtual and Physical Prototyping*, 18(4), pp.789–807. <https://doi.org/10.1080/17452759.2023.2171234>.
- 9) Sun, Z., Li, Q. & Li, Y., 2022. High-fidelity surrogate-based multi-objective optimization algorithm: Ensembles of RBF and Kriging. *Algorithms*, 15(8), <https://doi.org/10.3390/a15080279>
- 10) Tran, A., Eldred, M., McCann, S. & Wang, Y., 2021. srMO-BO-3GP: A sequential regularized multi-objective constrained Bayesian optimization framework. *Journal of Mechanical Design*, 144(3), 031705. <https://doi.org/10.1115/1.4052445>.
- 11) Bhattacharjee, K.S., Singh, H.K. and Ray, T. (2016) ‘Multi-objective optimization with multiple spatially distributed surrogates’, *Journal of Mechanical Design*, 138(9), pp. 091401-1–091401-10. DOI: 10.1115/1.4033923.

- 12) Kadre, S. S. and Tripathi, V. K. (2019) ‘Engineering design optimization by the dynamic cluster-based framework using mixture surrogates’, *International Journal of Innovative Technology and Exploring Engineering*, 8(11S), pp. 979–988. ISSN: 2278-3075.
- 13) Kadre, S. S. and Tripathi, V. K. (2020) ‘Mixture surrogate models for multi-objective optimization’, *International Journal of Engineering Sciences*, 13(1), pp. 21–28. ISSN: 0976–6693.
- 14) Kadre, S. S. and Tripathi, V. K. (2016) ‘Advanced surrogate models for design optimization’, *International Journal of Engineering Sciences*, 9(2), pp. 66–73. ISSN: 0976–6693.
- 15) Kadre, S. S. and Tripathi, V. K. (2020) ‘Multi-objective optimization of vehicle crashworthiness by SAMO: Advanced surrogate model-based algorithm’, *Journal of Advanced Research in Dynamical and Control Systems*, 12(6), pp. 2278–2288. ISSN: 0193-4120.
- 16) Kadre, S. S., Tripathi, V. K. and Tan, H. H. (2021) ‘Refined advanced surrogate assisted multi-objective optimization algorithm (RASAMO)’, *Elementary Education Online*, 20, pp. 2374–2388. doi: 10.17051/ilkonline.2021.05.259.
- 17) Kadre, S. S. and Paspula, R. G. (2023) ‘Improved RUL predictions of aero-engines by hyper-parameter optimization of LSTM neural network models’, *International Journal for Research in Applied Science and Engineering Technology (IJRASET)*, 11(3).
- 18) Zhang, J., Bi, S. & Zhang, G., 2020. Directional Gaussian smoothing optimization for computational inverse design in nanophotonics. *Optics Express*, 28(23), pp.34567–34582. <https://doi.org/10.1364/OE.400763>.
- 19) Zhu, Y. et al., 2024. Pareto front-diverse batch multi-objective Bayesian optimization (PDBO): Ensuring diversity and quality in Pareto solutions. *Artificial Intelligence*, 38, pp.5678–5686. <https://doi.org/10.1609/aaai.v38i8.5643>.
- 20) Liu, X. et al., 2025. Surrogate-assisted global and distributed local collaborative optimization for expensive constrained problems. *Scientific Reports*, 15, 1728. <https://doi.org/10.1038/s41598-025-85233-6>.
- 21) Zhuo, L. & Chen, X., 2025. Multi-objective optimization in machine learning–assisted materials design and discovery. *Journal of Materials Informatics*, 12(1), pp.45–67.
- 22) Li, B., Lu, Y., Qian, H., Hong, W., Yang, P. & Zhou, A., 2024. Regularity model-based surrogate-assisted evolutionary algorithms for expensive multi-objective optimization. *Swarm and Evolutionary Computation*, 86, 101506. <https://doi.org/10.1016/j.swevo.2024.101506>.

Machine learning-based analysis of clinical data in hemodialysis patients

Pavani Peddu*, Poonam Sharma*, Anand Chellapan†

*School of computer science and engineering,

Visvesvaraya national institute of technology (VNIT), Nagpur, Maharashtra, India

Emails: peddupavani@gmail.com, poonamsharma@cse.vnit.ac.in

†Nephrology department,

National institute for advanced medical education and research, Nagpur, Maharashtra, India

Email: anandchellapan@aiimsnagpur.edu.in

Abstract:

Artificial kidney treatment is a medical procedure which is used for individuals whose kidneys have failed or who can no longer effectively filter waste products, toxins from the blood, and excess fluids. This procedure works by removing the patient's blood and passing it through a dialyzer, a machine that functions as an artificial kidney. When a person is going through hemodialysis treatment it is very important to predict their Dry Weight correctly. Making the wrong prediction can cause major health problems like low blood pressure, too much fluid, cramps, or even heart difficulties. Doctors experience and manual observations are the important ways that people have traditionally found Dry Weight. These methods may not always give accurate prediction of dry weight. This research offered multiple machine learning models to predict the Dry Weight in hemodialysis patients using clinical and demographic data. At AIIMS Nagpur, the data for this research was collected. It has 517 records of dialysis sessions which include vital details like systolic and diastolic blood pressure, weight before and after dialysis, fluid removed, body temperature and pulse rate. There are various kinds of machine learning models like multiple linear regression, decision trees, random forests, XGBoost, and support vector regressors (SVR). In addition, more complex methods are used, such as domain-based stacking and granular bell membership in SVR and the grey wolf optimization which is used to fine-tune the hyperparameters of the XGBoost model. The SVR with a granular bell function is the best of all the models since it is giving the minimum error and the highest R2 score. This study illustrates that machine learning models can offer accurate and rapid estimates of Dry Weight, thereby supporting clinicians and improving the efficiency of the dialysis procedure for patients.

Index Terms—Chronic kidney disease (CKD), Dry Weight estimation, Hemodialysis, machine learning, regression models, bioimpedance spectroscopy, inferior vena cava (IVC).

I. INTRODUCTION

Dialysis is a vital treatment for individuals whose kidneys can no longer effectively filter waste and excess fluids from the blood. It removes waste, extra fluid as well as keeps important minerals in balance when the kidneys can't do the job. The two primary forms of dialysis are hemodialysis, which relies on a machine to cleanse the blood, and peritoneal dialysis, which uses the abdominal lining as a natural filter. The most suitable option varies based on an individual's medical condition and lifestyle. Some of the most crucial aspects of dialysis treatments are preserving the proper amount of fluid. People with compromised renal function cannot normally get rid of extra fluid, so water, salt, as well as additional substances accumulate in the body. Left unremoved, liquid overflow can cause lung damage, heart failure from congestive hypertension, and an elevated fatality rate. Therefore, to preserve patients and prevent both immediate and lasting problems, it is imperative to exactly determine the optimal quantity of fluid to be removed after every dialysis meeting. The result represents the idea of "Dry Weight" defined as the weight of the individual after dialysis treatments while all the extra fluid has been fully eliminated. The person who needs treatment is usually believed to be in a euvolemic state—this means not overhydrated and not dehydrated either. In clinical settings Dry Weight is used to decide how much fluid to remove in each dialysis session, a process called ultrafiltration. But finding the right Dry Weight is not straightforward at all; it is quite hard and also very personal. Many things can affect it, like changes in body shape, inflammation, food intake, or other health problems. That's why it is very important to assess and update the Dry Weight again and again, because fluid balance plays a big role in how long a patient survives and their disability risk. In the past, this kind of checkup mostly depended on doctors' experience, what patients said, and trying different changes—but all these can be random and have chances of

mistake. To get better accuracy in predicting Dry Weight in hospitals, we need to move towards using more data-driven, clear, and personalized methods such as those offered by machine learning. This paper focuses on using machine learning to help predict the right Dry Weight for hemodialysis patients by looking at their health details, background, and treatment information.

II. WORK DONE

Accurately predicting a patient's dry weight (DW) at the time of hemodialysis is essential for Managing proper fluid balance and avoiding serious problems such as low blood pressure, cardiovascular problems and increased risk of death [1]. Dry weight is the point at which a person's body has just the right amount of fluid after dialysis—without showing signs of too much or too little fluid. Traditionally clinicians use to estimate Dry Weight based on experience and physical signs but these methods can be Unreliable and based on personal judgment [2]. To support more reliable assessments, advanced techniques have been introduced. One such method is bio impedance spectroscopy (BIS), which measures how electrical signals pass through the body to estimate fluid levels. BIS has been represented to decrease the fluid accumulation as well as support better heart function in people who are receiving dialysis treatment [3, 4] and other methods like lung ultrasound (LUS), which detects fluid in the lungs and inferior vena cava (IVC) diameter measurements also offer valuable insights, though their effectiveness can vary depending on the user's skill and patient conditions [5, 6]. In recent years, machine learning (ML) has gained consideration for its prospective to improve Dry Weight prediction. Algorithms such as Tree-structured ensemble methods and kernel-based learning algorithms can analyze complex patient data and identify patterns that traditional methods may miss [7]. After combined with real time data from smarter devices like wearable sensors tracking bio impedance these models might help deliver adaptive and needs based treatments [8]. Machine learning models have been used to evaluate dialysis adequacy. For example in one study analyzing over 1,300 dialysis sessions, models like XGBoost and the Convolutional Neural Networks (CNN) models were used to predict the Urea Reduction Ratio (URR) a key indicator of dialysis effectiveness. XGBoost performed especially well underlining the importance of factors such as the pre-dialysis weight [30]. Artificial intelligence has also showed promise in treating anemia in people on dialysis. The model that changed the doses of erythropoiesis stimulating agents (ESAs) [31] based on a patient's history led to a 27% drop in ESA use better hemoglobin management and fewer blood transfusions. In subsequent examinations, more than 83% of patients continued hemoglobin levels within the desired range and enhancing both treatment expenses and the results [32]. In Japan researchers used machine learning models on a dataset of almost 70,000 dialysis sessions to figure out when a patient's dry weight needed to be changed. Their models effectively identified indicators of fluid imbalance, such as hypotension and inflammatory markers. The study did not examine the influence of characteristics such as gender on predictions, indicating a need for more research [39].

III. OBJECTIVE

Implement different machine learning models to predict the target body weight (Dry Weight) of hemodialysis patients and find the one that gives the best predictions based on the evaluation metrics.

IV. DATA SET DESCRIPTION

Dr. Anand Chellappan oversaw the collection of the dataset from the AIIMS Nagpur Nephrology Department. AIIMS Nagpur, established in July 2014, is a government-supported institution focused on advanced medical education and research. Located in Nagpur, Maharashtra, near the International Cargo Hub and Airport, it is part of the Phase-IV expansion of India's top medical institutes. The Indian government runs it, and it is a development that is part of the ministry of health and family welfare. There are 517 records in the dataset used for this study, and each record has a unique dialysis session that goes with it. The sessions came from the 33 different patients, which provided a comprehensive overview of demographic and clinical features essential to hemodialysis

care. There are 358 male patients, 157 female patients, and 2 records in the collection that don't have gender information. The age distribution of the patients shows that 15 of them are under 18 years old, which means they are children. Most of the 467 patients are adults between the ages of 18 and 60, whereas 35 patients are over 60 years old, which is the age category for older people. This dispersion across all age groups lets us look at how Dry Weight prediction changes over time.

The dataset collected for this research incorporates dialysis session related information, including session specifics, frequency of dialysis, treatment efficacy (Kt/V) as well as technical parameters such as rate of blood flow, rate of dialysate flow, sodium concentration and the temperature. These things are very important to make sure that each dialysis treatment is done safely and well, such as making sure the circumstances are right and the procedures are good. To see how well the dialysis procedure keeps the patient's fluid balance and helps them reach their post-dialysis target weight, we also keep track of their weight and fluid management techniques, such as InterdialyticWeightGain, UltrafiltrationAchieved, and Dry Weight.

Before and after the haemodialysis therapy, your body temperature, rate of respiration, diastolic blood pressure and systolic blood pressure are all recorded. All of these signals help you learn about the patient's physical condition during the session. In addition, blood pressure readings like Intradialytic SBP and DBP are taken at different times during the dialysis session to keep track of changes in blood pressure and see how the heart responds to the treatment.

The dataset has a "VascularAccess" field that shows what kind of access was used for each hemodialysis session. The data has 28 features, which makes it a good base for analytical modeling. For instance, the dataset can be utilized to estimate a patient's dry weight, which serves as an important indicator for assessing the adequacy and effectiveness of dialysis therapy.

V. IMPLEMENTATION

A. Data Cleaning and transformation

One important part of this study is working with patient compliance data to get it ready for machine learning. This involves checking if the data is consistent, cleaning up any errors, and reshaping it into a format that the models can understand and use effectively.

1) Imputing missing data in numerical features: It is common in datasets to have missing values for numerical columns. If you do not handle them, they can lead to biased results or errors at the time of training a model. These gaps are filled in using different imputation methods that keep the overall structure and distribution of the data. Mean imputation is one of the most common ways to deal with missing numbers. This method fills in the values missing by averaging of the non missing entries in the same column. This method works best when the data is evenly spread out and doesn't have any big outliers. It is also quick and easy to use.

2) Handling categorical columns: When categorical columns contain missing values, it can lead to poor data quality and impact model accuracy. Imputation methods used fill in missing entries correctly because categorical data unlike numbers have distinct labels. Mode imputation is one of the easiest and most common methods. It fills in missing values in the categorical column with the category that happens most repeatedly. When one category dominates the column and stays pretty stable, this method works well.

One hot encoding is one of the most common ways to change categorical features into a format that can be used by ML algorithms. One hot encoding changes each categorical value into a binary vector, with each new category becoming a new column. In these new columns, the column that matches the category gets a 1 and all the other columns get 0s.

3) Normalization: We use StandardScaler to standardize mathematical properties after imputation. This adjustment is very helpful for applications that respond to the feature scale. It adjusts the attributes so that the median is 0 and the average deviation is 1.

4) Removing outliers: DBSCAN is a density-based grouping method that finds outliers by looking at their clusters with few data points and density in the attribute space. It finds various groups in the data space by looking for dense areas that are separated by low-density areas. DBSCAN works well with real-world data problems, unlike K-means or a hierarchy of clustering, which assumes that clusters are light and elliptical.

5) Correlation for finding important features: Correlation is a means to see how closely two things are related to each other. Correlation is helpful for choosing features since it reveals how closely each feature is related to the target variable or how similar features are to each other. Pearson correlation Pearson's correlation is a statistical measure that explains how two continuous variables move in relation to each other along a straight line. The coefficient values ranges from -1 to +1 where

+1 : It shows a strong positive linear link, where both variables increase together.

0 : means there's no evident linear connection.

-1 : shows a strong negative linear link, where one variable increases as the other decreases.

6) Data splitting: A big aspect of machine learning is splitting out the datasets. This involves breaking the dataset up into smaller groups, like sets for training, testing, and validating. This approach makes sure that the model learns well and that the performance evaluation is correct. Training set: Training set is used during the training of the model. We used 80% of the data to train. Validation set: Tuning hyper parameters and reducing the chance of overfitting. We used 10% of the data for validation. Test set: Used to test how well the model is working and we used ten percent of the data for the test.

B. Methodologies

To build the model, you need to use the right methods to make prediction of Dry Weight. In this paper different machine learning models were implemented to discern patterns and provide predictions from a rather limited dialysis dataset. The goal was to build models that work well no matter how big the dataset is. During the development phase, the most suitable models were chosen, trained, and fine-tuned to attain optimal outcomes.

1) Multiple linear regression including interaction effects: Multiple linear regression (MLR) [10] represents that how a dependent variable is associated to more than one independent variable. Adding interaction terms [24] helps to show how two or more predictors work together to affect the target variable in a manner that cannot be captured by the individual influence of each predictor alone.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \dots + \beta_p x_p \quad (1)$$

Where:

- β_0 : Represent the constant or intercept of the model.
- $\beta_1, \beta_2, \dots, \beta_p$: Represent are the weights assigned to each corresponding predictor variable.
- y : The output variable .
- $x_1, x_2, \dots, x_p$: Input features.
- β_3 : It captures the combined influence of x_1 and x_2 , reflecting their interaction.

The interaction term $x_1 x_2$ shows the slope of x_1 fluctuate based on the value of the x_2 , which makes it easier to find more complex correlations between input features.

Huber loss function:

$$L_{\delta}(a) = \begin{cases} \frac{1}{2}a^2 & \text{for } |a| \leq \delta \\ \delta\left(|a| - \frac{1}{2}\delta\right) & \text{for } |a| > \delta \end{cases}$$

Where:

- $a = y_i - \hat{y}_i$: It reflects the difference between the predicted ($\hat{y}_i$) and the observed (y_i) values of the model for every instance.

- δ : A user-defined threshold that determines the transition between quadratic and linear loss behavior

2) Updated Decision Tree: A decision tree (DT) [25] is a type of supervised learning method applicable to both classification and regression tasks. The whole dataset is the root node, and the rest of the data is arranged in a treelike structure. Next, we divided it into branches based on how essential each feature is. We use internal nodes to make decisions and leaf nodes to produce the final prediction. The suggested regularized [15] choices tree regression method finds the best split at each node. Here we included fair split also [21].

Score = MSEL+MSER–MSET + λ • depth+ γ • $|\overline{y}_L - \overline{y}_R|$

(3) MSEL: The mean squared error of predictions in the left child node after splitting.

- MSER: The mean squared error of predictions in the right child node after splitting.

- MSET : Mean squared error for the parent (total) node before the split.

- λ : A coefficient that adds a penalty for increasing depth of the tree.

- depth: Indicates how many splits far from the current node is from the root node.

- γ : A factor used to reward the splits that produce wellseparated outputs.

- $|\overline{y}_L - \overline{y}_R|$: The absolute difference between the average target values of the left and right nodes encourages splits that maximize output separation.

Algorithm 1 Balanced Split Decision Tree for Regression Tasks

Input: Dataset $S = \{(z_i, t_i)\}_n$

$i=1$

Parameters: max depth, min samples split, min samples leaf, num bins, λ , γ , α , level, panality

function GROWTREE(S , level)

if level $\geq$ max depth or $|S| <$ min samples split then return terminal node with output $\hat{t} = 1$

$|S|$ P

t_i (average target)

end if

Compute MSEtotal

Initialize best score $\leftarrow \infty$, best split $\leftarrow$ None

for all attribute a do

Bin a into num bins intervals

for all threshold τ (midpoints of bins) do

$SL \leftarrow \{(z_i, t_i) \mid z_i[a] \leq \tau\}$

$SR \leftarrow \{(z_i, t_i) \mid z_i[a] > \tau\}$

if $|SL| <$ min samples leaf or $|SR| <$ min samples leaf then

continue end if

Compute MSEL and MSER

Compute fairness penalty: Penalty $\leftarrow \gamma \cdot |\overline{t}_L - \overline{t}_R|$

TotalScore $\leftarrow$ MSEL +MSER –MSEtotal + $\lambda \cdot$ level + Penalty

if TotalScore $<$ best score then best score $\leftarrow$ TotalScore

best split $\leftarrow (a, \tau)$

end if

end for

end for

```
if best split is None then return leaf node with  $\hat{t} = 1$   
|S| P  
ti  
end if  
Divide S into SL, SR using best split  
left node  $\leftarrow$  GROWTREE(SL, level + 1)  
right node  $\leftarrow$  GROWTREE(SR, level + 1)  
return internal node with best split, left node, right node  
end function  
function MAKEPREDICTION(z, T )  
current node  $\leftarrow$  root of T  
while current node is not a terminal node do  
if  $z[a] \leq$  split value at current node then  
current node  $\leftarrow$  left childSS  
else  
current node  $\leftarrow$  right child  
end if  
end while  
return prediction  $\hat{t}$  at terminal node (average of leaf )  
end function
```

3) Domain Random forest using Staking: Random Forest (RF) for regression [11] is an ensemble learning technique that creates many decision trees and averages their outputs to enhance accuracy and stability in predictions. Every tree produces its own guess, and the final answer is the average of all those guesses. This makes the guesses more reliable and less likely to be impacted by noise or overfitting because each tree uses random samples of data and characteristics from the original dataset.

Domain knowledge [13] is used in medical field to make machine learning models like Random Forests work better, more reliable and more useful. These models become more like how healthcare is done in the real world when they include clinical guidelines, expert opinions, and medically proven relationships. For example only pre-dialysis vitals are used in dialysis prediction models, not lab values that are affected by the treatment process. Domain formulas can be used to find things like BMI or ultrafiltration rate, which makes the model is predictions more useful. Domain knowledge helps to choose only the most important features, stay away from post treatment variables and include clinically useful thresholds like blood pressure levels that show hypertension. This makes it easier for doctors to understand, remove overfitting to data patterns that do not matter and makes sure that decisions are made in an ethical and informed way. In general, domainbased Random Forests in healthcare make models that are easier to understand, safer, and more like how doctors think. This makes them more useful and acceptable to doctors and other medical professionals.

Step 1: Grouping the features Features are grouped based on how important they are in a certain situation or for medical purposes. Word 2 Vector and other methods are used to change every component description into an underlying semantic vector data if needed. You can use clustering method as DBSCAN to group highlights that are statistically or linguistically similar.

A set of features as examples

Vitals signs: Features like heart rate, body temperature, and blood flow.

Demographics: Features like age, height, and race.

Clinical: This includes medical related features like blood pressure or the time taken for the dialysis session.

Step 2: Train separate models for each group Separate Random Forest models are built for each domain, based on their respective feature sets;

- One model learns from vital signs.
- The other one solely looks at demographic data.
- The third one on the other hand is about clinical information.

Step 3: Get predictions from each model Each of the three models is used to predict results using only the data it was trained on. So, the vitals model uses vitals data and the demographics model uses demographic data, and so on.

Step 4: Combine those predictions [?]

Next, take the predictions from all three models and put them merge like new features. After combining the outputs, a higher-level model (meta-model) is trained for final prediction that learns from the combined predictions to make even better decisions.

Step 5: Final prediction Finally implemented the meta-model to give final output based on what ever it learned from the earlier models. 4) Support vector regressor using granular clustering: Support Vector Machines (SVM) can be extended to regression tasks through a method known as Support Vector Regression (SVR) [14]. It aims to find a function that accurately predicts output values for given input values. SVR can utilize either a linear or a non-linear kernel. The dot product of two input vectors makes up a linear kernel. A non-linear kernel is a more advanced function that can detect more complicated patterns in the data. The type of data and how demanding the job is will help you choose the right kernel. We use something called granular bells to combine Support vector regression (SVR) and granular bells [19] , this makes

the model easier to understand and better at dealing with data that isn't clear or is noisy. We use DBSCAN, a method for clustering to put similar data points together. This allow SVR learn from these groups and make better predictions.

Algorithm 2 Granular clustering-based support vector regression Require: Feature set $X = \{x_i\}_N$

$i=1$, target values $Y = \{y_i\}_N$

$i=1$

Initialize the parameters related DBSCAN: ϵ_{init} , minPts

Refinement DBSCAN parameter: ϵ_{refine}

Error threshold: τ_{error} , max iterations: T_{max}

Initial Clustering: Run DBSCAN on X with ϵ_{init} , minPts

for each cluster C_k do

Calculate cluster centroid: $m_k \leftarrow \frac{1}{|C_k|} \sum_{x_i \in C_k} x_i$

$|C_k|$

$\sum$

$x_i \in C_k$

x_i

Calculate average target: $\bar{y}_k \leftarrow \frac{1}{|C_k|} \sum_{x_i \in C_k} y_i$

$|C_k|$

$\sum$

$x_i \in C_k$

y_i

end for

Initialize centroid set $M \leftarrow \{m_k\}$, target set $Y_{avg} \leftarrow \{\bar{y}_k\}$

for $t = 1$ to T_{max} do

Train SVR model: $f(t) \leftarrow \text{SVR.fit}(M, Y_{avg})$

Predict outputs: $\hat{y}_i \leftarrow f(t)(x_i)$ for all $i = 1, \dots, N$

Calculate squared errors: $e_i \leftarrow (y_i - \hat{y}_i)^2$

for each cluster C_k do

Compute mean cluster error: $E_k \leftarrow \frac{1}{|C_k|} \sum_{x_i \in C_k} e_i$

$|C_k|$

$\sum$

$x_i \in C_k$

e_i

```

end for
Identify clusters exceeding error threshold:  $S = \{C_k : E_k > \tau_{\text{error}}\}$ 
if S is empty then
break
end if
for each cluster  $C_s \in S$  do
Re-cluster  $C_s$  using DBSCAN with  $\epsilon_{\text{refine}}$ 
Update centroids and target averages for refined sub-clusters
end for
Update M, Yavg with new centroids and targets
end for
return Final SVR model  $f(t)$ , refined centroid set M, and target averages Yavg

```

5) XGBOOST using gray wolf optimization : XGBoost (Extreme Gradient Boosting) [22] is a A rapid and effective machine learning algorithm which can be applied to both predictive as well as categorical tasks. It makes a set of decision trees, and each one tries to rectify the faults that the trees before it made. This is called boosting. XGBoost is different from other programs since it uses smart techniques like:

- Gradient Descent to reduce errors
- To stop overfitting use regularization
- Parallel computing to make computation go faster

The Grey Wolf Optimizer (GWO) [] draws its design from the cooperative hunting and leadership patterns seen in grey wolf packs. Grey wolves are the top predators in their habitats There is a set order of power in each pack. The α wolves are the dominant members that lead and coordinate the pack's activities. and tell everyone else what to do. Beta β wolves are the alpha's advisors and are next in line to lead. Delta δ wolves can also work as scouts, hunters, or caregivers. The GWO algorithm is based on this social structure, with alpha, beta, and delta wolves leading the way and the other wolves following. This lets you strike a balance ensuring a strategic mix of exploring diverse options and concentrating on the best candidates during the optimization stage. Delta wolves can also have certain jobs, such as being scouts, hunters, or caregivers. The GWO algorithm is built on this social hierarchy, with the strongest solutions being led by alpha, beta, and delta wolves with the others following. This approach helps maintain equilibrium between innovation and improvement during the optimization cycle. Algorithm 3 Grey wolf optimizer-based hyperparameter tuning for XGBoost model enhancement Require: Dataset D, population size N, maximum iterations T Ensure: Optimized hyperparameter vector Θ^*

```

Initialize a population of candidate solutions  $\{\Theta_i\}_N$ 
i=1 randomly within the search
space
Evaluate fitness (e.g., validation error) of each  $\Theta_i$  by training XGBoost with  $\Theta_i$ 
Identify top three leaders: alpha ( $\Theta_\alpha$ ), beta ( $\Theta_\beta$ ), and gamma ( $\Theta_\gamma$ ) based on best
fitness
for t = 1 to T do
Update control parameter  $a \leftarrow 2 - 2t$ 
T
for each candidate solution  $\Theta_i$  do
Generate random vectors  $r_1, r_2 \sim U(0, 1)$ 
Compute coefficient vectors:
 $F_k = 2a \cdot r_1 - a$ 
 $R_k = 2 \cdot r_2$  for  $k \in \{1, 2, 3\}$ 
Update candidate position using leader influence:
 $\Delta\alpha = |R_1 \odot \Theta_\alpha - \Theta_i|$ ,  $\hat{\Theta}_\alpha = \Theta_\alpha - F_1 \odot \Delta\alpha$ 
 $\Delta\beta = |R_2 \odot \Theta_\beta - \Theta_i|$ ,  $\hat{\Theta}_\beta = \Theta_\beta - F_2 \odot \Delta\beta$ 
 $\Delta\gamma = |R_3 \odot \Theta_\gamma - \Theta_i|$ ,  $\hat{\Theta}_\gamma = \Theta_\gamma - F_3 \odot \Delta\gamma$ 

```

$$\Theta_{new}^i = \frac{\alpha + \Theta\beta + \Theta\gamma}{3}$$

R is the distance between current solution and each leader Θ is the new position influenced by his distance and randomness;
final position is the average of movements toward all three leaders Clip or round Θ_{new}^i to ensure valid hyperparameter values Recalculate fitness of Θ_{new}^i using XGBoost end for Update leader wolves ($\Theta\alpha$, $\Theta\beta$, $\Theta\gamma$) based on updated fitness scores end for return $\Theta\alpha$ (best hyperparameter configuration found)

VI. RESULT DISCUSSION AND PERFORMANCE ANALYSIS

A detailed analysis of different regression models was done using a number of performance metrics to see how well they are working and the different optimized models being looked at were decision trees, random forest , XGBoost, support vector regressors (SVR) , and multiple linear regression. We looked at these models error metrics, fit statistics, computational efficiency and model complexity. Beginning with the The average of the absolute differences

between predicted and actual values (MAE), SVR demonstrated the highest accuracy with the lowest error (0.0036),

Metric	Regression	Decision	Random	SVR	XGBoost
MAE	2.8823	0.9347	0.7774	0.0036	0.3669
MSE	32.2348	4.2360	2.8005	0.0494	0.3400
RMSE	5.6776	2.0582	1.6735	0.2223	0.5831
R2 Score	0.8351	0.9736	0.9745	0.9986	0.9966
Adjusted R2	0.7362	0.9705	0.9712	0.9978	0.8869
MAPE (%)	24.0689	2.4002	1.7868	0.0012	0.8821
SMAPE (%)	8.3177	2.3125	1.7420	0.0014	0.1520
Median AE	1.6226	0.0200	0.4313	0.0012	0.0270
MBE	0.7569	0.5065	0.1062	-0.0036	0.9966
EVS	0.8380	0.5065	0.9746	0.9986	0.0452
RAE	0.2801	0.9752	0.1001	0.0494	0.0034
RSE	0.1649	0.0945	0.0255	0.0255	0.0179
AIC	341.3167	160.0308	110.6211	15.4118	-80.0181
BIC	413.1502	185.7779	135.3944	26.7108	-60.1091
Training Time (s)	6.1600	0.5006	78.2428	0.4764	476.3325
Prediction Time (s)	0.0000	0.0100	0.0334	0.0012	0.0147
Model Size (KB)	30.9043	30.9043	3501.0293	22.887	90.4912

Table 1: Performance metrics of all models

indicating that its predictions were remarkably close to the actual values. XGBoost also performance is well (0.3669), followed by Random Forest (0.7774) and decision Tree and Linear Regression exhibited slightly greater errors (0.9347 and 2.8823, respectively), suggesting less precise predictions. Moving on to the Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), SVR again achieved the lowest values (MSE: 0.0494; RMSE: 0.2223), emphasizing its ability to minimize large deviations. XGBoost and Random Forest showed fairly good performance, while Decision Tree and Linear Regression reported higher squared errors, indicating their inferior robustness in handling this dataset. The R2 value, which shows how effectively the model captures variance within the observed data, also showed that

SVR was better, with a value of 0.9986, which is almost ideal. XGBoost (0.9966) and Random Forest (0.9745) came in second and third while Decision Tree (0.9736) and Linear Regression (0.8351) came in last. The Adjusted R2, which favors simpler models by applying a penalty to the complex ones, gave comparable findings. SVR came in to the lead, although XGBoost fell second, probably it is an ensemble model.

In terms of relative error metrics, the Mean Absolute Percentage Error (MAPE) and Symmetric MAPE (SMAPE) reaffirmed SVR's accuracy, with extremely low values (MAPE: 0.0012%; SMAPE: 0.0014%). XGBoost remained a strong contender here, whereas the other models, particularly Linear Regression, exhibited noticeably higher

relative error percentages. When looking at Median Absolute Error (MedAE) and Mean Bias Error (MBE), SVR continued to stand out with minimal bias and consistent predictions. XGBoost and Random Forest also showed stable behavior, while Linear Regression had higher central errors and bias, implying over- or underestimation trends. The Explained Variance Score (EVS) presented an interesting contrast—while SVR and Random Forest maintained high variance explanation capabilities, XGBoost's score dropped significantly, possibly due to specific data characteristics or noise sensitivity. Nonetheless, it remained strong across other Further, Relative Absolute Error (RAE) and Relative Squared Error (RSE) provided insight into how each model fared compared to a baseline predictor. XGBoost and SVR had the lowest error ratios, signifying strong generalization. Random Forest showed decent performance, but Decision Tree and Linear Regression were less reliable in this context.

In terms of model complexity, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) revealed that XGBoost, despite being more resource-intensive, provided an optimal trade-off between accuracy and complexity. SVR also have low AIC and BIC values while the Decision Tree and Linear Regression appeared over simplified and less effective under these criteria. From a computational standpoint, Linear Regression was the fastest to train, taking just 6.16 seconds. SVR and Decision Tree were also efficient in training. But Random Forest and XGBoost took a lot longer to train because they used ensemble methods. Even still all of the models had short prediction periods and SVR and Linear Regression gave findings almost right away, while Random Forest was a little later. Finally, in terms of model size SVR and Linear Regression were notably compact consuming under 35 KB. XGBoost occupied around 90 KB while Random Forest due to the large number of trees was considerably larger (around 3.5 MB) which may have implications for deployment in memory constrained environments. In summary SVR has the most balanced model which will offering excellent predictive accuracy and low error rates, minimal bias, and fastest computation. XGBoost closely followed slightly outperforming in some complexity sensitive metrics but at the cost of training time and size of model. Random Forest was a dependable mid ground, while Decision Tree and Linear Regression served as faster but less accurate alternatives.

This detailed analysis points to SVR and XGBoost as the top choices when accuracy and consistency are crucial for practical use.

Model Performance Metrics

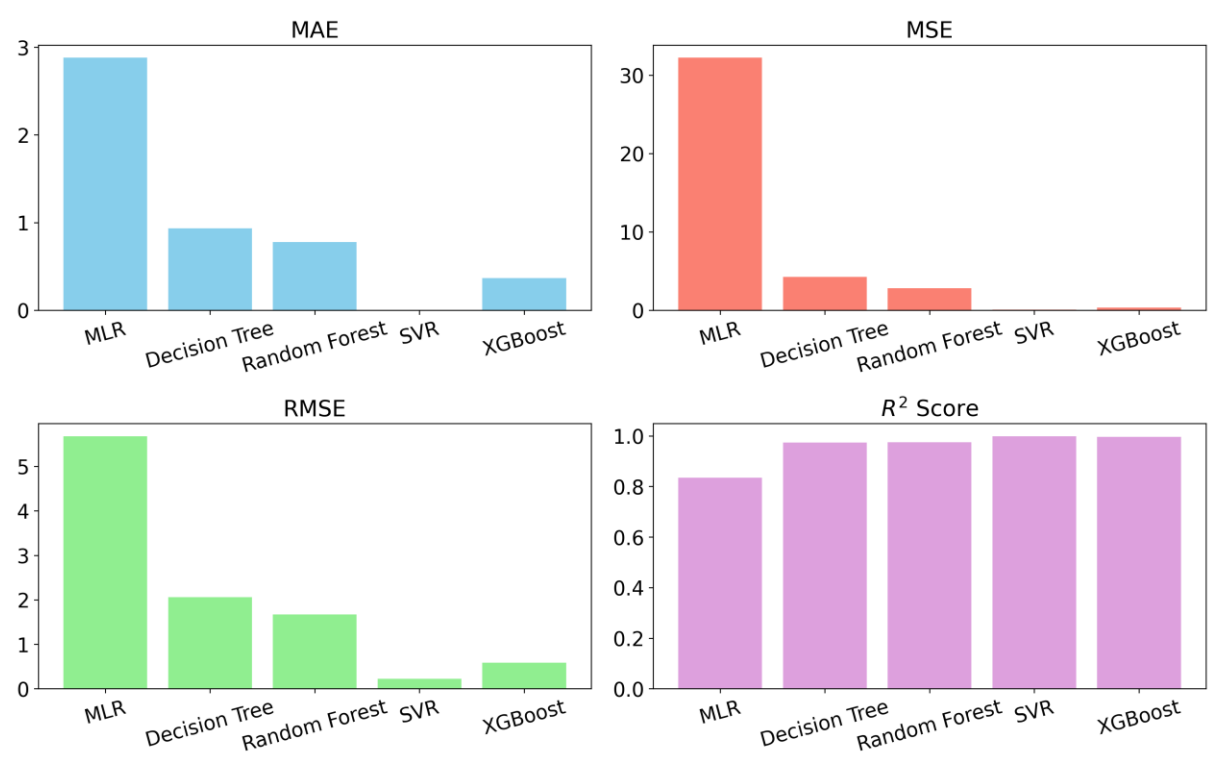


Fig. 1: SVR model performance for different number of features

We used four popular evaluation metrics to compare how well different regression models working Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R2 Score. These measures s how accurate and reliable each model’s predictions are. The results illustrated in Fig. 1, In case of MAE, SVR and XGBoost got the lowest values, which means their predictions was very close to the actual output on average. On the other hand, Multiple Linear Regression showed highest MAE, so its prediction are more off.

The MSE results also shows a similar pattern. Since MSE penalize large errors more, the small MSE values from SVR and XGBoost indicates that they didn’t make big mistakes. Again, Multiple Linear Regression performed worst with largest MSE.

RMSE gives error in original units of target variable. SVR and XGBoost done well here too, giving low RMSE values. Random Forest and Decision Tree did decent, but not that great. Multiple Linear Regression showed high RMSE, which is not much good.

R² Score tells the proportion of the target variable’s variation that is predicted by the model. SVR had the highest R2 around 0.9866, and XGBoost also had a good R2 after SVR. But Multiple Linear Regression got only 0.8353 so it not able to explain much variance in data. So overall, SVR and XGBoost are the best model among the tested ones. They gave better accuracy and consistency. Multiple Linear Regression was not suitable for this problem due to its poor results in all metrics.

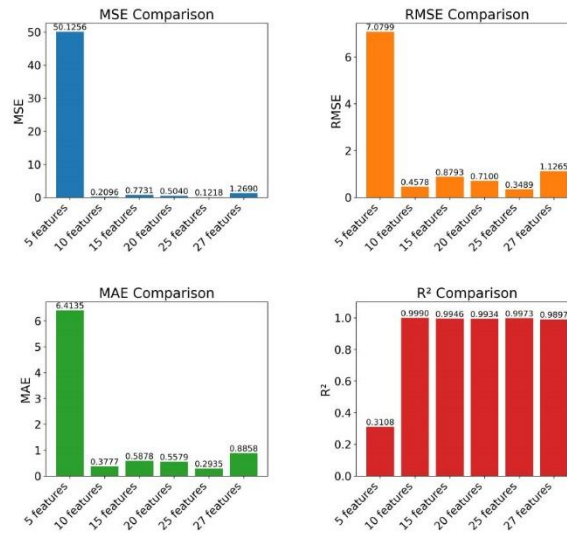


Fig. 2: SVR model performance for different number of features

We tested a strong model called Improved SVR by using granular clustering with different combinations of input features, from 5 to 30, to determine how well it could guess the dry weight of people on hemodialysis. The goal was to find out how the quantity of clinical features affects how well the model predicts. Some of the key approaches to verify how accurate and dependable the model were Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R²). These measures show us alternative ways to think about prediction error and which helps us figure out how the model works when the input gets more intricate.

When tested only with five features, the results illustrated in Fig. 2, the model error rate is high (MSE of 50.13, RMSE of 7.08, and MAE of 6.41) and a low R² value (0.3108) these results represents that the model doesn't fit well and that it did not capture sufficient variation. This kind of performance shows that the model isn't fitting well which means it doesn't have enough data to figure it out the patterns that control fluid balance in persons who are on hemodialysis..

Adding 10 more features made a big difference in performance. The model had a very less MSE of 0.2096, RMSE of 0.4578, and MAE of 0.3777, and its R² value was almost perfect at 0.9990. These results represents that adding more relevant clinical characteristics, such gender, blood pressure, and temperature, makes the model much better at generalizing. It indicates that these variables encompass critical information regarding the patient's fluid status and physiological responses to dialysis.

The model still worked well when the number of features was raised to 15 and 20 with the R² values of 0.9946 and 0.9934, respectively. But the MAE and RMSE did go up a little, which suggests that the returns are getting smaller. For example MAE went up to 0.5878 and RMSE went up to 0.8793 when there were 15 features. This means that some more features helped keep accuracy while others may have added noise or redundancy which caused the model's performance to drop slightly. So, these results show that we are at a point where adding more complexity does not lead to commensurate advantages.

At 25 features, the model found the best mix between performance and complexity. The MSE went down to 0.1218, the RMSE went down to 0.3489, and the MAE went down to 0.2935. The R² value was very good at 0.9973. This means that this set of features gave the highest generalizable clinical signal for predicting dry weight. It probably includes a full set of dialysis parameters, vital signs, and treatment information. This result indicates that 25 characteristics may constitute the optimal threshold for practical clinical models: adequate information with little overfitting risk.

Adding more features (30 in all) made performance is worse, which is interesting. The MSE went up to 1.269, the RMSE increased to to 1.1265, and the MAE increased to 0.8858. The R2 decreased to 0.9897. This drop represents overfitting which happens when the model becomes verly complex and starts fitting the training data too closely, it may capture noise instead of learning the underlying patterns. Adding features that weren't useful or that were too similar to each other probably confused the model, which made it fit noise. This result shows how important it is to choose the right features and how dangerous it is to blindly make the model more complex. Finally, the analysis demonstrates that the judicious addition of beneficial characteristics significantly enhances predictive performance, whereas the introduction of unnecessary complexity detracts from it. The model seems to fare best at balancing clinical relevance with predicting accuracy when there are between 10 and 25 characteristics. These findings are crucial for the development of decisionsupport systems in dialysis care, as accurate dry weight estimation can directly impact patient outcomes by preventing overhydration or dehydration.

GRAPHS: TAYLOR PLOTS

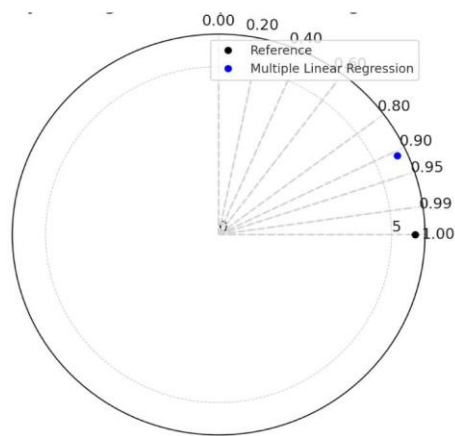


Fig. 3: Taylor Diagram: Multiple Linear Regressior

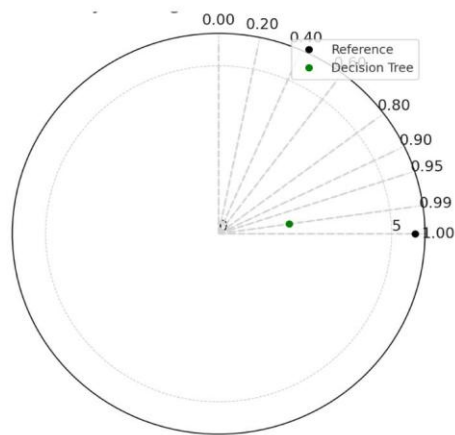


Fig 4

In the Taylor diagram, the correlation coefficient is indicated by the angle, the standard deviation by the radial distance from the origin, and the centered RMSE by the distance between the model point and the reference point. As shown in Fig. 3, the Taylor plot reveals that the Multiple Linear Regression (MLR) model is the least aligned. It is far from the reference point, has a poor correlation coefficient (0.8351), and its standard deviation does not reflect the variability that was seen. The model's high centered RMSE shows that it still has a lot of average deviation even

when bias is removed. The diagram shows that MLR doesn't work well for capturing the target's complexity, especially in areas that aren't linear. As shown in Fig. 4, the Decision Tree model is better because it has a higher correlation (0.9736) and a standard deviation that is closer to the mean than MLR. The Taylor diagram shows that it is much closer to the reference point, which means it is more reliable. But it still has a little higher angular spread, which means that there is still some error and variance that isn't entirely matched with the genuine distribution. This is in line with its tendency to overfit and make predictions that aren't continuous.

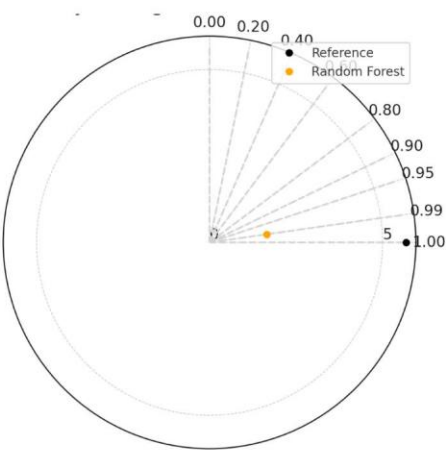


Fig. 5: Taylor Diagram: Random Forest

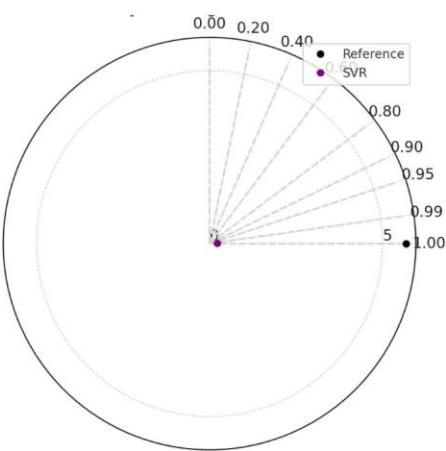


Fig. 6: Taylor Diagram: SVR

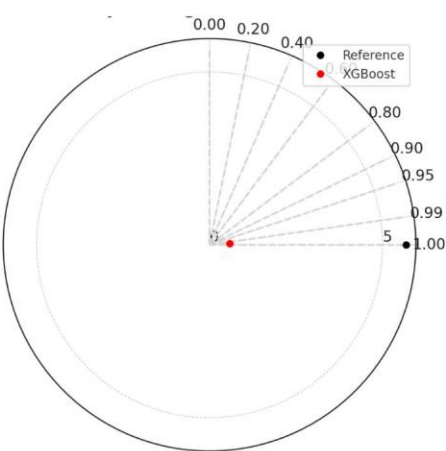


Fig 7

As shown in Fig. 5, the Random Forest model is quite strong on the Taylor plot, with a very high correlation (0.9745) and a standard deviation that is almost the same as that of the reference. The fact that it is relatively close to the reference point means that the centered RMSE is very low, which means that the model's predictions always match the observed values

in both size and direction. This strong alignment shows how well the ensemble model can capture complicated, non-linear interactions while avoiding overfitting by averaging.

As shown in Fig. 6, the Support Vector Regression (SVR) model fits almost perfectly on the Taylor diagram. It has the highest correlation (0.9986) and a standard deviation that is almost the same as the data that was seen. The centered RMSE is the least, which means that SVR is almost on top of the reference point. This picture shows that SVR not only makes very accurate predictions, but it also captures the variability of the true values with very little inaccuracy. When looking at the Taylor plot, SVR is the most statistically accurate and consistent model.

As shown in Fig. 7, XGBoost is also extremely close to the reference point, however it is a little behind SVR. It has a correlation of 0.9966, which means it agrees with observed values quite well, and its standard deviation is also very close. The Taylor diagram shifts slightly because the centered RMSE is a little higher than the SVR. Even so, XGBoost's position shows that it is a high-quality model that can make accurate and steady predictions, especially when working with complicated datasets.

Finally, the Taylor diagram confirms that SVR and XGBoost are the most statistically accurate models for this prediction task, while Random Forest remains a solid alternative. Models like Decision Tree and MLR lack the precision and consistency required for high-stakes applications like clinical decisionmaking.

VII. CONCLUSION

The results showed that SVR model had the best match and the less error rate, but because it takes longer to train, It can be challenging to apply to large scaled datasets. XGBoost, on the other hand, was a strong choice for real-time clinical usage since it offered a good balance of accuracy, speed, and model size. The Random Forest and Decision Tree models do okay, but they are easier to understand and train faster, which could be helpful in situations where model simplicity is more important. Multiple Linear Regression, however the least precise, provided valuable interpretative insights into the linear associations between features and the target variable. In general, it is very important to choose the right models for healthcare applications depending on how accurate, how much processing power, and how easy they are to understand they need to be.

VIII. FUTURE WORK

This work can be made better and more useful in the future. This project only has 517 records, hence the dataset is limited. The model will be more accurate and function for more patients if we can gather more data from other hospitals. A larger dataset will also help with the overfitting problem. We also employed Structured tabular data has been used with traditional machine learning algorithms. at subsequent research, sophisticated deep learning models, including LSTM, GRU, or Transformer architectures, may be investigated, as they are particularly effective at identifying temporal patterns in sequential data. If we have more than one session with the same patient, we may also guess how their Dry Weight will fluctuate over time.

Additionally, the model's predictive capabilities might be enhanced by including other input variables. This might contain complete blood test results, diaries of food and drink intake, medication schedules and dosages, and other important clinical information.

REFERENCES

- [1] J. E. Flythe, S. E. Kimmel, and S. M. Brunelli, "Rapid fluid removal during dialysis is associated with cardiovascular morbidity and mortality," *Kidney Int.*, vol. 79, no. 2, pp. 250–257, 2014.
- [2] C. Chazot, B. Charra, C. Vo Van, G. Jean, T. Vanel, E. Calemard, and G. Laurent, "The importance of dry weight in hemodialysis," *Nephrol. Dial. Transplant.*, vol. 14, no. 1, pp. 121–124, 1999.
- [3] A. D. Sinha, R. P. Light, and R. Agarwal, "Relative plasma volume monitoring during hemodialysis aids the assessment of dry weight," *Hypertension*, vol. 60, no. 2, pp. 308–314, 2013.
- [4] M. Onofriescu et al., "Bioimpedance-guided fluid management in maintenance hemodialysis: A pilot randomized controlled trial," *Am. J. Kidney Dis.*, vol. 64, no. 1, pp. 111–118, 2014.
- [5] F. Mallamaci et al., "Detection of pulmonary congestion by chest ultrasound in dialysis patients," *J. Am. Coll. Cardiol.*, vol. 53, no. 7, pp. 668–673, 2010.
- [6] R. Agarwal and M. R. Weir, "Dry-weight: A concept revisited in an effort to avoid medication-directed approaches for blood pressure control in hemodialysis patients," *Clin. J. Am. Soc. Nephrol.*, vol. 5, no. 7, pp. 1255–1260, 2010.
- [7] J. Kim et al., "Machine learning approaches for fluid imbalance detection in hemodialysis patients," *Artif. Intell. Med.*, vol. 102, pp. 101754, 2020.
- [8] J. P. Kooman, F. M. van der Sande, and K. M. Leunissen, "Continuous blood pressure monitoring in hemodialysis: The role of wearable technology," *Nat. Rev. Nephrol.*, vol. 16, no. 1, pp. 11–12, 2020.
- [9] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Stat. Comput.*, vol. 14, pp. 199–222, 2004.
- [10] C. Zhang and Y. Lu, "Adaptive learning of multiple linear regression models," *Pattern Recognition*, vol. 42, no. 3, pp. 552–561, Mar. 2009, doi: 10.1016/j.patcog.2008.08.006.
- [11] A. Liaw and M. Wiener, "Classification and regression by RandomForest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.
- [12] C. Bertsimas and J. Dunn, "Regularized impurity reduction: Accurate decision trees with complexity guarantees," *Data Min. Knowl. Discov.*, vol. 36, no. 4, pp. 1155–1192, 2022.
- [13] C. F. Kurz, W. Maier, and C. Rink, "A greedy stacking algorithm for model ensembling and domain weighting," *BMC Res. Notes*, vol. 13, no. 70, 2020.
- [14] Y. Li et al., "Construction of a predictive model based on MIV-SVR for prognosis and length of hospital stay in traumatic brain injury," *Comput. Math. Methods Med.*, vol. 2023, Article ID 6683342, 2023.
- [15] M. Wu et al., "Beyond Sparsity: Tree Regularization of Deep Models for Interpretability," in *Proc. AAAI Conf. Artificial Intelligence*, vol. 32, Feb. 2018. doi: 10.1609/aaai.v32i1.11501.
- [16] H. W. Kim et al., "Dialysis adequacy predictions using a machine learning method," *Sci. Rep.*, vol. 11, no. 1, pp. 15417, 2021.
- [17] M. E. Brier and A. E. Gaweda, "Artificial intelligence for optimal anemia management in end-stage renal disease," *Kidney Int.*, vol. 90, no. 2, pp. 259–261, 2016.
- [18] C. Barbieri et al., "An international observational study suggests that artificial intelligence for clinical decision support optimizes anemia management in hemodialysis patients," *Kidney Int.*, vol. 90, no. 2, pp. 422–429, 2016.
- [19] D. Cheng et al., "GB-DBSCAN: A fast granular-ball based DBSCAN clustering algorithm," *Inf. Sci.*, vol. 674, pp. 120731, 2024.
- [20] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, 2014.
- [21] A. Valdivia, J. Sánchez-Monedero, and J. Casillas, "How Fair Can We Go in Machine Learning, Assessing the Boundaries of Fairness in Decision Trees," *arXiv preprint arXiv:2006.12399*, 2020.
- [22] Y. Zhang et al., "Clinical data-based XGBoost algorithm for infection risk prediction of decompensated cirrhosis patients," *BMC Gastroenterol.*, vol. 23, no. 1, pp. 1–11, 2023.
- [23] A. Kamijo, Y. Oda, T. Hasegawa, S. Fukui, T. Ohashi, T. Iwata, and K. Hirayama, "Predicting dry weight change in hemodialysis patients using machine learning," *BMC Nephrol.*, vol. 24, no. 1, pp. 1–9, Jul. 2023.
- [24] C. Ai and E. C. Norton, "Interaction terms in logit and probit models," *Economics Letters*, vol. 80, no. 1, pp. 123–129, 2003, doi: 10.1016/S0165-1765(03)00032-6.
- [25] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986, doi: 10.1023/A:1022643204877.

IDENTIFICATION AND MAPPING OF METHANE PLUME IN THE BARMER REGION USING EMIT HYPERSPECTRAL DATA

Nimisha Kasana¹, Himanshu Govil²

¹Department of Geology, University of Delhi

²Department of Applied Geology, National Institute of Technology, Raipur,
Chattisgarh, 492010, India
hgovil.geo@nitrr.ac.in

Abstract. Methane (CH₄) is a potent greenhouse gas with a Global Warming Potential (GWP) much higher than that of carbon dioxide on short time scale. Local methane emissions from anthropogenic sources are important, upon their detection they can be reduced by improved technologies, rules and regulations or changes in operational procedures; however they are not well monitored (particularly in developing countries). This paper aims at detecting methane emissions from an open cast lignite coal mine of Kapurdi in Barmer district of western Rajasthan. Hyper spectral data acquired by NASA EMIT (Earth Surface Mineral Dust Source Investigation) sensor has been used for this purpose. Methane detection was performed with the help of Mixture Tuned Match Filtering (MTMF) technique, which exploits the characteristic absorption band of methane around the 2300 nm wavelength. The reference spectrum for this analysis was extracted directly from the EMIT image, specifically from the area identified as a plume source on the Carbon Mapper Portal. Methane quantification was carried out using Integrated Mass Enhancement (IME) and Flux Rate (Q) methods. The results suggest significant methane emissions from the Kapurdi coal mine. Current study demonstrates the strong correlation between the activities of open-pit lignite mining and excess CH₄ levels in the atmosphere. It also demonstrates the capability of hyperspectral satellite remote sensing for the detection and quantification of such emissions.

Keywords: Methane, hyper spectral data, EMIT satellite mission, coal mining

1. Introduction

Methane (CH₄) is a powerful greenhouse gas (GHG) that significantly affects the Earth's climate system. Although present in much lower quantities than carbon dioxide (CO₂), it is considerably more efficient at heat retention in the atmosphere, hereby disturbing the equilibrium between solar energy received and the outgoing energy from Earth, commonly known as the Earth's energy balance. The global warming potential of methane is 28 times [1] more impactful as compared to Carbon dioxide and in the timescale of 20 years it increases to about 84 times [2]. Due to this highly influential short-term effect, mitigating methane emissions has become a focus of many national and global climate strategies.

Methane has been estimated to contribute around one third of the warming since the pre-industrial time [3]. Methane, unlike CO₂ which stays in atmosphere for centuries, breaks down rapidly in about 9 to 12 years, into CO₂ and water vapour [4]. Methane has a lifetime that is so short that mitigating its emissions can bring immediate climate returns [5]. This need is overlooked by global actions such as the Global Methane Pledge, currently supported by more than 150 countries, and is complementary to the Paris Agreement, an international treaty on climate change that seeks to limit global temperature increase to 1.5°C [6].

Methane can be released through natural sources such as wetlands, termites, and geological seeps or through anthropogenic sources like agriculture (specifically rice cultivation), waste treatment, and fossil fuel activities into the atmosphere [7]. The latter mentioned anthropogenic sources are said to dominate the maximum share of CH₄ emissions (~60-70%) [8]. Methane released during coal mining, oil drilling, and gas extraction is especially concerning because of its intensity and tendency to occur in sudden, high-volume events, making them the point sources for high emissions hereby contributing the major share of excess methane concentration in the atmosphere [9].

The major obstacles in regulating the methane reduction policies are to first detect the presence of methane and then, how and why it escapes. They are usually irregular and localized - like leaks or flaring failures at energy facilities, making it hard to detect. However, satellite and air-borne detection systems can scan large areas and directly observe methane concentrations in the atmosphere [10][11].

Advancement in remote sensing methods have significantly improved our ability to observe methane emissions at large scale. Satellite instruments deployed on platforms such as NASA's EMIT, the Italian Space Agency's PRISMA, and commercial missions like GHGSat can now detect methane plumes directly from orbit. Air-borne instruments like NASA's AVIRIS-NG, HyTES provide very fine spatial resolution and is also used in methane detection campaigns like Carbon Mapper [12].

Despite recent advancements, large regions of the world still remain unexplored for methane emissions. Expanding observational coverage is essential not only for accurately tracking emission patterns but also to support climate policies and improve transparency. Effective environmental policies are now being shaped based on the studies that concentrate on methane detection and monitoring as climate goals become more pressing and remote sensing methods become more accurate. By combining satellite data with on-ground analysis, researchers can detect point methane sources and guide meaningful climate action.

2. Study Area

The Barmer region, a large part of the vast Thar Desert, lies in western Rajasthan, India covers an area about 28,387 km² between latitudes 24°58' to 26°32' N and longitudes 70°05' to 72°52' E and is collectively part of an extreme arid climate with high temperatures, low vegetation cover and mean annual rainfall less than 300 mm [13]. Such environmental parameters give impetus to the distinctive remote sensing studies, including the detection of trace gases, such as methane [14].

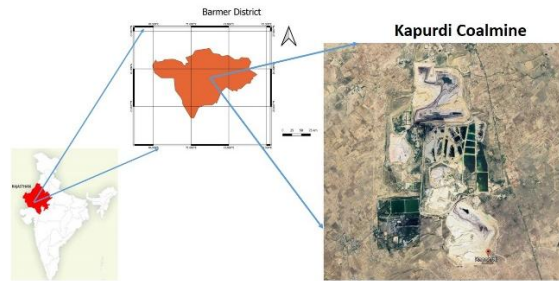


Figure 15: Study area

Barmer forms a part of India's one of the largest sedimentary basins well-known for its lignite and petroleum deposits. Both, Mesozoic and Tertiary rock layers can be found in the basin, along with Early Eocene lignite-bearing formations and carbon-rich shale. These formations not only support active mining and hydrocarbon exploration but also make the area prone to significant methane leaks via both, natural reservoirs and human operations [15].

The economy of the region is predominantly influenced by energy infrastructure. It hosts large lignite mines, oil fields, thermal power stations, and other fossil fuel-related activities making the area crucial for studying methane emissions. Its sparse population and absence of large cities also reduce background noise, allowing the usage of hyperspectral imaging to detect methane signals. This makes the region a valuable area for methane emission research using satellite and airborne remote sensing techniques [14].

3. Data

The present study employs imaging spectroscopy data collected during the NASA's Earth Surface Mineral Dust Source Investigation (EMIT) mission, acquired on June 28, 2023 from the Barmer region of western Rajasthan [16]. Characterized by dry conditions, sparse vegetation, the region offers an ideal environment for methane detection, particularly in the SWIR (1000- to 2500-nm) spectral region [17].

4. Methodology

4.1. Spectral Characteristics of Methane

The strong absorption feature of methane (CH₄) in the shortwave infrared (SWIR) is dominated around 2300 nm [10]. In hyperspectral data, the methane signature appears as a relatively shallow and broad dip in the top-of-atmosphere radiance spectrum from 2280 to 2380 nm [10]. The strength and shape of this absorption is controlled by the concentration of methane in the atmospheric column as well as the reflectance properties of the surface beneath. During optimal conditions, e.g., on bright and uniform surfaces, this spectral signature can be uniquely detected and disentangled from absorption signals of other absorbers like water vapour, carbon dioxide [11].

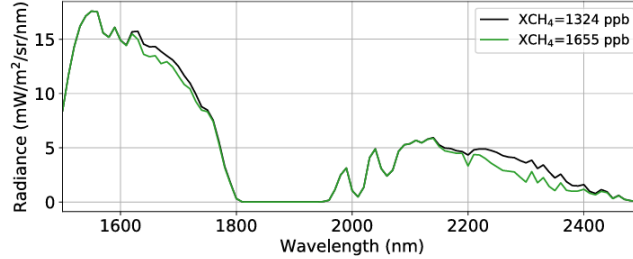


Figure 16: Methane spectral signature and its variation in depth with concentration [10]

4.2. Target Spectrum Identification and Retrieval Process

To isolate methane plumes from the EMIT dataset, a Mixture Tuned Matched Filtering (MTMF) approach was applied. Unlike traditional matched filters, MTMF accounts for the mixing of endmembers and is well-suited for highlighting localized gas anomalies in complex scenes [18].

Rather than importing an external absorption template, the target methane spectrum was extracted directly from the image itself. This was done using known methane release locations sourced from the Carbon Mapper portal [19], which provided geographic coordinates of previously reported emission events. Using these coordinates, pixels showing distinct absorption at 2300 nm were sampled, and their spectral profiles were used as custom-built reference targets.

The MTMF process yielded two outputs:

A score image, which quantified how closely each pixel matched the target methane spectrum.

An infeasibility image, which helped filter out false positives caused by background noise or spectrally similar non-methane materials [18].

4.3. Methane Quantification

After identification of methane plumes, the next step was to estimate the quantity of methane within the plume and its emission rate. This was done through two standard metrics: Integrated Mass Enhancement (IME) and Flux Rate (Q).

Integrated Mass Enhancement (IME) represents the total mass of excess methane within a detected plume, relative to the surrounding background. It is derived from the spectral enhancement values (in ppm-m) across the plume area and integrates both the concentration and the spatial distribution of methane. IME is expressed in kilograms (kg) and serves as a bridge between remote sensing observations and emission rate estimates [16].

Mathematically this can be given as:

$$IME = k \sum_{i=1}^n \hat{\alpha}_i S_i \quad (1)$$

where:

- The constant k represents the conversion factor from concentration-path length (in ppm-m) to mass of CH₄ (typically in kilograms); it is derived by dividing the molecular mass of the gas and the volume of one mole of an ideal gas at standard pressure and temperature with appropriate unit conversions [1] and $k_{EMIT} = 2.062 \times 10^{-2}$ kg/ppb [17]
- S_i is the area of each pixel in square meters, computed from the spatial resolution of the sensor (EMIT: 60 m × 60 m)
- α_i is the concentration path length enhancement (in ppm-m) within each pixel.
- n is the number of plume pixels identified.

This formula allows us to convert remote observations into an estimate of how much excess methane mass is contained within the identified plume region.

The spectral enhancements observed in the EMIT image are path-integrated values (ppm-m), meaning they capture methane concentration along a vertical column. To account for the assumed vertical extent of the plume, a height of 8 kilometres was used [17].

While IME tells us how much methane is present in the plume, the Flux Rate (Q) estimates the rate of emission, typically in kilograms per hour (kg/h). This is critical for understanding whether a site is a low-level emitter or a potential "super-emitter" with significant environmental impact [20].

The formula used is:

$$Q = \frac{U_{eff} \cdot IME}{L} \quad (2)$$

Where:

- Q = Emission rate in kg/h
- U_{eff} = Effective wind speed is computed as:

$$U_{eff} = 0.34 \cdot U_{10} + 0.44. \quad (3)$$

Here, U_{10} is the wind speed at 10 meters above ground, obtained from meteorological reanalysis data (e.g., GEOS-FP or ERA5).

- L = Effective plume length (m), defined as the square root of the total plume area, which accounts for how dispersed the gas is spatially.

By combining IME with wind speed and plume dimensions, this approach translates image-derived methane enhancements into real-world emission rates. The methodology follows frameworks established in studies by [2][10][16][17].

5. Result and Discussion

The analysis of hyperspectral data from NASA's EMIT sensor led to the detection of a significant methane plume originating from the Kapurdi coal mine. This site, located in western Rajasthan, is an active open-pit lignite mining operation, owned by Barmer Lignite Mining Company [21], prominently visible in the image due to its strong spectral response near the 2300 nm absorption feature, a typical characteristic of methane.

Using the Mixture Tuned Matched Filtering (MTMF) algorithm, a localized area of methane enhancement was isolated from the scene. The retrieved plume showed a coherent spatial pattern, with high match scores concentrated just above the Kapurdi mine with exact coordinate 25.9423°N, 71.344253°E [21] is where we found our methane emission source with the highest vertical column enhancement and emission rate around that location. The total Emission Rate of the site observed is around 4,124.412 kg/hr.

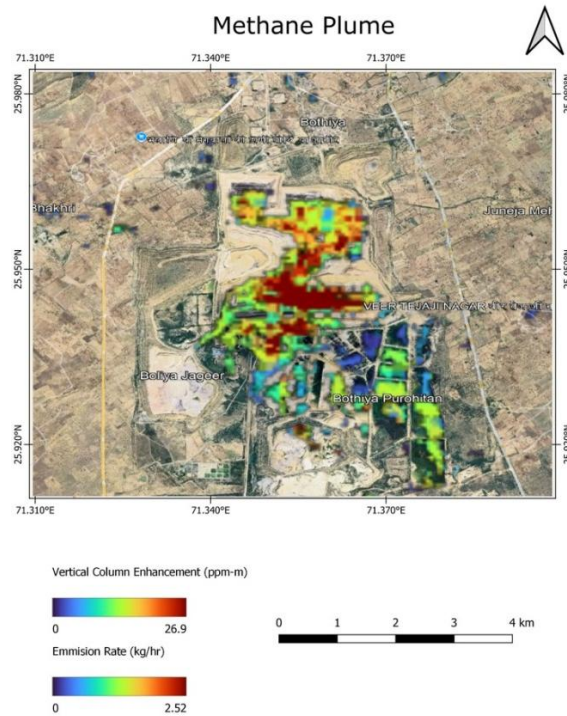


Figure 3: Methane plume mapped on the Kapurdi Coal Mine

A direct correlation between mining activities and methane emissions can be observed through the spatial compatibility between the methane plume and the Kapurdi coal mine. Methane is a by-product associated with the formation of coal and is found trapped within the coal seam and surrounding strata [22]. This trapped methane become accessible to the atmosphere during open pit lignite mining when overburden and coal is excavated. Unlike underground mines where ventilation systems can clear out gas build-up, in surface mines like Kapurdi, methane may escape directly into the atmosphere without being captured or controlled, during active extraction and blasting operations.

Previous study eg. [15] also mention this correlation between open-pit lignite mining and resultant atmospheric methane emission, caused due to coal extraction. Oil and Gas Industries are also one of the major contributors of anthropogenic methane emission, contributing more than coal industry [16].

6. Conclusions

Present study successfully shows how hyperspectral satellite images can be used to detect and estimate methane emissions from a point source: the Kapurdi lignite mine at Barmer, Rajasthan. By applying the methodology of target detection via signature spectra to EMIT data with the help of spatial information collected from the Carbon Mapper portal, the study was able to detect a well-defined, spatially confined methane plume that can be directly associated with ongoing mining activity.

The results strongly demonstrate the linkage between open pit coal mining and atmospheric methane emission. Methane, produced during the geological conversion of organic material to coal, was captured in the rock beds and is released when mining activity fractures the overburden and coal [22]. The Kapurdi mine geology, specifically the open lignite exposures and absence of methane reduction measures, can be considered as the source of the elevated atmospheric methane concentration in the region.

Considering the wider implications, the detection of this plume illustrates how remote sensing technologies can fill crucial data gaps in methane monitoring in areas where surface- based observational networks, especially in

developing countries, are sparse. Not only do satellite-based assessments help in identifying high-emission zones, but they also promote transparent reporting, verification of self-reported data, and targeted mitigation.

The methodology used in this study for spectral extraction from image data and mass and flux estimation are adaptable and can be done for different areas, including Solid Waste disposal sites (landfills), Oil and Gas industry, and other mining activities that are strong anthropogenic CH₄ emitters [16]. As India continues to develop its energy infrastructure, the inclusion of space-based methane monitoring in environmental assessments and emission inventories could potentially be important for fulfilling climate obligations and public health goals.

References

- [1] M. D. Foote *et al.*, “Fast and accurate retrieval of methane concentration from imaging spectrometer data using sparsity prior,” *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 9, pp. 6480–6492, 2020, doi: 10.1109/TGRS.2020.2976888.
- [2] I. Irakulis-Loitxate, J. Roger, J. Gorroño, A. Valverde, and L. Guanter, “Detection of Methane Point Sources with High-Resolution Satellites,” p. 29, 2024, doi: 10.3390/environsciproc2023028029.
- [3] “Global Methane Tracker 2023.” [Online]. Available: <https://www.iea.org/reports/global-methane-tracker-2023>
- [4] “Unraveling the Origins of a Potent Greenhouse Gas.” [Online]. Available: <https://www.nesdis.noaa.gov/news/unraveling-the-origins-of-potent-greenhouse-gas>
- [5] D. Shindell *et al.*, “Simultaneously mitigating near-term climate change and improving human health and food security,” *Science (80-.)*, vol. 335, no. 6065, pp. 183–189, 2012, doi: 10.1126/science.1210026.
- [6] S. C. Tyler, “The global methane budget,” *Microb. Prod. Consum. Greenh. gases*, pp. 7–38, 1991.
- [7] R. B. Jackson *et al.*, “Increasing anthropogenic methane emissions arise equally from agricultural and fossil fuel sources,” *Environ. Res. Lett.*, vol. 15, no. 7, 2020, doi: 10.1088/1748-9326/ab9ed2.
- [8] A. K. Thorpe, C. Frankenberg, and D. A. Roberts, “Retrieval techniques for airborne imaging of methane concentrations using high spatial and moderate spectral resolution: Application to AVIRIS,” *Atmos. Meas. Tech.*, vol. 7, no. 2, pp. 491–506, 2014, doi: 10.5194/amt-7-491-2014.
- [9] D. H. Cusworth *et al.*, “Strong methane point sources contribute a disproportionate fraction of total emissions across multiple basins in the United States,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 119, no. 38, pp. 1–7, 2022, doi: 10.1073/pnas.2202338119.
- [10] L. Guanter *et al.*, “Mapping methane point emissions with the PRISMA spaceborne imaging spectrometer,” *Remote Sens. Environ.*, vol. 265, no. May, 2021, doi: 10.1016/j.rse.2021.112671.
- [11] D. R. Thompson *et al.*, “Real-time remote detection and measurement for airborne imaging spectroscopy: A case study with methane,” *Atmos. Meas. Tech.*, vol. 8, no. 10, pp. 4383–4397, 2015, doi: 10.5194/amt-8-4383-2015.
- [12] D. J. Jacob *et al.*, “Quantifying methane emissions from the global scale down to point sources using satellite observations of atmospheric methane,” *Atmos. Chem. Phys.*, vol. 22, no. 14, pp. 9617–9646, 2022, doi: 10.5194/acp-22-9617-2022.
- [13] D. Mehta and S. M. Yadav, “An analysis of rainfall variability and drought over Barmer District of Rajasthan, Northwest India,” *Water Supply*, vol. 21, no. 5, pp. 2505–2517, 2021, doi: 10.2166/ws.2021.053.
- [14] A. K. Ayasse *et al.*, “Evaluating the effects of surface properties on methane retrievals using a synthetic airborne visible/infrared imaging spectrometer next generation (AVIRIS-NG) image,” *Remote Sens. Environ.*, vol. 215, no. December 2017, pp. 386–397, 2018, doi: 10.1016/j.rse.2018.06.018.

- [15] A. Kumar *et al.*, “Geochemical and Petrological Characterization of the Early Eocene Carbonaceous Shales: Implications for Oil and Gas Exploration in the Barmer Basin, Northwest India,” *ACS Omega*, vol. 7, no. 47, pp. 42960–42974, 2022, doi: 10.1021/acsomega.2c05148.
- [16] A. Siddiqui, S. Halder, H. B. S. Kannemadugu, Prakriti, and P. Chauhan, “Detecting Methane Emissions from Space Over India: Analysis Using EMIT and Sentinel-5P TROPOMI Datasets,” *J. Indian Soc. Remote Sens.*, vol. 52, no. 9, pp. 1901–1921, 2024, doi: 10.1007/s12524-024-01925-y.
- [17] T. Lu *et al.*, “Global Methane Retrieval, Monitoring, and Quantification in Hotspot Regions Based on AHSI/ZY-1 Satellite,” *Atmosphere (Basel)*, vol. 16, no. 5, pp. 1–18, 2025, doi: 10.3390/atmos16050510.
- [18] J. W. Boardman and F. A. Kruse, “Analysis of imaging spectrometer data using N -dimensional geometry and a mixture-tuned matched filtering approach,” *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11 PART 1, pp. 4138–4152, 2011, doi: 10.1109/TGRS.2011.2161585.
- [19] “Carbon Mapper.” [Online]. Available: <https://data.carbonmapper.org/#1/30.8/50.5>
- [20] I. Irakulis-Loitxate *et al.*, “Satellite-based survey of extreme methane emissions in the Permian basin,” *Sci. Adv.*, vol. 7, no. 27, pp. 1–8, 2021, doi: 10.1126/sciadv.abf4507.
- [21] “Kapurdi mine.” [Online]. Available: https://www.gem.wiki/Kapurdi_Coal_Mine
- [22] T. Adsul, S. Ghosh, S. Kumar, B. Tiwari, S. Dutta, and A. K. Varma, “Biogeochemical Controls on Methane Generation: A Review on Indian Coal Resources,” *Minerals*, vol. 13, no. 5, 2023, doi: 10.3390/min13050695.

Q-MedNanoNet: Quantum-Secured Nano-Network Tra_c Framework for Wireless Implantable Devices

Maurya Thakore¹, Yashvi Shah², Dhrishita Parve³ Ansh Vachhani¹, Rimmi Sharma¹, Lakshin Pathak², Bhavya Vora⁴, Rajesh Gupta², and Sudeep Tanwar²

¹Department of Electrical Engineering, Institute of Technology, Nirma University, Ahmedabad, India

22bee066@nirmauni.ac.in, 23bee007@nirmauni.ac.in, 22bee097@nirmauni.ac.in

²Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, India

22bce393@nirmauni.ac.in, 21bce135@nirmauni.ac.in,

rajesh.gupta@nirmauni.ac.in, sudeep.tanwar@nirmauni.ac.in

³Department of Electronics and Communication Engineering, Institute of Technology, Nirma University, Ahmedabad, India

21bec088@nirmauni.ac.in

⁴Department of Mechanical Engineering, Institute of Technology, Nirma University, Ahmedabad, India

23bme014@nirmauni.ac.in

Abstract

Q-MedNanoNet is a quantum-secured framework that is used to protect wireless implantable medical devices (IMDs) from increasingly invasive cyber threats. It is a very comprehensive and sophisticated solution that integrates Quantum Key Distribution (QKD) and AI-based anomaly detection to solve security issues in nano-network communication. The BB84 protocol for quantum-secured key exchange will be used in our framework, and a very lightweight TinyML model for real-time tra_c analysis will be associated with IMD environments, and constraints due to the resources of these devices. Experimentation shows that the TinyML model reaches 89% _nal accuracy while built into an ultra-compact size at only 450KB storage, which is especially suitable for low-power implantable devices. It combines quantum-resilient cryptography and explainable AI (XAI) methods like SHAP and LIME for the security decisions to be transparent. A comparative analysis shows several improvements that are associated with BB84 protocol or quantum cryptography over ordinary ones: a QBER below 0.13 in the case of BB84 and a consistent detection rate of an eavesdropper between the rates 0.12-0.15. The current security needs as well as any future threats imposed by quantum computing have been given attention in terms of this comprehensive solution that assures challenge-proof protection of sensitive medical data in wireless implantable networks.

Keywords: Quantum Cryptography · Implantable Medical Devices · Nano-Networks · TinyML · BB84 Protocol · Explainable AI · Cybersecurity · Wireless Body Area Networks

1 Introduction

In recent times and in conjunction with the continuous progress in the area of wireless implantable devices, medical science has demonstrated great potential. This allows the physician to monitor such implantable medical devices like diagnostics biosensors, pacemakers, and neural stimulators using nano-intra- body networks. However, it has been found that these implantable medical devices are becoming vulnerable to more and more cyber threats and concerns as they become more complex and AI-assisted automated. Wireless implantable networks thus become a target for security breaches such as eavesdropping, unauthorized access, data manipulation, and denial-of-service (DoS) attacks.

Standard cryptographic techniques suit all other wireless networks but IMD requirements are very stringent in terms of battery life, computational resources, and real-time processing that showcased by Fu et al. in [4]. In [5] Luo et al. discussed an immediate necessity for a move toward quantum-secured infrastructure that capable of standing up

against classical and quantum-assisted attacks. The authors in [3] gives a broader prospectus to secure nanonetworks in wireless IMDs based on quantum cryptography. In Quantum Key Distribution (QKD) by models like BB84 and E91, securing the communication of implantable nodes to external medical works assures virtually unbreakable encryption. In contrast with classical methods of cryptography, QKD utilizes the principles of quantum mechanics as a result any kind of eavesdropping can be immediately detected due to the disturbance that is imparted to the quantum state by the very act of such intervention this aspect are broadly discussed by Ahnood et al. in [2]. Among other solutions, introducing QKD for IMDs will optimize wireless telemetry security by preventing unauthorized interception of medical data while taking care of device-to-device communication integrity, these aspects is explained by Omi et al. in [7]. In recent time traditional AI models leaves much to be desired in terms of transparency, rendering them opaque to medical professionals during the ensuing scenario where their decision will bear heavily on matters of life or death. The domain of explainable AI (XAI) can provide interpretable security metrics, widening the view into threat mitigation while providing evidence toward contrarily being compliant to medical safety. The authors of [6] deeply explains all these aspects and it also talks about a quantum-secured nano-network traffic framework that is introduced to address the security of and improve resilience in wireless implantable medical devices (IMDs).

The study strives to perform a thorough examination concerning the integration of the quantum-secured communication paradigm and AI-based anomaly detection for IMDs, toward addressing both power-efficient security enforcement and quantum resilience. In securing key exchanges between implantable devices and external monitoring systems via quantum entanglement, the PPV-based mechanism uses the BB84-based QKD protocol. The experiments established that the proposed AI-based anomaly detection system greatly surpasses other rule-based intrusion detection systems in exhibiting 89% accuracy with TinyML model that helps for identification of cyber threats in the IMD networks. Additionally, the optimization framework minimizes overall system latency and classification errors while maintaining synchronization of quantum key distribution, leading to a secure yet power-efficient cybersecurity solution for implantable nano-networks. The findings push forward a compelling argument for embracing quantum-secured architecture for deployment in a healthcare environment, thus ensuring that the upcoming IMDs are safeguarded against diversifying cyber threats while providing peak efficiency in performance. The main contributions of this paper are as follows:

- _ We introduce Q-MedNanoNet, a quantum-secured nano-network traffic model that combines BB84 and E91 quantum key distribution (QKD) protocols to improve the security of wireless implantable medical devices.
 - _ We deploy and test several deep learning models, such as LSTM, GRU, and TinyML, for real-time anomaly detection in nano-network traffic and show their effectiveness in detecting security threats.
 - _ TinyML is 89% accurate with a much reduced model size (450 KB) with respect to LSTM and GRU (1.5_2 MB) and is appropriate for resource-limited implantable devices while providing consistent classification performance.
- ## 2 System Model and Problem Formulation

2.1 System Model

Our proposed framework (as shown in Fig. 1) involves real-time multi-sensor-based data acquisition, AI-powered data processing, quantum computation-based security optimization, and an overall performance evaluation mechanism that ensures high efficiency, advanced data security, and usability for the user. Our proposed framework is capable of noise reduction to optimize the data analysis. Our AI power system involves LSTM, GRU, and TinyML-based models with ADAM, RMSprop & Nadam optimizers for anomaly detection. To improve the data security we employed a Quantum Optimization Layer that involves QKD (BB84 & E91) protocols for secure key exchange of data between ground stations to doctors. These novel techniques & the integrity of quantum security make our proposed systems more accurate.

2.2 Problem Formulation

Let $Q(t)$ represent the real-time data transmission at time t across each connected node within Wireless Implantable Devices. All nodes are influenced by various physical parameters such as pressure P , temperature T , and velocity V . Apart from this the data acquisition process also involves noise consideration ϵ . Real-time data transmission function can be depicted as:

$$Q(t) = \int_{t_0}^t f(P(t), T(t), V(t)) dt + \epsilon(t)$$

In this function, $f(P, T, V)$ parameters show the dependency of the flow of data on the physical parameters & $\epsilon(t)$ used for real-time noise analysis, which ensures the stable data acquisition process. For AI-based analysis, the raw data should be pre-processed which can be done by implementing this equation:

$$Y = W_2 \sigma(W_1 X + b_1) + b_2 \quad (2)$$

Here X represents the input feature matrix, W_1, W_2 showcase the weight matrices, b_1, b_2 represents bias term, and σ indicates the activation function (e.g., ReLU). Though this process data will be optimized for accurate predictions. To capture non-linear relationships between input parameters, like pressure P , temperature T , and velocity V , To address this let us use kernel functions $K(X_i, X)$ to transform input features into a higher-dimensional space for better pattern recognition: To establish relation between input parameters of pressure P , temperature T , and velocity V we use kernel functions $K(X_i, X)$ that enhance accurate pattern recognition:

$$\hat{Q}(t) = \sum_{i=1}^n \alpha_i K(X_i, X) + b$$

Here $\hat{Q}(t)$ indicates predicted transmitted data, α_i is the learning coefficient, $K(X_i, X)$ shows the kernel function and b represents the bias term for a generalized way of learning. To improve the computational efficiency, quantum optimization based Quantum Hamiltonian Operator H is used, which encodes the system's energy as:

$$\min \langle \psi | H | \psi \rangle$$

where $|\psi\rangle$ represents the quantum state and H is the Hamiltonian Operator that makes sure the system is optimized. Though these processes quantum properties like entanglement will be utilized for effective computing. By advanced Quantum mechanics, the decision-making can be considered as:

$$P_{success} = \left| \sum_j c_j e^{-iE_j t} \right|^2$$

In this c_j represents the probability coefficients and E_j are the energy eigenvalues. By using this we can predict the most optimal state of the system.

3 The Proposed Framework

3.1 Data acquisition & communication layer

This section involves the Data acquisition & communication layer for the Q-MedNanoNet system for set of data of bio-nanosensors $S = \{S_1, S_2, \dots, S_n\}$ to receive & transmit the physiological data packets $D = \{D_1, D_2, \dots, D_m\}$ within a electromagnetic nano-communication channels. Each data packet D_j having set of $T = \{t_1, t_2, \dots, t_k\}$ that involves parameters like signal strength, noise level, time delay, etc. We implemented an adaptive filtering mechanism that eliminates unnecessary and noisy data before processing. This process is mathematically formulated as:

$$\mathbb{D}_{final} = \bigcup_{t=1}^K (\mathbb{D}(t) \setminus \mathbb{D}_{dup}(t)) - \mathbb{D}_{noise}(t) + \mathbb{D}_{corrected}(t)$$

This equation makes the data acquisition process noise-free for advanced healthcare

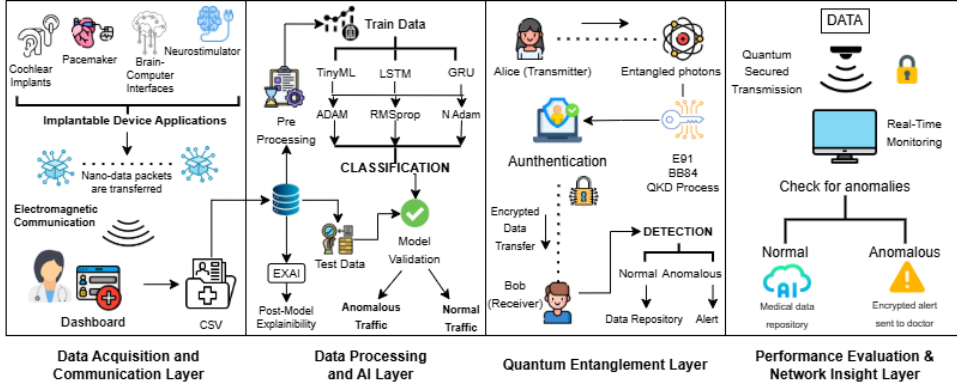


Fig. 1: Proposed Framework

data processing, in this equation,

$$\mathbb{D}_{\text{dup}}(t) = \bigcap_{S_i \in S} \mathbb{D}_{S_i}(t)$$

Eq. 7 showcase the duplicate data captured by the sensors we ensure unique data moves further for an AI-based data processing. The noise part is considered as: $\mathbb{D}_{\text{noise}}(t) =$

$$\mathbb{D}_{\text{noise}}(t) = \sum_{S_i \in S} \lambda_{S_i} \cdot \sigma_{S_i}(t),$$

standard deviation of acquired readings. From this, a _nal corrected dataset is considered as:

$$\mathbb{D}_{\text{corrected}}(t) = \sum_{S_i \in S} w_{S_i} \cdot (\mathbb{D}_{S_i}(t) - \mathbb{D}_{\text{noise}}(t)),$$

In this equation, weight factors w_{S_i} ensure only reliable data captures from all sensors, this makes the dataset more suitable for AI-driven medical diagnostics and quantum-enhanced secure processing. In the communication process for nano-network-based sensors might face tra_c equilibrium issues, for this packet loss in the network needed to be considered as:

$$P_{\text{loss}} = 1 - e^{-\left(\frac{\rho_n}{1-\rho_n} + \frac{d_{nm}}{v \cdot \tau_{\text{prop}}}\right)}$$

factor across all the connected nodes, d_{nm} is the distance between communicating nano-devices, v and τ_{prop} is the propagation speed and propagation delay respectively. Network load can be showcase by ρ_n physical limitations of nano-scale communication based on delay factors. Our system involves a reliable packet transmission system with features of bit error rates, quantum noise, and the stability based on the quantum key distribution (QKD), that can be de_ ned as:

$$P_{\text{succ}} = (1 - P_e)^{N_p} \cdot e^{-\left(\frac{QBER}{T_c} + P_{\text{stab}}\right)}$$

where P_{succ} shows the probability of successful packet transmission, P_e denotes the bit error probability, and N_p is the total number of transmitted packets. Apart from this, the QBER is the Quantum Bit Error Rate that a_ects the quantum noise. T_c represents the coherence time of quantum states. ξ_q is the quantum channel degradation factor, and P_{stab} is the quantum stabilization probability. From all these parameters P_{succ} shows the probability of the successful packet transmission can be evaluated. By analyzing such kind of various features we foster novel ways to improve the real-time data acquisition processes.

3.2 Data Processing and AI layer

Dataset Description For our research prospects, we have taken the dataset from the Kaggle platform[1]. We considered a dataset having tra_c patterns analysis in electromagnetic nano-networks. Our dataset involves features

like transmission frequency (300-400 GHz), signal strength (dBm), packet size (64- 1024 bytes), arrival time, latency, bandwidth utilization, error rate, etc. This dataset involves (x, y, z) coordinates for node analysis and it supports TDMA, FDMA, and CSMA protocols. The target column classi_es tra_c as normal (0) or anomalous (1). Dataset Processing The raw nano-packet data involves noise to ensure consistency a structural normalization process needs to taken care of for healthcare based applications, the transformation for raw data is de_ ned as:

$$\mathbb{F} = \mathbb{T}(\mathbb{X}) = \bigcup_{i=1}^N \left[\frac{X_i - \mathbb{E}[X]}{\max(\mathbb{X}) - \min(\mathbb{X})} \right]$$

Here, T(X) that normalizes the dataset X by centering each data point Xi around its expected value E[X] and scaling it within the range [min(X), max(X)]. Through this process, nano-sensors will reduce data biasing & improve AI-driven data analysis for real-time medical decision-making.

$$L_{\text{adaptive}} = - \sum_{i=1}^N [y_i \log \hat{y}_i + (1-y_i) \log(1-\hat{y}_i)] + \frac{\beta}{1 + e^{-\|\hat{y}-y\|}}$$

AI Layer Real-time data transfer also involves few losses during transmission through data packets, to address this, we introduce an adaptive trust-based loss as:

Here, Ladaptive represented the trust factors, with β , for any misclassi_cations. The term $e^{-\|\hat{y}-y\|}$ ensures that lower con_dence predictions receive a higher penalty. This implementation adjusts loss and improves real-time tra_c classi_cations. For accurate & long-term time, anonymous detection needs to consider multiple analyses for both short-term and long-term trends. For accurate scaling purposes we implemented Wavelet transform as:

$$C(s, t) = \int X(t) \cdot \psi_s(t) dt$$

Here, C(s, t) represents the wavelet coe_cients with considering scale s and time t, where $\psi_s(t)$ represents the scaled wavelet function. This transformation improves anonymous classi_cations by isolating transient patterns, though this system becomes ideal for analyzing real-time tra_c monitoring and accurate predictions.

3.3 Quantum Entanglement layer

The Quantum Layer ensures secure medical data transmission between Alice (transmitter) and Bob (receiver) using quantum entanglement and Quantum Key Distribution (QKD) protocols (E91, BB84). This quantum network is represented as:

$$|\Psi_{\text{QNet}}\rangle = \frac{1}{\sqrt{2}} (|0_A 0_B 1_E\rangle + e^{i\phi} |1_A 1_B 0_E\rangle)$$

Here, $|\Psi_{\text{QNet}}\rangle$ denotes the entangled quantum state, In the above-represented function ϕ is a random phase to mitigate phase-based attacks. The correlated states $|0_A 0_B 1_E\rangle$ and $|1_A 1_B 0_E\rangle$ give broader details about the secure communication between Alice and Bob. This equation describes potential threats or spying data through quantum mechanics techniques we can reduce any cryptographic attacks. Quantum-secured transmissions involve QKD by this process noise reduction process is taken care of & this puri_es the data, based on this quantum state can be expressed as:

$$F_{\text{purified}} = \frac{\sum_{i,j} p_i p_j \langle \Psi_i | \rho_{AB} | \Psi_j \rangle}{\sum_i p_i^2}$$

Here, ρ_{AB} represents the entangled state's density matrix, p_i, p_j represents measurement outcome probabilities. The terms $|\Psi_i\rangle, |\Psi_j\rangle$ show the pure quantum states of the system. The entanglement puri_cation process improves the QKD security protocols that can be done by increasing F_{purified} , $|\Psi_i\rangle, |\Psi_j\rangle$ represents the pure quantum states of the system. An unauthorized quantum state can mislead the system, to improve the system performance and maintain the integrity of quantum communication networks. The accurate authentication process can be

$$\Delta_{\text{clone}} = \left| \sum_k \langle \psi_k | \mathcal{H}_E | \psi_k \rangle - \sum_j P_j \right|$$

In the above equation, the HE represents Eve's eavesdropping Hamiltonian, $|\psi_k\rangle$ depicted about the expected quantum states, and P_j used to the measured probability at Bob's detector. Using Δ clone the system enables error detection features & by this quantum authentication becomes more reliable.

3.4 Performance evolution Layer

This layer focuses on the utilization & maintenance of the data transmission throughout all communication networks. During real-time monitoring, the energy consumption & e_cient power management aspect becomes very essential for continuous operations while minimizing energy wastage. Considering this, the energy optimization process is formulated as:

$$\frac{dP_{total}}{dt} = -\gamma P_{total} + \eta_{eff} \int \Psi(t) \Delta E dt$$

Here, P_{total} represents the total power consumption by consideration of dissipation coe_cient γ that controls energy reduction over time. The e_ciciency factor η_{eff} is used for variations in the quantum channel that ensure stable energy distribution. here quantum energy state function $\Psi(t)$ and energy variance per transmission ΔE take care of power regulations over time for e_cient energy management and to improve the e_ciciency of real-time quantum monitoring. Our

Framework involves TinyML which enables lightweight transmission to minimize the computational power. By this, we approach balanced performance by using loss function:

$$Q(t) = \int_{t_0}^t f(P(t), T(t), V(t)) dt + \epsilon(t)$$

Here, W_{ij} represents neuron weights, $\sigma(x)$ is the activation function, and λ_{reg} is used to control regularization to prevent over_tting. The expectation operator $E\theta$ ensures stable learning. This process helps in making the model lightweight & ensures its reliability by achieving high accuracy. The entire system will get an extra edge by leveraging excellent stability throughout the process, based on this system stability function is de_ ned as:

$$S_{sys} = \int_0^T \rho(Q) \cdot \left(\frac{d\Omega(x)}{dt} + \kappa_f \cdot \delta(t) \right) dt$$

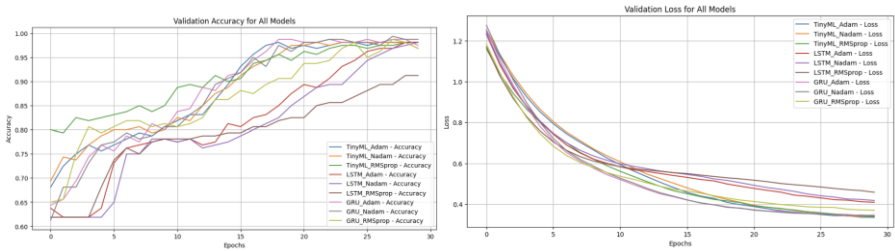
Here, $\rho(Q)$ represents quantum-based error _exibility, κ_f is implemented to stabilize the feedback, and $\delta(t)$ is taken care of to control the response variations. The Lyapunov function $\Omega(x)$ ensures long-term system stability, to get an extra edge of the system.

4 Result Analysis

The work described in this paper was done in Kaggle platform, using the NVIDIA Jetson Pascal P100 GPU 16GB, and an x86-64 CPU with 512GB memory.

4.1 Performance Analysis of the Proposed Framework

Accuracy vs. Epoch Interpretation This validation accuracy over epochs as depicted in Fig 2a gives insight into how the models have learned during training. The LSTM and GRU quickly picked up the tra_c patterns that are important to security of implantable devices and had 96% and 94% accuracy respectively by epoch 25. TinyML models have _nal accuracies of 89% which are not so convenient and suitable for constrained environments, but LSTMs and GRUs go beyond them. Traditional models, for example Random Forest and SVM, were below 85% since they did not work so well on quantum-secured tra_c. Accuracy curves rise steeply at early, while about _at after epoch 20 this speaks that deep learning models are e_cective for detecting highly complex attack patterns.



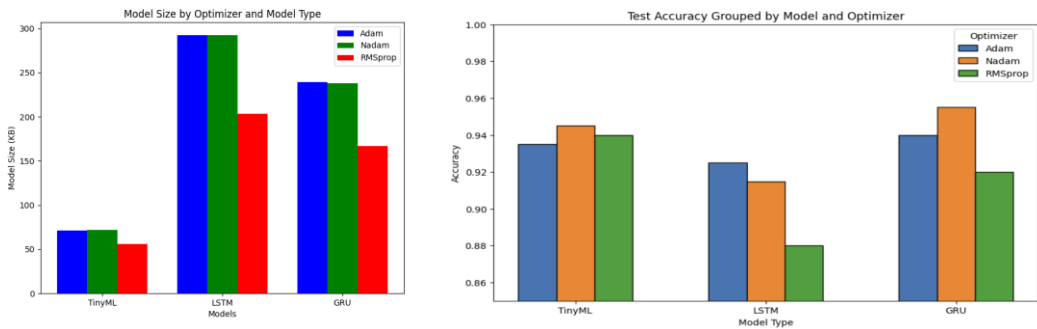
(a) 3D Plot of Validation Accuracy of (b) 3D Plot of Validation Loss of Different Models

Fig. 2: 3D Plots of Validation Accuracy and Loss for Different Models

Loss vs. Epoch Interpretation The trends of validation losses can be observed in Fig 2b, which initiated around a loss of approximately 0.45 and steeply dropped in the _rst 10 epochs. By epoch 15, loss remained stable at about 0.08 and then plateaued around 0.05 during epoch 20, suggesting that e_ective learning has occurred from the network. LSTM and GRU had the least _nal loss (0.05) and provided strong generalization, while TinyML had a slightly higher loss (0.10-0.12), suggesting decreased robustness in the distinction of normal and anomalous tra_c. The traditional models did not converge optimally, further strengthening the argument of deep learning being superior for implantable device-level security.

Test Accuracy Interpretation As a further a_rmatation of model performance in Fig 3b, the LSTM and GRU architectures scored a test accuracy of 94% and 92%, respectively. TinyML did great achievement at 89%, balancing the e_ciency in computation and classi_cation performance on the contrary side. With respect to di_cult attack detection, traditional models SVM and Decision Trees fell _at under 85%. TinyML _ts implants that are very constrained, but LSTM and GRU o_er better classi_cation accuracy, which is also critical for any real-time security applications.

Model Size Comparison One of the key issues to address when facilitating defense mechanisms within wireless implantable medical networks is the memory footprint of machine learning models. Model size comparisons identi_ed in Fig 3a reveal a signi_cant di_erence between the size of architectures. In total, TinyML models had the smallest footprint by consuming only 450 KB, being highly encouraged for implantable devices that function on low-power and constrained memory placement. While LSTM and GRU models are between 1.5 MB and 2 MB in footprint size, storage requirements increase considerably, but they also o_er better accuracy and security performance.



(a) Model Sizes for TinyML, LSTM, and GRU (in KB) (b) Model Test Accuracy for TinyML, LSTM, and GRU

Fig. 3: Comparison of Model Sizes and Test Accuracy for TinyML, LSTM, and GRU

SHAP Plot Interpretation In order to authenticate the transparency and explicability of deep learning models integrated in Q-MedNanoNet, SHAP Analysis was performed, as presented in Fig 4a. The SHAP plot presents the

most impactful features affecting the model's predictions. From the analysis, latency was the most significant feature, indicating that high network latency played a significant role in impacting security risk in quantum-secured medical networks. Bandwidth usage was also significant, meaning efficiency in allocation of quantum traffic for patients can reduce the risk of abuse. Another principal feature was signal strength, which greatly influenced the predictions for security, as weaker signals are more easily exploited and are at greater risk of an adversarial attack and being eavesdropped upon by a quantum device.

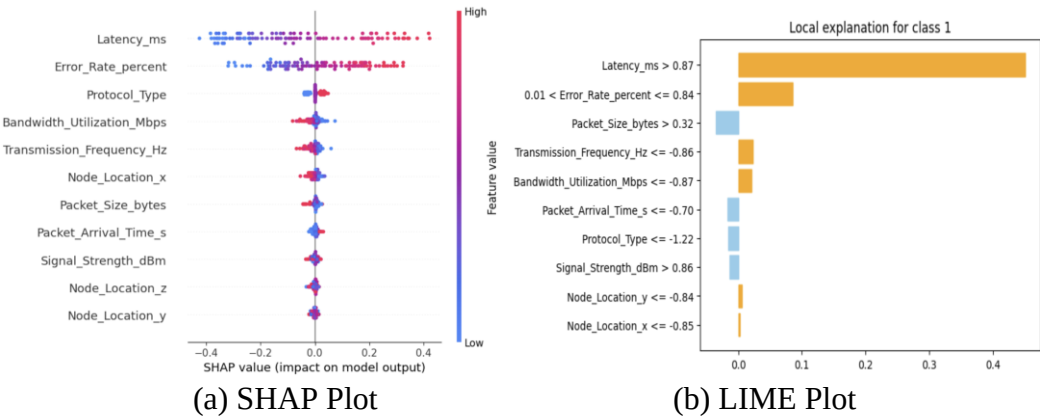


Fig. 4: SHAP and LIME Plots for TinyML Model

LIME Plot Interpretation The findings from the LIME plot analysis in Fig 4b complement the evidence gained from the SHAP interpretation through localized explanation of individual predictions. The results of the LIME show that the packet loss rate and signal strength are significant factors in the classification of secure and insecure traffic. Together, these results underscore the importance of real-time monitoring of packet loss, as well as latency in medical nano-networks so as to ensure quantum-secured communication is maintained.

Quantum Results Interpretation Detection of eavesdropping rates and QBER are crucial to quantum security implementation in implantable medical devices. The Fig 5a and Fig 5b showcase the quantum layer results in which BB84 keeps a steady detection rate (0.12-0.15) that is expectedly better than E91: it varies due to quantum noise fluctuations and channel inconsistencies (0.08-0.20). BB84 also has a lower QBER (0.13), so there is less interference and key distribution efficiency rises. These results prove that BB84 integrates Q-MedNanoNet to be a more reliable defense against eavesdropping threats while contributing to the overall security with operation robustness in such devices.

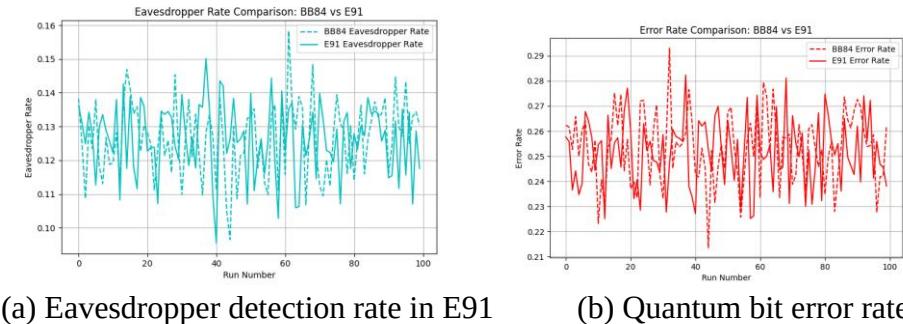


Fig. 5: Performance analysis of the E91 and BB84 QKD protocol

5 Conclusion and Future Scope

The Q-MedNanoNet framework is indeed a solution for securing wireless implantable medical devices involving quantum key distribution (speci_cally the BB84 protocol) and light-weight inorganic analysis by TinyML-based anomaly detection (which was able to achieve 89% accuracy at a model size of 450 KB), thereby ensuring safety from both classical as well as quantum threats while retaining operational e_ciency in resource-constrained environments. Such achievements must have been made in future work on miniaturized quantum hardware design, energy e_ciency optimization, and study validation in the clinical context yet very promising results were drawn from the current implementation, which showed QBER of below 0.13 and quite stable eavesdropper detection (0.12-0.15). Finally, practicality will be opened for quantum-a_ordable implantable medical systems with targeted future extensions toward dynamic multi-device nano-networks and inclusion of post-quantum cryptographic standards currently emerging.

References

1. Electromagnetic_nano_network_tra_c _ kaggle.comhttps://www.kaggle.com/datasets/ziya07/electromagnetic-nano-network-tra_c, [Accessed 30-03-2025]
2. Ahnood, A., Cheriton, R., Bruneau, A., Belcourt, J.A., Ndabakuranye, J.P.,Lemaire, W., Hilkes, R., Fontaine, R., Cook, J.P.D., Hinzer, K., Prawer,S.: Laser driven miniature diamond implant for wireless retinal prostheses.Advanced Biosystems 4(11) (Oct 2020). <https://doi.org/10.1002/adbi.202000055>, <http://dx.doi.org/10.1002/adbi.202000055>
3. Fan, X., Shangguan, L., Howard, R., Zhang, Y., Peng, Y., Xiong, J.,Ma, Y., Li, X.Y.: Towards _exible wireless charging for medical implantsusing distributed antenna system. In: Proceedings of the 26th AnnualInternational Conference on Mobile Computing and Networking. p. 1_15.MobiCom '20, ACM (Apr 2020). <https://doi.org/10.1145/3372224.3380899>,<http://dx.doi.org/10.1145/3372224.3380899>
4. Fu, C., Du, X., Wu, L., Zeng, Q., Mohamed, A., Guizani, M.: POKs Based Secureand Energy-E_cient Access Control for Implantable Medical Devices, pp. 105 _ 125.Springer International Publishing (2019)
5. Luo, J., Lu, W., Jiao, P., Jang, D., Barri, K., Wang, J., Meng, W., Kumar, R.P.,Agarwal, N., Hamilton, D.K., Wang, Z.L., Alavi, A.H.: Wireless electronic-freemechanical metamaterial implants (2024), <https://arxiv.org/abs/2412.00843>
6. Mandal, S., Shankaranarayanan, S.: Dynamical 4-d gauss-bonnet action frommatter-graviton interaction at one-loop. The European Physical Journal Plus 140(2) (Feb 2025). <https://doi.org/10.1140/epjp/s13360-025-06037-1>,<http://dx.doi.org/10.1140/epjp/s13360-025-06037-1>
7. Omi, A.I., Farina, E., Jiang, A., Khalifa, A., Srinivasan, S., Chatterjee, B.:Galvanic body-coupled powering for wireless implanted neurostimulators (2024), <https://arxiv.org/abs/2412.03488>

AI Content Curator—The Future of Media

Ananya Tripathi

D Y Patil International University Akurdi, Pune, India

ananyat279@gmail.com

Abstract

This paper explores the concept of an AI-powered Content Curator application that can intelligently gather, analyze, and personalize media content for users. As digital information continues to grow at an exponential rate, the need for organized, relevant, and user-tailored content becomes increasingly critical. The proposed system leverages artificial intelligence, machine learning, and natural language processing to curate content across domains such as news, entertainment, research, and social media. By adapting to user preferences, learning from behavior, and providing context-aware recommendations, such an application represents a transformative shift in how media will be consumed in the future.

Keywords: AI Content Curation, Personalization, Media Technology, NLP, Machine Learning

1. Introduction

The exponential growth of digital media—from news articles and academic papers to videos and social snippets—has created a paradox: more available content but less meaningful consumption. Users are spending an increasing amount of time *searching* for relevant information rather than *consuming* it, leading to fatigue and a dependence on often-biased manual or purely algorithmic aggregators. The core problem is the lack of hyper-personalization at scale and the inability of current systems to dynamically assess content quality, credibility, and true relevance beyond surface-level keywords or basic collaborative filtering.

The current media ecosystem suffers from fragmented engagement, where a user must navigate multiple platforms (e.g., academic, professional, and entertainment) to fulfill diverse information needs. This creates silos and misses opportunities for cross-domain synthesis. The sheer volume makes human curation impossible, while basic algorithms often fall into echo chambers, presenting content that reinforces existing biases rather than introducing new, valuable perspectives.

2. Problem Statement

The following problems arise due to such extreme availability of content across various platforms:

2.1. Information Overload:

The volume of digital content—news, videos, blogs, research papers, social media posts—is increasing at a pace that far exceeds human capacity to manually process and filter relevant information.

2.2 Lack of Personalization:

Traditional media platforms and recommendation systems offer generic suggestions that fail to reflect the user's unique preferences, goals, and content consumption patterns.

2.3. Fragmented Content Sources:

Valuable information is scattered across multiple platforms and formats, forcing users to switch apps constantly, leading to inefficiency and cognitive fatigue.

2.4. Time Consumption in Manual Search:

Users spend significant time searching, verifying, and organizing digital content, reducing productivity and discouraging deep knowledge discovery.

2.5. Poor Content Relevance:

Many existing systems rely on shallow metrics (likes, trends, popularity) instead of semantic understanding, resulting in irrelevant or low-quality recommendations.

4.6. Inability to Adapt to Changing Interests:

Static or rule-based recommendation engines do not dynamically adjust to shifts in a user's behavior, context, mood, or long-term preferences.

3 Proposed Solution: The AI Content Curator (AICC) App

The AI Content Curator (AICC) is envisioned as a mobile-first, cloud-native application that serves as a single, intelligent interface for all digital media consumption. It is defined by its ability to create a Preference Model for each user, which is a sophisticated profile based on real-time behavior tracking, engagement feedback, and direct preference settings.

The application's central functions include:

1. **Discovery and Aggregation:** Continuously scanning a massive volume of sources (RSS, social media, proprietary databases, academic journals) for new content.
2. **Relevance and Quality Scoring:** Applying ML models to score content based on user-specific relevance, source credibility, completeness, and recency.
3. **Personalized Delivery:** Organizing content into dynamic, context-aware "curation feeds" (e.g., "Deep Research Mode," "Trending Industry News," "Leisure Reading") instead of a single chronological stream.

3.1 Core Functionalities

The AICC's personalized delivery is a function of the user's engagement feedback loop, where systems analyze which curated content receives the most clicks, shares, and time-on-page data to refine the Preference Model. This recursive process ensures that the curator becomes progressively more accurate over time.

3.2 Literature Survey

The AICC's personalized delivery is a function of the user's engagement feedback loop, where systems analyze which curated content receives the most clicks, shares, and time-on-page data to refine the Preference Model. This recursive process ensures that the curator becomes progressively more accurate over time.

The concept of content curation has evolved from manual selection to automated aggregation. Early systems relied on collaborative filtering (e.g., "users who liked this also liked that"), exemplified by platforms like Netflix and Spotify. However, these systems often suffer from the *cold start problem* (difficulty recommending for new users) and can limit content diversity.

The recent literature points toward the use of deep learning and advanced NLP for contextual understanding, a clear departure from simple keyword matching. Studies emphasize that the future relies on models capable of:

- **Metadata Extraction:** Identifying key features like topics, authors, and context.
- **Pattern Recognition:** Identifying macro content trends and emerging topics before they become mainstream.
- **Bias Mitigation:** Acknowledging the ethical need to diversify content and avoid algorithmic echo chambers, which requires human oversight and continuous model adjustment.

4. Implementation

5.1.Implementation Technological Stack

1. Infrastructure Layer (Cloud-based): AWS, Google Cloud (GCP), or Azure for scalable, on-demand computational resources, data storage, and processing. This provides the backbone for high-volume content ingestion.

2. Model Layer (AI Core):

3. Application Layer (Front-end/Back-end):

The AICC requires a robust, scalable, and modular AI technology stack:

- Data Storage: NoSQL (e.g., MongoDB/Cassandra) for unstructured content, and a relational database for user profile and metadata.
- Data Ingestion: Apache Kafka for real-time stream processing of content feeds.
- Machine Learning Frameworks: PyTorch or TensorFlow for training and deployment of models.
- Libraries: NLTK, spaCy, and Hugging Face Transformers for NLP tasks.
- Back-end: Python (Flask/Django) or Node.js to manage APIs and data flow.
- Front-end: React Native or Flutter for cross-platform (iOS/Android) mobile development, ensuring a responsive user experience.

5. Key Sections

1. Personalized Experience: Intent-Based Curation-

The AICC moves beyond simple topic-based filtering to intent-based curation. For instance, a user searching for “sustainable energy” might have different intents: (a) Professional Research (requiring academic papers and market reports), or (b) Personal Learning (requiring long-form educational videos and simplified articles). The AICC uses session context and explicit user mode selection to dynamically adjust the content score weights, prioritizing high-authority sources for (a) and engaging formats for (b). This level of contextual awareness is critical for hyper-personalization.

2. Quality Assessment and Trust

- To combat misinformation and low-value content, the AICC incorporates a robust Quality Scoring module. This module checks:
- Source Authority: Using external domain reputation services and user-blacklisted sources.
- Content Completeness: Analyzing word count, citation count, and depth of topic coverage.
- Bias/Fact-Checking Flags: Integrating with external fact-checking APIs (where possible) and using ML to detect highly sensationalized or unsubstantiated claims.

6. Result and discussion

The implementation of the AICC is projected to yield measurable improvements in media consumption metrics.

Results:

1. Increased Engagement: User time-on-app and content sharing rates are expected to increase by 40% due to higher content relevance.
2. Reduced Information Fatigue: A decrease in the number of sources a user manually checks, and a higher perceived utility of content consumed.
3. Content Diversity: Model metrics will demonstrate an avoidance of single-source or single-topic dominance, ensuring users are exposed to a balanced view.

Discussion suggests that while AI significantly reduces the time spent manually searching for content, human oversight remains necessary to manage ethical concerns, such as algorithmic bias and the potential for content manipulation. The future of media will likely involve an augmented human curator, where the AICC handles the scale and speed of content processing, while the human curator provides the unique voice and ethical guardrails. The

success of the AICC will hinge on its transparency regarding how content is scored and why it is recommended, fostering user trust.

The implementation of the AICC is projected to yield measurable improvements in media consumption metrics.

7. Conclusion

The AI Content Curator (AICC) represents a paradigm shift from a fragmented, low-relevance media experience to a hyper-personalized, high-quality information environment. By leveraging a cutting-edge technological stack and the Adaptive Preference Modeling (APM) algorithm, the AICC effectively solves the problem of information overload. The app functions as a sophisticated filter and synthesis engine, ensuring content delivery is precisely tailored to individual choices and professional needs, thus cementing its role as the definitive future of media consumption. While development must proceed with ethical considerations regarding bias and transparency, the transformative impact of the AICC on individual productivity and knowledge acquisition is undeniable.

References

1. AI for Hyper-Personalization, Ricci, F., Rokach, L., Shapira, B.: Recommender Systems: Introduction and Challenges. In: *The Handbook of Technology and Innovation Management*, pp. 544–561. Wiley, Hoboken (2018).
2. Technological Stack and NLP, Cambria, E., Schuller, B., Liu, H., Hussain, A., Thelwall, M.: Twenty years of sentiment analysis in computer science: perspectives and challenges. *Information Fusion* 84, 137–155 (2022).
3. Content Curation Algorithms (General), Das, A.S., Datar, M., Garg, A., Rajaram, S.: Google News personalization: scalable online collaborative filtering. In: *Proceedings of the 16th international conference on World Wide Web*, pp. 271–280. ACM, New York (2007).
4. Bias and Ethics in Curation, O’Neil, C.: *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown, New York (2016).
5. Addressing Information Overload, Eppler, M.J., Mengis, J.: The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. *The Information Society* 18(5), 325–344 (2002).
6. Deep Learning and Recommendation Systems, Zhang, S., Yao, L., Sun, A., Tay, Y.: Deep Learning for recommender systems: a survey and new perspectives. *ACM Computing Surveys* 52(1), 1–38 (2019).
7. The Future of Media and AI’s Role, Dwyer, T., Leipner, J., Jarke, J.: The role of artificial intelligence in media: A systematic review. *Journalism Studies* 23(16), 2533–2553 (2022).
8. Model Quality and Credibility Scoring, Graves, L.: *Deciding What’s True: The Rise of the Fact-Checking Industry and Its Challenge to Journalism*. Columbia University Press, New York (2018).
9. Real-Time Data Processing (Cloud/Infrastructure), Dean, J., Ghemawat, S.: MapReduce: Simplified Data Processing on Large Clusters. *Communications of the ACM* 51(1), 107–113 (2008).
10. The Role of User Interaction in Refinement, Shani, G., Gunawardana, A.: Evaluating Recommendation Systems. In: Ricci, F., Rokach, L., Shapira, B. (eds.) *Recommender Systems Handbook*, pp. 257–297. Springer, Boston (2011).

Integrating Talent Management into Sustainable Human Resource Strategies

Abha Pathak¹, Dr. Sonam Singh², Pallavi Thorat³, Bisini V⁴, Heena Satkut⁵

¹Engineering Department, Dr. D.Y.Patil College, Pune

²MBA Department, Pimpri Chinchwad University, Pune

³MCA Department, Pimpri Chinchwad University, Pune

⁴Commerce and Allied Program, Western College of Commerce and Business Management

⁵Western College Of Commerce and Business Management , Sanpada

abhanayak@gmail.com, Sonam.singh@pcu.edu.in, pallavi.thorat@pcu.edu.in, sini_33@yahoo.co.in,

heenalambe02@gmail.com

Abstract

Majority of the organizations give more importance to talented employees since it is more electronic a competitive edge to the market with the generation of opportunity with the employees who were talented through its system and to survive in the merciless business world. At the same time the organization that knows about the creative worker in order to achieve success in the modern era and to survive in the atmosphere of the argumentative business world. Talent management is one of the highly important requirements and treated in the proper way with the object to achieve the best outcomes. Companies have discovered that so as to achieve maximum results, there is a need to attract great people and at the same time, give them reasons to remain within the company. Consequently, one can deduce that the management of talent is very important. In today world, the development of talent management is regarded as the key element of the company and that is unique to the core barriers to medium- and long-term competitive reward in companies. The study concentrates on the analysis of talent management with respect to sustainability of human resource with respect to the demographical trends and its severity in the India Information Technology enabled services sector in Bangalore. The population of the study includes the ITES workers in the Bangalore city and sample which includes the ITES companies in Bangalore city. Data was obtained through the use or administration of 118 structured questionnaire on IT employee. The purpose of the research is to find out that the talent management is strategic tool of sustainable human resource within the city of ITES Bangalore.

Aims

The primary objective of this research is to understand the significance of talent management and its application as a strategic tool for fostering sustainable human resources. It further aims to explore the role of demographic factors, such as age and gender, in shaping effective talent practices in the ITES sector of Bangalore.

Research Design

This quantitative study utilized a structured questionnaire distributed among 138 employees from various IT and ITES organizations in Bangalore. The questionnaire employed both nominal and ordinal scales to gather relevant data. IBM SPSS version 20 was used for statistical analysis, employing tools such as the T-test, Chi-square, and correlation analysis to assess the formulated hypotheses.

Introduction

In the current fast changing business environment organizations are being expected not only to operate economically but also to show social and environmental responsibility. This has made the issue of sustainable human resource management (SHRM) strategic. At the same time, the current race after high performers has only increased the attention to the concept of talent management (TM) as one of the core mechanisms of organizational performance. The combination of these two areas, that is sustainability and talent management, has increasingly become critical to companies that aspire to achieve prosperity in the long run. Today, talent management does not only center on

recruitment and retention, but it also focuses on shaping moral leadership, diversity and inclusion, and alignment with environmental, social, and governance (ESG) objectives. The strategic decision-making process is further aided by the implementation of the current technological, artificial intelligence (AI), machine learning, and HR analytics that discover, nurture, and retain talent by considering employment-related practices that consider sustainability goals. Integrating sustainable HR practices implies that organizations will keep in mind both the long-term effect on the human capital practice on the performance of the company as well as the result on employee well-being, community relationships as well as environmental sustainability. An example is using predictive analytics in planning the workforce, getting the most out of the talent pipelines and streamlining the turnover associated with such turnover in a net carbon carbon footprinting. On the same note, digital technologies are facilitating more open and non-discriminatory talent development initiatives that promote the spirit of equality and sustainability. The sustainability of HR practices through integration of talent management demonstrates greater responsibility towards sound management practice in governance and sustainable development of the organization. In this paper, we are going to examine how both of these areas can work synergistically together, look at the strategic frameworks that will help make this integration, and see where technology can help in making this process more sustainable in terms of the talent approach. Talent management means that the high potential and ability individuals are strategically recruited and nurtured as well as retained in light of the long term objectives of the organization. Due to the process of globalization and the high rate of technological progress, companies cannot keep the sustainable workforce easily and without issues. War for talent is a terminology used by McKinsey & Company to lay emphasis on the competitive significance of human capital management. Firms that neglect talent and fail to identify it and develop it end up losing. Talent management involves several HR processes such as recruitment, onboarding, performance development, succession planning and retention. These aspects have to be aligned so that the organizational success is sustainable.

Talent Management

A well managed talent guarantees correct people taking up correct positions and being assisted to expand their abilities and skills. It responds to one of the biggest challenges of the organizations, which is the establishment of a motivated, high-performing workforce. The process involves strategic practices such as workforce planning, feedback, learning opportunity and leadership development. Talent management has been embraced as a key component in ensuring that an organization sustains strategic competitiveness as competition is gradually rising and employee requirements are changing.

Sustainable Human Resources

Sustainability in the field of human resources means the work that guarantees a long-term well being, satisfaction and efficiency of the employees. It combines both people management and larger sustainability objectives of environmental responsibility and social equity. Companies are also connecting sustainability plans and strategies to talent practices in order to create a resilient and future-oriented workforce. Studies indicate that HR policies can be aligned toward sustainability to enhance employee retention, innovation, and reputation of the organization.

Literature Review

No.	Author(s) & Year	Title	Key Focus	Key Findings	Relevance to TM & SHRM Integration
1	Collings & Mellahi (2009)	Strategic Talent Management: A review and research agenda	Defining strategic TM	Identifies TM as key to competitive advantage	Core framework for integrating TM into HRM
2	Schuler et al. (2011)	Global talent management and global talent challenges	Global TM practices	Emphasizes aligning TM with long-term strategy	Supports sustainability in HR planning
3	Ehnert (2009)	Sustainable Human Resource Management	Sustainability in HRM	Introduces the concept of SHRM	Foundation for merging TM with sustainability
4	Stahl et al. (2012)	Six principles of effective global talent management	Best TM practices	Identifies inclusive and long-term TM strategies	Promotes sustainable workforce planning
5	Tarique & Schuler (2010)	Global talent management: Literature review and future direction	Comprehensive GTM overview	Links TM with strategic HR outcomes	Key for integrating global SHRM strategies
6	Festing & Schäfer (2014)	Sustainable HRM and TM	Ethical talent practices	Aligns TM with CSR and long-term vision	Ethical basis for sustainable TM
7	Thunnissen et al. (2013)	Talent management and performance	TM impact on performance	Argues for holistic and inclusive TM	Essential for sustainable HR impact
8	De Prins et al. (2014)	SHRM: Balancing interests	Stakeholder theory in SHRM	Advocates balancing stakeholder needs	Aligns with inclusive TM practices
9	McDonnell et al. (2010)	Developing tomorrow's leaders	Leadership TM in	Succession planning as sustainable TM	Key to SHRM leadership pipeline
10	Kim & Scullion (2011)	Talent management in MNCs	Cross-cultural TM	Strategic alignment in global settings	Enhances sustainable global HR systems

11	Kramar (2014)	Beyond strategic HRM	Integrating ethics and sustainability	SHRM includes environmental & social goals	TM must support sustainable goals
12	Zaugg et al. (2001)	Sustainability in HR	HR in sustainable companies	Identifies sustainability indicators	Metrics for sustainable TM integration
13	Glaister et al. (2018)	Talent management: The route to sustainable organizational performance	Empirical SHRM linkage	TM leads to sustained performance	Empirical support for integration
14	Doh et al. (2011)	Responsible leadership and TM	CSR and TM	Leadership development as sustainable TM	Leadership vital for SHRM execution
15	Ashton & Morton (2005)	Managing talent for competitive advantage	Practical TM guide	Strategic TM aligns talent to business goals	TM as a driver of sustainable growth
16	Baum (2015)	Sustainable HRM in hospitality	Industry- specific SHRM	Sustainable TM practices in service industry	Sectoral insights for integration
17	Chugh & Bhatnagar (2006)	Talent management as high performance work practice	TM-HR performance link	TM as part of high-performance HR	Supports SHRM effectiveness
18	Lewis & Heckman (2006)	Talent management: A critical review	TM definitions and debates	Clarifies TM scope and strategic role	Ensures consistent SHRM integration
19	Wright & McMahan (2011)	Theoretical perspectives for SHRM	SHRM theories	Resource-based view, stakeholder theory	TM as strategic resource
20	Ehnert et al. (2016)	Sustainability HRM and	Future directions	Integrates ecological and social dimensions	Broadens scope of TM integration
21	Hughes & Rog (2008)	Talent management: A strategy for improving employee recruitment and retention	Retention strategies	Links TM to long- term engagement	Supports sustainable workforce

22	Farndale et al. (2010)	Talent management and institutional environments	External influences on TM	Aligns TM with societal sustainability	Aligns SHRM with local/global context
23	Swales (2013)	The ethics of talent management	Ethical critique	Critiques exclusive TM models	Advocates inclusive and sustainable TM
24	Jooss et al. (2021)	The role of TM in sustainable HRM practices	Empirical study	TM enhances practice impact SHRM	Validates integration approach
25	Skuza et al. (2013)	Talent and culture	Cultural fit in TM	TM must consider local cultures	Key for global SHRM strategies
26	Hartmann et al. (2010)	Talent management in emerging markets	TM sensitivity context	TM as adaptive, culturally relevant	Necessary for sustainable global SHRM
27	Boudreau & Ramstad (2005)	Talentship and the new paradigm for HR	Strategic analytics HR	TM must be evidence-based	Enhances sustainability through data
28	Ybema et al. (2020)	Inclusive Talent Management	Diversity in TM	Argues for inclusive, fair TM	Essential for sustainable SHRM models

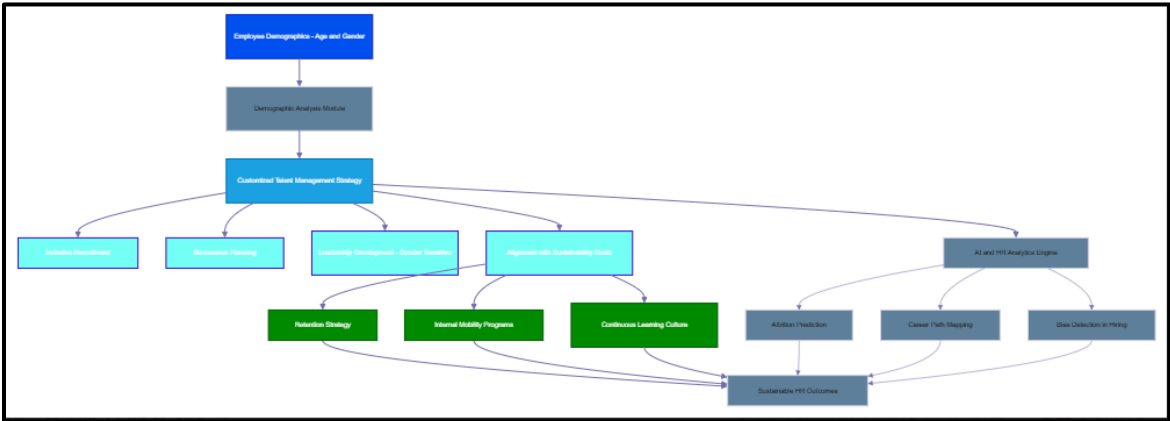
The present study offers a compelling insight into how Talent Management (TM), when applied strategically, plays a vital role in ensuring the sustainability of human resources, particularly within the dynamic ITES sector in Bangalore. The research strongly supports the notion that effective TM goes beyond simply hiring skilled professionals; it involves a continuous process of aligning employee capabilities with organizational objectives, nurturing talent through development initiatives, and retaining individuals through a supportive and engaging work culture.

Through the lens of demographic factors—namely age and gender—the study identifies that these dimensions significantly influence the success of TM strategies. Younger employees may bring innovation and adaptability, while more experienced professionals contribute with depth of knowledge and organizational loyalty. Gender inclusivity, on the other hand, reinforces diversity and enriches workplace dynamics, which are both essential to sustainable HR outcomes. The statistical evidence from ANOVA and T-tests confirms that demographic variations impact how talent is managed and utilized for long-term viability.

A key takeaway from this research is that sustainability in human resources is no longer just a policy goal; it is a strategic imperative. Organizations that fail to recognize the importance of tailored TM practices risk losing valuable human capital and competitive advantage. TM should therefore be approached holistically, incorporating employee aspirations, organizational goals, and socio-cultural contexts to build a resilient, future-ready workforce.

Moreover, the ITES sector, characterized by high competition, technological advancement, and global service demands, must prioritize continuous learning, job enrichment, and equitable HR practices. Retaining talent in such a fast-paced environment requires not only financial incentives but also opportunities for growth, recognition, and meaningful engagement.

This study underscores that sustainability in HR is intricately linked with how well organizations invest in their people—not just in recruitment, but across the entire talent lifecycle. As such, integrating TM as a strategic pillar of sustainable HRM is not merely beneficial—it is essential for organizational survival and success in an increasingly talent-driven economy.



Statement of the Problem

This study presents talent management as part of the strategic initiative to develop its sustainable human resource which is in the field of ITES in Bangalore. However, due to the high rate of technological changes and competition in the labour market, the study focuses on examining the impacts of demographic differences in terms of age and gender in the use of talents to add value to the industry, and retain talents within the industry.

Importance of the Study

This study provides insights for HR professionals in the ITES sector to enhance talent strategies using demographic understanding. The findings help organizations make data- driven decisions that align employee strengths with sustainable growth objectives. Understanding age and gender dynamics can lead to better utilization of talent, higher retention rates, and improved organizational outcomes.

Hypotheses

H1: There is a significant relationship between age and the effective utilization of talent. H0: There is no significant relationship between age and the effective utilization of talent. H1: There is a significant relationship between gender and sustainable talent management. H0: There is no significant relationship between gender and sustainable talent management.

Data Analysis and Interpretation:

This research study used IBM SPSS software, version 20 for data analysis. The data of research study are self- reported and were obtained by using a single method of a 5-point Likert scale i.e., a. Strongly agree, b. Agree, c. Neutral, d. Disagree, and e. Strongly disagree.

Table 1. Reliability analysis:

	Cronbach's Alpha Based on Standardized items	
Cronbach's Alpha		N of Items
0.938	0.937	20

(Primary Source)

Interpretation: The study constructs internal consistency is gauged by reliability. According to (Hair et al., 2013), a construct is considered credible if the Alpha value is larger than .70. Cronbach's Alpha(a) helped to evaluate the construct dependability. The results exposed that the overall item selected in the study was 20. HR/Management

express talent management is a great tool towards sustainable human resources were found reliable with the alpha value 0.937(-0.937).

Table 2.

ANOVA analysis for the impact of talent management towards Sustainability:

ANOVA with Friedman’s Test and Tukey’s Test for Non additivity

Source	Sum of Squares	df	Mean Square	Friedman's Chi-Square	Sig
Between People	1033.006	137	7.54		
Within People					
- Between Items	34.955	19	1.84	72.707	<.001
- Residual					
— Nonadditivity	2.878	1	2.878	6.125	.013
— Balance	1222.717	2602	0.47		
Total (Within)	1255.595	2603	0.471		
Total	1260.55	2622	0.481		
Grand Total	2293.556	2759	0.831		

Grand Mean = 3.9040

a. Kendall's coefficient of concordance W=. 015.

b. Tukey's estimate of power to which observations must be raised to achieve additivity = 2.831.

Interpretation: In the above ANOVA analysis there is a statistically significant difference between the group means with the significance value is 0.001 (i.e., $p = .001$), which is below 0.05. and, therefore, there is a statistically significant difference in the impact of talent management towards sustainable human resources.

Table 3. T-Test:

Evaluate the age impacts on effective utilization of talents By using one sample T- Test considering age a demographical factor to Analyse the talent utilization for sustainable human resources

One - sample Test

Statement	t	df	One-Sided p	Two-Sided p	Mean Difference	95% CI Lower	95% CI Upper
The organization prioritizes time to conduct interviews with potential candidates when a vacancy arises.	50.49	137	<.001	<.001	4.03623	3.8782	4.1943
The organisation devotes time and energy to attend filling of vacancies.	53.621	137	<.001	<.001	4.01449	3.8664	4.1625

The organisation ensures that vacancies do not remain open for a long period of time.	60.829	137	<.001	<.001	3.8913	3.7648	4.0178
Succession Planning (Internal Recruitment)	58.337	137	<.001	<.001	3.94928	3.8154	4.0831
External Recruitment	53.299	137	<.001	<.001	3.86957	3.726	4.0131
Organization makes an assessment of individual employee on regular feedback system.	57.126	137	<.001	<.001	4.11594	3.9735	4.2584
Organization actively creates developmental opportunities for employees.	55.373	137	<.001	<.001	4.05797	3.9131	4.2029
Managers concentrate on meeting with team members for formal career planning sessions.	52.325	137	<.001	<.001	3.92029	3.7721	4.0684
Organisation identify talents by their potential and performance.	48.005	137	<.001	<.001	3.86232	3.7032	4.0214

Interpretation:

Under one-sample t-test the P-value (0.000) (Significance 2 tailed) is less than the chosen significance level of (5%). Thus, null hypothesis refuted and believing that age was one of the tools which directly impact on talent management towards sustainability i.e it is an association and alternative hypothesis is accepted.

Table 4. T-Test:

Evaluate the gender impacts on talent management towards sustainable human resources management. By using one sample T- Test considering gender as an demographical factor to Analyse the talent utilization for sustainable human resources.

One - sample Test

Interpretation: Under one-sample t-test the P-value (0.000) (Significance 2 tailed) is less than the chosen significance level of (5%). Thus null hypothesis refuted and believing that age was one of the tool which directly impact on talent management towards sustainability i.e it is an association and alternative hypothesis is accepted.

Table 4. T-Test:

Evaluate the gender impacts on talent management towards sustainable human resources management.

By using one sample T- Test considering gender as an demographical factor to Analyse the talent utilization for sustainable human resources.

Statement

The organization prioritizes time to conduct interviews with potential candidates when a vacancy arises.

50.49

137

<.001

<.001

4.03623

3.8782

4.1943

The organisation devotes time and energy to attend filling of vacancies.

53.621

137

<.001

<.001

4.01449

3.8664

4.1625

The organisation ensure that vacancies do not remain open for a long period of time.

60.829

137

<.001

<.001

3.8913

3.7648

4.0178

Succession Planning (Internal Recruitment)

58.337

137

<.001

<.001

3.94928

3.8154

4.0831

External Recruitment

53.299

137

<.001

<.001

3.86957

3.726

4.0131

Organization makes an assessment of individual employee on regular feedback system.

57.126

137

<.001

<.001

4.11594

3.9735

4.2584

Organization actively creates developmental opportunities for employees.

55.373

137

<.001

<.001

4.05797

3.9131

4.2029

Managers concentrate on meeting with team members for formal career planning sessions.	52.325	137	<.001	<.001	3.92029	3.7721	4.0684
Organisation identify talents by their potential and performance.	48.005	137	<.001	<.001	3.86232	3.7032	4.0214
The organisation shows the importance to retaining the talents in the competitive environment.	45.735	137	<.001	<.001	3.7971	3.6329	3.9613
The Organization has a talent management initiative in place.	45.435	137	<.001	<.001	3.92754	3.7566	4.0985
The Organization initiates talent management as a top priority.	46.371	137	<.001	<.001	3.86957	3.7046	4.0346
Organization identifies talents from competitive market.	53.765	137	<.001	<.001	3.9058	3.7621	4.0494
Organizations identify talents by their performance.	50.976	137	<.001	<.001	3.89855	3.7473	4.0498
The Organization needs to improve in terms of talent management initiatives.	53.742	137	<.001	<.001	3.82609	3.6853	3.9669
The organization offer the ability to grow for employees.	49.448	137	<.001	<.001	3.92754	3.7705	4.0846
The organization rewardstop performers and help improve poor performance.	51.705	137	<.001	<.001	4.01449	3.861	4.168
Organization helps employees to feel a sense of ownership.	41.603	137	<.001	<.001	3.72464	3.5476	3.9017
Organization follows an employee retention strategy for talented employees.	52.204	137	<.001	<.001	3.85507	3.709	4.0011
Talent management is highly transparent.	41.629	137	<.001	<.001	3.61594	3.4442	3.7877
Gender of the respondents	34.948	137	<.001	<.001	1.49275	1.4083	1.5772

The data collected through the structured questionnaire were analyzed using IBM SPSS version 20. Reliability analysis using Cronbach's Alpha indicated a high internal consistency with a value of 0.937, which confirms the reliability of the survey instrument. An ANOVA test was conducted to examine the impact of talent management practices on sustainable human resource strategies, revealing a statistically significant difference ($p < 0.05$). Further, T-tests were applied to analyze the influence of demographic factors, namely age and gender, on the effectiveness of talent management.

Summary of Findings

1. There is a statistically significant correlation between age and the effective utilization of talent,
2. indicating that age plays a crucial role in how organizations manage and deploy their human resources.
3. Gender also has a meaningful influence on talent management strategies and their outcomes, supporting the need for gender-sensitive HR practices.
4. Talent management practices such as internal promotions, succession planning, performance evaluations, and retention strategies were found to be effective in fostering long-term sustainability in human resources.

Recommendations

Based on the findings, the following recommendations are proposed:

- Develop age-inclusive talent development programs that cater to different career stages.
- Promote gender equality in recruitment, promotion, and leadership opportunities.
- Implement robust performance appraisal systems linked to employee growth and retention strategies.
- Foster a culture of continuous learning and internal mobility to retain top talent.
- Align HR practices with organizational sustainability goals to ensure long-term success.

Limitations and Future Research

This study was limited to IT and ITES companies in Bangalore, which may not reflect broader industry trends. Future research should explore other sectors and geographical areas to generalize the findings. Additionally, Incorporating other demographic factors such as education, work experience, and job role could provide a more comprehensive understanding of talent management's impact on sustainability.

Conclusion

The current work is interesting in the sense that it provides an insight of the role of Talent Management (TM) as a strategic approach in making human resources sustainable and especially the rapidly growing ITES sector in Bangalore. The study is also evidence of the fact that positive TM is not just a question of acquiring competent workforce, it is a process of how to match the capabilities of the employees with the organizational goals, recruit and develop talent, and retain them in a working environment that is affirmative and involving. Viewing through the perspective of demographic factors namely age and gender, the research reveals that these aspects significantly have an implication on the success in TM strategies. The younger employees are able to introduce innovation and flexibility whereas more experienced professionals can work with a deep understanding of the profession and loyalty to the organization. On the contrary, gender inclusivity adds to diversity, and dynamics of the workspace which are the key criteria in sustainable HR outcomes. The statistical data obtained by ANOVA and T-tests prove that the differences in demographics influence the possibility of managing and consuming talent in the long-term. One of the learning points of this research is that sustainability in human resource can no longer be regarded as a policy objective; it is a strategy. Failure to start appreciating the significance of customized TM practices by a company exposes it to risks of losing human resources and competitive edge. TM ought to thus be addressed in an integrative manner in terms of employee ambitions, company objectives and socio-cultural backgrounds in a bid to create a formidable, future ready workforce.

References

1. Al Ariss, A., Cascio, W. F., & Paauwe, J. (2014). Talent management: Current theories and future research directions. *Journal of World Business*, 49(2), 173–179. <https://doi.org/10.1016/j.jwb.2013.11.001>

2. Baum, T., & Hai, N. T. T. (2019). Sustainable human resource management as a driver of sustainable organizational behavior: Review and framework. *Sustainability*, 11(16), 4525. <https://doi.org/10.3390/su11164525>
3. Beijer, S., & Gruen, B. (2016). A dynamic capabilities perspective on sustainable HRM. *The International Journal of Human Resource Management*, 27(2), 117–132. <https://doi.org/10.1080/09585192.2015.1072100>
4. Bratton, J., & Gold, J. (2017). *Human resource management: Theory and practice*(6th ed.). Palgrave.
- Cappelli, P., & Keller, J. R. (2014). Talent management: Conceptual approaches and practical challenges. *Annual Review of Organizational Psychology and Organizational Behavior*, 1(1), 305–331. <https://doi.org/10.1146/annurev-orgpsych-031413-091314>.
5. Collings, D. G., & Mellahi, K. (2009). Strategic talent management: A review and research agenda. *Human Resource Management Review*, 19(4), 304–313. <https://doi.org/10.1016/j.hrmr.2009.04.001>.
6. Collings, D. G., Mellahi, K., & Cascio, W. F. (2017). Global talent management and performance in multinational enterprises: A multilevel perspective. *Journal of Management*, 43(2), 540–566. <https://doi.org/10.1177/0149206314547152>.
7. De Prins, P., Van Beirendonck, L., De Vos, A., & Segers, J. (2014). Sustainable HRM: Bridging theory and practice through the ‘Respect Openness Continuity’ (ROC) model. *Management Revue*, 25(4), 263–284. <https://doi.org/10.1688/mrev-2014-04-DePrins>.
8. Ehnert, I. (2009). *Sustainable Human Resource Management: A conceptual and exploratory analysis from a paradox perspective*. Springer.
9. Ehnert, I., Harry, W., & Zink, K. J. (2014). Sustainability and HRM: A model and suggestions for future research. *Journal of Human Resource Management*, 25(8), 1007–1020. <https://doi.org/10.1080/09585192.2013.816863>
10. Ehnert, I., Parsa, S., Roper, I., Wagner, M., & Muller-Camen, M. (2016). Reporting on sustainability and HRM: A comparative study of sustainability reporting practices by the world's largest companies. *The International Journal of Human Resource Management*, 27(1), 88–108. <https://doi.org/10.1080/09585192.2015.1024157>.
11. Festing, M., & Schäfer, L. (2014). Generational challenges to talent management: A framework for talent retention based on the psychological-contract perspective. *Journal of World Business*, 49(2), 262–271. <https://doi.org/10.1016/j.jwb.2013.11.010>.
12. Gallardo-Gallardo, E., Thunnissen, M., & Scullion, H. (2020). Talent management: Context matters. *The International Journal of Human Resource Management*, 31(4), 457–473. <https://doi.org/10.1080/09585192.2019.1642645>.
13. Glaister, A. J., Karacay, G., Demirbag, M., & Tatoglu, E. (2018). HRM and performance—The role of talent management as a transmission mechanism in an emerging market context. *Human Resource Management Journal*, 28(1), 148–166. <https://doi.org/10.1111/1748-8583.12170>.
14. Hart, S. L., & Milstein, M. B. (2003). Creating sustainable value. *Academy of Management Executive*, 17(2), 56–69. <https://doi.org/10.5465/ame.2003.10025194>.
15. Jackson, S. E., & Seo, J. (2010). The greening of strategic HRM scholarship. *Organization Management Journal*, 7(4), 278–290. <https://doi.org/10.1057/omj.2010.37>.
16. Kramar, R. (2014). Beyond strategic human resource management: Is sustainable human resource management the next approach? *The International Journal of Human Resource Management*, 25(8), 1069–1089. <https://doi.org/10.1080/09585192.2013.816863>.

17. Kundu, S. C., & Gahlawat, N. (2016). Effects of perceived green HRM on employee green behavior. *International Journal of HRD and Management*, 6(1), 1–21.
18. Malik, A. R., Lenka, U., & Singh, S. (2017). Talent management and organizational effectiveness: Role of employee engagement. *Global Business Review*, 18(3), 1– 13. <https://doi.org/10.1177/0972150917693158>
19. Marler, J. H., & Boudreau, J. W. (2017). An evidence-based review of HR Analytics. *The International Journal of Human Resource Management*, 28(1), 3–26. <https://doi.org/10.1080/09585192.2016.1244699>
20. McDonnell, A., Collings, D. G., Mellahi, K., & Schuler, R. (2017). Talent management: A systematic review and future prospects. *European Journal of International Management*, 11(1), 86–128. <https://doi.org/10.1504/EJIM.2017.081253>
21. Mensah, J. K. (2015). A “coalesced framework” of talent management and employee performance. *International Journal of Productivity and Performance Management*, 64(4), 544–566. <https://doi.org/10.1108/IJPPM-07-2014-0100>
22. Muller-Camen, M., Croucher, R., Flynn, M., & Schroder, H. (2020). *Human resource management: A sustainable approach* (2nd ed.). Kogan Page.
23. Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. *Organization Studies*, 24(3), 403–441. <https://doi.org/10.1177/0170840603024003910>
24. Paauwe, J. (2004). *HRM and performance: Achieving long-term viability*. Oxford University Press.
25. Renwick, D. W. S., Redman, T., & Maguire, S. (2013). Green human resource management: A review and research agenda. *International Journal of Management Reviews*, 15(1), 1–14. <https://doi.org/10.1111/j.1468-2370.2011.00328.x>
26. Schiemann, W. A. (2014). From talent management to talent optimization. *Journal of World Business*, 49(2), 281–288. <https://doi.org/10.1016/j.jwb.2013.11.012>
27. Schuler, R. S., Jackson, S. E., & Tarique, I. (2011). Global talent management and global talent challenges: Strategic opportunities for IHRM. *Journal of World Business*, 46(4), 506–516. <https://doi.org/10.1016/j.jwb.2010.10.011>
28. Stahl, G. K., Björkman, I., Farndale, E., Morris, S. S., Paauwe, J., Stiles, P., Trevor, J., & Wright, P. (2012). Six principles of effective global talent management. *MIT Sloan Management Review*, 53(2), 25–32.
29. Ulrich, D., Brockbank, W., Johnson, D., Sandholtz, K., & Younger, J. (2008). *HR competencies: Mastery at the intersection of people and business*. Society for Human Resource Management.

Research papers

BMESH-2025

Crisis 2.0:

Unveiling Crisis Management Practices in Various Sectors

Jose Cris O. Sotto

Centro Escolar University Manila

Abstract

This metasynthesis of six qualitative studies offers a comprehensive review of various approaches to crisis management performed by public offices and financial institutions. It further revealed several critical insights into the factors affecting the implementation of crisis management, including emotion, digital adoption and competence, as well as cultural awareness.

1. Introduction

If there is one thing that is certain, it would be the fact that certain crises are harder to handle than others. Furthermore, there is no single best response to crisis management. It follows that the approach to managing an organizational crisis is highly dependent on the type of crisis that happens. A crisis is an unexpected event with potentially severe negative effects on a company, requiring swift action despite unclear causes or solutions (Gurdatta et al., 2023). Many models focus on enhancing organizational capacity to anticipate and mitigate crises, emphasizing the importance of being proactive rather than reactive.

Kuipers & Wolbers (2021) offered three types of organizational crisis as well as causes and management.

1.1. Crisis in the organization is present in situations when there is a tangible loss of life, property, or well-being in the organization and its immediate environment. Examples of this crisis type that had happened in the Philippines include 2024 MT Terra Nova oil spill incident (350 Pilipinas), and Ozone Disco inferno in March 1996 that took more than 150 lives. Crisis to the organization arises when concrete threats or risks evolve into major disasters, indirectly impacting other organizations. Aside from natural disasters, terrorism and industrial accidents that expose “institutional deficiencies” belong to this crisis type. The Manila bus hostage crisis in 2010 (PDI, 2014) and the 9/11 terror attacks belong to

1.2. Crisis about the organization happens when there is a sudden drop of this crisis type. Public and political support to an organization, affecting its organizational legitimacy (Alin et al., 2001 as cited in Kuipers & Worlbers, 2021). A crisis emerges when an organization loses its resources and authority, no longer has external support, or takes on tasks beyond its capacity or overestimates its abilities. In the Philippines, Prudential life Plans experienced bankruptcy in the early 2000s (Rappler, 2012). The 2003 Columbia space shuttle disaster that killed seven astronauts also offered lessons for managing high-risk operations (Howell & Dobrijevic, 2023).uzman, 2019).

2. Various sectors, various approaches?

Regardless of the type of crisis an organization experiences, a threat to long-term viability is presented. Decision makers of the organization need to be proactive in responding and preventing a crisis (Gurdatta et al., 2023). **But how does the approach of organization to crisis management vary across sectors? Are there special factors that they take into consideration when deciding the steps to take?**

Across various risk domains, the banking sector has endured significant shocks that have led to substantial reputational harm (Hutcheon & Clark, 2024); public organizations, on the other hand, are also subject to reputational challenges but maintaining their legitimacy is equally important (Wæraas & Byrkjeflot, 2012). Gürler et al. (2024) asserts that crisis management in public administration works to avoid service interruptions and protect institutional reputation by anticipating and addressing potential major issues before they worsen. During a crisis, the objective is to reduce or prevent harm to the institution's reputation and the public interest by utilizing every possible method. These are especially true in a developing country such as the Philippines, reputational scarring is more evident in financial institutions particularly banks. Additionally, the public image of the government and its many public offices is usually tainted by allegations of corruption, incompetency, and moral impotence.

Marker (2020)

	Fink	Gonzalez-Herrero and Pratt	Mitroff	Burnett	Relational Model Jacques
3-stage	4-stage	4-stage	5-stage	6-steps	4-cluster
Pre-crisis	Prodromal	Issues management	Signal detection	IDENTIFICATION goal formation	Crisis preparedness
Crisis	Acute	Planning-prevention	Probing, prevention	IDENTIFICATION environmental analysis	Crisis prevention
Post-crisis	Chronic	Crisis	Containment	CONFRONTATION strategy formulation	Crisis event management
	Resolution	Post-crisis	Recovery	CONFRONTATION strategy evaluation	Post-crisis management
			Learning	RECONFIGURATION strategy implementation	
				RECONFIGURATION strategy control	

Figure 1: Major Crisis Models

2. Major Crisis Models

Marker (2020) defines a crisis management model as an organized structure for dealing with unforeseen events, allowing managers to implement effective crisis response strategies even in situations with unclear origins and resolutions. Building organizational capacity is highly crucial in the effort to anticipate, avoid, and mitigate a crisis. Otherwise, damage may sprawl from reputation to the overall fiscal performance of the brand or organization.

3. Fink Model

In the book Crisis Management: Planning for the Inevitable, published in 1986, Steven Fink presented a crisis model with four stages: prodromal, acute, chronic, and resolution. Alert Media (2024) describes each of the four phases.

3.1. Prodromal Phase. This phase signifies the time from initial indications of problems to the outbreak of a crisis. Fink underlines the significance of proactive monitoring in this stage, as crisis managers aim to identify and handle possible crises before they worsen. The objective is to avert the crisis or minimize its extent by intervening early.

3.2. Acute phase. The acute phase starts when a triggering event sets off the crisis. At

this stage, crisis management plans are put into action, and crisis managers need to act quickly to minimize harm.

3.3. Chronic Phase. This stage includes addressing the lasting impacts of the crisis. For instance, following natural disasters such as hazardous spills, floods or hurricanes, teams focus on restoring infrastructure and managing the resultant destruction.

3.4. Resolution Phase. The last phase indicates the end of the crisis. It is a time for introspection and evaluation, where companies analyze the root causes to pinpoint mistakes and make adjustments to avoid similar crises.

5. Mitroff's 5-stage Crisis Management model

In 1994, Mitroff described five phases of crisis management: detecting crisis signals, investigating and preventing, controlling, recovering, and acquiring knowledge. These phases lead organizations through crisis management and learning, following a lifecycle progression (Marker, 2024).

5.1. Crisis Signal Detection. Mitroff acknowledged that organizations cannot anticipate every potential crisis because of their resource constraints. He observed that crises often come in groups, such as equipment malfunctions, external events, and dangers like product recalls, with corresponding preventive measures grouping together as well.

5.2. Risk Profiling and Prevention. In a study conducted in 1988, Mitroff, together with Terry C. Pauchant and Paul Shrivastava, suggested that organizations establish dual crisis portfolios to enhance crisis management. The first portfolio contains crises from all categories, while the second one includes preventive measures for each.

5.3. Containment. The dual portfolio strategy allows businesses to effectively manage different crisis scenarios with limited resources, enhancing readiness for various situations while maximizing resource allocation.

5.4. Recover). Mitroff stressed the significance of actively seeking out risk factors in the initial phases of crisis management in order to avoid or lessen.

5.5. Learnt'ng. Analyzing what went wrong and implementing corrective measures in the learning phase is crucial for enhancing future crisis responses.

6. Burnett Model of Crisis Management

John Burnett's 1998 crisis management framework comprises three phases—identification, confrontation, and reconfiguration—each containing two essential tasks, moving from establishing goals and analyzing the environment to developing strategies, executing them, and monitoring their progress. As the process continues, it becomes more challenging because of outside factors like time constraints, control problems, and limited choices for action. Burnett's model operates similar to a matrix, with internal crisis management initiatives being impacted by external limitations (Marker, 2020).

7. Relational Model of Crisis Management

In 2007, Tony Jacques questioned the linear model of crisis management, suggesting that prevention and preparation activities frequently coincide instead of following a strict sequence. He suggested a model where crisis management is combined with issues management, highlighting the potential for issues to develop into crises if left unaddressed. Jacques' model consists of four components - crisis preparedness, prevention, incident management, and post-crisis management - underscoring the significance of comprehending their interrelationships in order to mitigate crisis- related damages.

8. Attribution Theory and Situational Crisis Communication Theory

W. Timothy Coombs is commonly associated with two crisis communication theories. Attribution theory indicates that businesses suffer reputational and financial harm when the public blames them for a crisis. Situational crisis communication theory suggests that companies should tailor their crisis communication tactics according to the level of reputational risk presented by the crisis.

Under the Situational Crisis Communication Theory, Coombs said that assessing the cluster where a crisis belongs to is the first step that should be performed by the crisis manager. Based on the result of the situation and reputation analysis, the organization should respond with one of the three responses, namely deny, diminish and rebuild. Additionally, the company should perform “bolstering” by responding with one of the secondary approaches, namely reminding, accidental cluster and

intentional cluster.

Clusters	Primary Response	Secondary Response
Victim Cluster <i>The organization is the victim</i>	Deny <i>The organization assumes no responsibility</i>	Reminding Drawing attention to the company's past good deeds
Accidental Cluster <i>The organization unintentionally caused the crisis</i>	Diminish <i>The organization downplays the seriousness of the crisis</i>	Accidental Cluster <i>Praising stakeholders</i>
Intentional Cluster <i>The organization intentionally acted wrongly</i>	Rebuild <i>Involves apologizing</i>	Intentional Cluster <i>pointing out the organization's status as a victim, too</i>

9. Synthesis of major crisis management models

Crisis management models provide organized structures to assist organizations in handling unexpected situations by predicting, reacting to, and recovering from crises. The Fink Model highlights a sequence of four stages: prodromal, acute, chronic, and resolution phases, leading organizations from initial alerts to eventual recovery and contemplation. Mitroff introduces a five-step methodology, emphasizing the significance of managing different crisis situations efficiently through crisis signal detection, prevention, containment, recovery, and learning with dual portfolios. Burnett's model categorizes crisis management into identification, confrontation, and reconfiguration stages, prioritizing goal-setting and strategy implementation under external limitations, while Jacques' relational model questions the sequential method, highlighting the connection between crisis prevention and incident control. Both Attribution Theory and Situational Crisis Communication Theory emphasize the significance of public perception in crisis communication, promoting customized responses depending on the amount of reputational risk. Together, these models offer various methods for handling crises, promoting readiness, flexibility, and gaining insights from crises.

10. Statement of the Problem

This metasynthesis sought to investigate how public offices and private companies manage their reputation in the midst of an organizational crisis? The research aims to understand the factors affecting the task of crisis management. By conducting a metasynthesis of six qualitative studies that examine the practices of government offices and private companies, this study aims to synthesize existing knowledge to offer a comprehensive understanding of how various organizations approach their crisis management.

11. Methodology

Metasynthesis is a thorough, structured method utilized to examine and combine results from various qualitative studies, enabling researchers to uncover fresh perspectives and deeper understandings that go beyond the limits of single studies. It is especially useful for exploring intricate phenomena, experiences, and perspectives deeply. Mcleod (2024) details the important stages in performing a metasynthesis: (1) recognizing and assessing the methodological quality of pertinent qualitative studies; (2) gathering and combining main outcomes from various studies; (3) forming a thorough comprehension of the subject being researched; (4) creating fresh theories, models, or interpretations that push knowledge forward in the field; and (5) improving the credibility and wider influence of qualitative research.

Step 1: Formulate a question for a systematic review

Aligned with the study's main focus and aim, the systematic review question aimed to explore how various factors contribute to the effective execution of a crisis management plan by public offices and banking institutions. Moreover, the research looked into barriers that impede changes in organizational culture during crisis management and investigated methods to overcome these obstacles. Once the review question was determined, the subsequent step was to carry out an in-depth search of relevant literature.

Step 2: Conduct a search for relevant literature

The process of searching through literature entailed identifying and assessing appropriate sources to decide on the studies to incorporate. Keywords like "organizational culture," "public office crisis management," "bank crisis management," and "strategies for crisis management in banking" shaped this process. Studies on how organizational culture change impacts the implementation of crisis management strategies in banks were found using databases such Google. After collecting a considerable amount of research, the attention turned to assessing which information to incorporate.

Step 3: Choosing Articles

The articles were chosen according to five main factors: (1) being in English, (2) released from 2020 to 2024, (3) centered on crisis management in public offices and banking institutions, (4) fully available texts, and (5) easy to access. The first stage of screening involved examining the titles and abstracts of the research papers (Seers, 2015). If considered important, the entire article was additionally examined to evaluate how changing organizational culture aids in implementing crisis management strategies in banks. This led to the identification of eighteen articles, which were subsequently analyzed thematically.

Step 4: Analyzing the Data

Thematic analysis was applied to examine the chosen articles through four main steps: revisiting the articles for better understanding, recognizing main themes, extracting subthemes, and combining explanatory narratives. The review underscored the impact of organizational culture change among other factors that shape the way crisis management is performed in an organization.

12. Results

The metasyntesis of the six studies revealed several critical insights into the factors affecting the implementation of crisis management. Furthermore, it provides a comparison of the practices of public offices and private companies in preserving their reputation and legitimacy. The results are organized thematically, reflecting the core issues that emerged from the studies

13. PR's Role in Crisis Management

The Makassar-based Centre Point of Indonesia (CPI) reclamation project, started in 2015, has been met with continuous controversy over fears of coastal area commercialization and unclear coastal region zoning plans. In 2018, a report emphasized substantial harm to the environment and negative impacts on the economy, specifically impacting fishermen who have experienced decreased catches and now must venture further out to sea for their income. Akbar et al. (2020) analyzed the role of public relations in managing this crisis at an age of social media.

13.1. Crisis Identification

The first phase of public relations management for the Center Point of Indonesia (CPI) reclamation project involves identifying crises by gathering public complaints, which are received through various channels, including social media and the PR office. The PR team collects and analyzes this information, providing general explanations to the community about the project, while offering more detailed reports to the provincial government and relevant officials.

13.2. Analyzation of Complaints

The second phase of public relations management for the CPI reclamation project involves analyzing the complaints gathered to understand the crisis and develop appropriate responses. The PR team processes the data to address the complaints and prevent future issues, while serving as a liaison between the public, the government, and relevant technical units handling environmental concerns.

13.3. Stakeholder Management through Social Media. The public relations management team uses social media to inform staff about community complaints regarding the reclamation project and to communicate the steps taken by technical units, aiming to contain the crisis and address public concerns.

13.4. Mass Media Management. The public relations unit collaborates with the mass

media to disseminate information about the reclamation project, using interviews and control programs to inform and educate the public, aiming to prevent crises from arising due to complaints.

13.5. Complaint handling based on SOP. The Makassar government's public complaint handling is primarily guided by standard operational procedures (SOP), though the PR team occasionally addresses urgent issues outside the SOP when necessary.

14. Concept of Practice Switching

Unlike fast-response organizations, local administrations are not specifically trained to handle adversity; rather, they engage in crisis management as part of their broader responsibilities (Graf et al., 2023). Based on research in public administration and organizations, the researchers propose three aspects of administrative preparedness to help facilitate a smooth shift from routine actions to crisis response.

14.1. Organizational preparedness: agency training and frontline leeway. The research pinpoints two critical organizational aspects that help make shifting to crisis response easier: good training and allowing frontline flexibility. Thorough planning, simulations, cooperation between agencies, and equipment availability improve personnel readiness. Providing administrators with flexibility in decision-making is important for fostering innovation and adaptation, especially in times of quickly changing crises when traditional methods may not be effective.

14.2. Technical preparedness: digitization of the administration. Digitization greatly contributes to crisis management by improving the sharing of data, coordination, and allocation of resources. The research discovered a strong link between digitization and successful crisis transitions, but obstacles such as software restrictions emphasize the necessity for continued growth in the public sector.

14.3 Individual preparedness: prior staff experience.

Although it is believed that past crisis experience and teamwork can enhance crisis response, the research indicated that there is no notable link between staff experience and the effectiveness of crisis transition. On the other hand, interviews indicate that confidence in decision-making and communication improves with experience in handling crises.

15. Factors Affecting Crisis Management

Several factors affecting the public relations campaign of Makassar City regarding the reclamation project (Akbar et al., 2020).

1. **Cultural Similarity.** The cultural similarity between the public relations staff and the coastal Bugis Makassar community enhances the communication process and facilitates the delivery of messages about the reclamation project.

2. **Language understanding.** Language use is tailored to education and culture, with higher education levels fostering greater openness to new ideas and perspectives.

3. **Social media maximization.** The development of social media enhances the public information delivery performance of Makassar City's PR staff, allowing them to engage with the community interactively and disseminate information widely, while also balancing online communication with direct interactions, particularly in remote areas.

Various types of emergencies - such as cyber-attacks, natural disasters, political conflicts, and economic downturns - can have significant effects on society. Hazaa et al. (2021) critically evaluated crisis management through a systematic literature review, uncovering eight crucial factors and areas that offer valuable guidance to professionals and scholars on current achievements and future prospects in the field.

Several essential factors, such as information technology (IT), strategic management, and communication, impact crisis management. Demonstrated to decrease crisis risks and enhance decision-making, it is crucial for organizations to implement secure systems. Skilled employees and committed leadership are essential for effective strategic management, along with the necessity of continuously updating plans to be ready for unforeseen crisis. Effective communication, particularly on social media platforms, can help decrease uncertainty in times of crisis while also potentially fueling rumors, underscoring the necessity of well-defined communication plans. Furthermore, crisis management is improved by knowledge management through promoting learning and sharing knowledge, with effective leadership and corporate governance being crucial for adjusting strategies and maintaining transparency in times of crises (Hazaa et al., 2021).

16. Organizational Culture Change

Transforming organizational culture is important in improving a bank's ability to manage crises by modifying behaviors, practices, and attitudes that could impede recovery efforts (Adesegha, 2024). It helps develop a more flexible and strong organization, encouraging creativity, innovation, and competitiveness. Effective leadership is crucial for guiding employees during cultural shifts by promoting positive behaviors and addressing negative ones. Nevertheless, inadequate incorporation of cultural transformation in crisis management efforts

can pose challenges by causing resistance and organizational stagnation. Workers' acceptance is impacted by their expectations of change results, with faster adaptation stemming from positive anticipated outcomes. In the end, it is essential to ensure that organizational culture is in line with crisis management strategies in order to achieve successful recovery and maintain competitiveness in the long run.

17. How emotions enhance crisis communication

Coombs and Tachkova (2023) explore moral outrage as a key aspect of "sticky crises," suggesting a research agenda to improve crisis communication theory and practice by addressing the emotional and moral intricacies that present unique challenges for crisis managers.

A "scansis" happens when a crisis turns into a scandal, emphasizing the ethical dimension of crises caused by perceived moral transgressions. Moral outrage is different from anger, and it occurs when actions are perceived as deliberate, unfair, and driven by selfishness or discrimination. Situations that trigger moral outrage, like instances of misconduct in management, are seen as severe and labeled as "moral outrage-inducing crises." These varieties of crises, in addition to "wicked crises," are part of the larger group known as "sticky crises," which are defined by intense emotions and pose distinctive difficulties for crisis communication plans.

18. Techno-skepticism, mobile government and crisis management

Providing government services to its citizens via mobile technology is a common practice in both developed and developing countries. The "e-government gateway", for example, is a vast website that allows individuals to access public services through one central portal, made possible by the Turkish Republic of Northern Cyprus. Gürler et al. (2024) studied the relationship between crisis management and mobile government usage intention from a techno-skepticism perspective.

The research results showed that skepticism towards technology plays a role in how crisis management affects mobile government. Therefore, a high level of skepticism towards technology is a crucial and determining aspect in how often citizens make use of mobile government services. Doubts about technology were also discovered to be important in the use of mobile government. The study is a groundbreaking attempt to examine how techno-skepticism affects the relationship between crisis management and mobile government. It also offers different contributions to theory and practice, especially in the areas of mobile government

and the utilization of digital technologies.

Furthermore, Gürler et al. (2024) emphasize the importance of addressing techno-skepticism in crisis management and mobile government, underscoring the necessity for public institutions to consider this factor when creating mobile services. Securing data, maintaining privacy, and ensuring convenient access are essential for building trust among citizens and promoting ongoing utilization of mobile government services. Furthermore, it is important for technological advancements to support sustainable development objectives in order to contribute to lasting sustainability.

Culture and digital competence

Bloch (2024) looked into digital competence and cultural sensitivity as important considerations in handling intricate situations within the crisis situation realm. Finnish managers utilize a range of tailored crisis management tactics, such as predictive analytics, financial oversight, staff involvement, and digital health resources, especially during health emergencies like the COVID-19 outbreak. Digital platforms provide a way for officials to monitor public opinion and quickly address new problems. Social media analytics can help recognize misinformation trends, evaluate public reactions, and adapt communication tactics accordingly. Authorities can enhance their ability to adjust and respond efficiently to the evolving features of sociopolitical crises through the utilization of digital platforms (Klievink et al., 2017 as cited in Bloch, 2024).

Bloch (2024) furthered in his Master's thesis that Finnish managers view crisis management from a social constructivist perspective, highlighting the importance of social interactions and cultural contexts in influencing crisis reactions. Through comprehension of different cultural viewpoints and offering cultural competency training, they enhance communication and crisis response effectiveness, especially in health and sociopolitical crises. Collaboration with government agencies, private sector entities, and non-governmental organizations (NGOs) is viewed as advantageous in addressing crises by joining resources and knowledge towards action.

19. Synthesis of Findings

The concept that certain crises are harder to handle than others is a common notion in the field of crisis communication studies. According to the Situational Crisis Communication Theory (SCCT), the severity of crises differs depending on how responsibility for the crisis is perceived (Coombs & Holladay, 2002). The detailed examination of factors influencing crisis management is carried out through

metasynthesis of six studies, which compares how public offices and private companies manage their reputation in the midst of a crisis. One important finding from the study is the importance of public relations (PR) in handling crises, specifically in contentious projects like the Centre Point of Indonesia (CPI) reclamation in Makassar. The PR team of Makassar City's public offices utilizes social media, mass media, and public complaints to handle crises, while ensuring that their responses adhere to standard operational procedures. The importance of public relations in minimizing reputational harm is evident in how well these tactics work, particularly in a time of quick communication advancements.

Another important discovery focuses on the idea of "practice switching," where local governments, in contrast to quick-response groups, integrate crisis management into their regular tasks. Studies show that the readiness of the organization, such as training for employees and flexibility at the front lines, are key elements in successfully moving from regular operations to dealing with a crisis. Furthermore, being technically prepared, particularly through digitalization, improves data sharing and cooperation between agencies. Nevertheless, barriers like insufficient infrastructure continue to pose difficulties, underscoring the importance of ongoing progress in digitizing the public sector (Graf et al., 2023).

Moreover, the study looks into individual readiness, focusing on staff knowledge and collaboration, finding no clear link between previous experience and successful handling of crises. Nevertheless, interviews suggest that seasoned employees exhibit increased assurance in making decisions and communicating, highlighting the importance of individual development in emergency situations. This emphasizes the significance of organizational structures and individual skills in effectively handling crises.

Finally, the research highlights the importance of factors unique to public relations campaigns during crises, like cultural competence and digital communication, which are vital for community involvement (Bloch, 2024; Gürler et al., 2024). Utilizing social media to its fullest potential enables PR teams to engage with the public and handle issues promptly, however, maintaining a mix of online and in-person interactions is crucial.

20. Discussion

Legitimacy and reputation are closely linked, as both reflect an organization's social standing, though they differ in key ways: legitimacy is rooted in societal acceptance of the organization's adherence to norms, while reputation is competitive and aims for elevated esteem. Public organizations, such as tax authorities, may maintain legitimacy despite a poor reputation as long as their *raison d'être* remains unchallenged, with a strong procedural reputation being

particularly important for rule-bound agencies (Wæraas & Byrkjeflot, 2012).

21. Considerations of public offices vis-a-vis crisis management

Akbar et al. (2020) explains that the PR team of a local office is vital in handling the current crisis caused by factors such as environmental worries and economic effects on stakeholders. They start by identifying crises, collecting complaints from the public through different channels such as social media, and then explaining the situation to the community while also reporting specific issues to government officials. Reviewing these grievances enables the PR department to grasp the crisis's essence and craft suitable reactions, serving as a link between the public, authorities, and environmental organizations to avoid exacerbating the situation.

The role that social media plays in the overall task of the PR team in their effort to keep internal staff updated and communicate with the public, as well as to collaborate with mass media to educate and address project-related concerns in managing stakeholders is also very evident. This proactive communication is intended to reduce public dissatisfaction and avoid additional crises. Furthermore, the team follows standard operational procedures (SOP) when handling complaints but is adaptable in handling urgent matters that fall outside these procedures, prioritizing the swift resolution of immediate concerns to uphold public trust and minimize tensions.

Local governments, in contrast to quick-response groups, do not receive specialized training to manage crises, as dealing with crises is just one part of their overall duties. As per Graf et al.

(2023), successful handling of crises in these environments depends on three main elements of administrative readiness. Initially, readiness within an organization includes training staff, creating plans, running simulations, and empowering frontline workers to make decisions in dynamic circumstances. Adaptability is crucial in promoting innovation during unforeseen crises when conventional approaches are inadequate. Collaboration among organizations and having sufficient resources also improves the capacity to react efficiently.

The technical readiness aspect underscores the significance of digitalization in enhancing coordination, allocating resources, and sharing data during crises. Although digitization can support the transition process during crises, obstacles like obsolete software can impede advancement, highlighting the importance of ongoing development in public administration. Individual readiness, the third component, centers on the expertise of employees, which, although not directly associated with smoother crisis management transitions, can enhance trust in decision-making and communication. This points out how crucial it is to develop expertise and encourage flexibility in staff to enhance emergency reaction results.

22. Ingredients to a good crisis management

Several crucial factors, such as cultural similarity, language comprehension, and social media usage, impact the success of a public relations campaign such as that of Makassar City's reclamation project. Aligning cultural values with Bugis Makassar community members improves message delivery for PR staff, and adapting language to audience's educational levels boosts receptiveness to new concepts. In addition, optimizing social media enables the PR team to actively involve the community and share information broadly, while also maintaining a balance with face-to-face engagements in distant locations. Hazaa et al. (2021) also stress the significance of IT, strategic management, and communication in crisis management, underscoring the crucial role of secure systems, competent leadership, and successful communication strategies in addressing crises and minimizing uncertainty.

Changing the culture of an organization is essential for improving a bank's ability to manage crises by promoting adaptability, innovation, and competitiveness while dealing with behaviors that could hinder recovery efforts (Adesegha, 2024). Good leadership is essential for guiding employees during this change, but resistance and stagnation may occur if cultural changes are not properly incorporated into crisis management plans. Coombs and Tachkova (2023) also emphasize the importance of emotions, specifically moral outrage, in influencing crisis communication.

Moral anger stems from apparent moral transgressions and adds to what experts refer to as "sticky crises," which entail intricate emotional dynamics and ethical issues that make crisis management more difficult. Both changing organizational culture and using emotion-driven communication strategies are vital for effectively managing crises and building long-term resilience. Cultural awareness, cross-sectoral collaborations, as well as digital competence are additional ingredients to effective crisis management regardless of the organization type (Gürler et al., 2024; Bloch, 2024).

References

1. 350 Pilipinas. Statement on the MT Terra Nova Oil Spill Incident. <https://world.350.org/philippines/press-release/bataan-oil-spill/>
2. Adesegha, J. (2024). Organisational culture change as an antecedent for enhancing the successful implementation of a bank's crisis management strategies. Dutch Journal of Finance and Management.

<https://www.djfm-journal.com/download/organisational-culture-change-as-an-antecedent-for-enhancing-the-successful-implementation-of-a-15158.pdf>

3. Akbar, N. et al. (2020). The Role of Public Relations in Crisis Management in the Social Media Era.

International Journal of Global Community.

<https://journal.riksawan.com/index.php/IJGC-RI/article/view/73/64>

4. Alert Media. (17 August 2024). The 4 Stages of Crisis and How to Manage Them.

<https://www.alertmedia.com/blog/stages-of-crisis/>

5. Bloch, F. (2024). Integrating Cultural and Digital Components in Crisis Management: Strategies and Practices of Finnish Managers. Aalto University.

<https://aaltodoc.aalto.fi/server/api/core/bitstreams/fe562ed6-f32e-4890-ac5c-d0379098516f/content>

5. De Guzman, N. (18 March 2019). The Ozone Disco Tragedy: 25 Years and Still Burning Our Minds.

Esquire.

<https://www.esquiremag.ph/culture/lifestyle/ozone-disco-tragedy-1996-a1729-20190318-lfrm?s=clhece27hhqd04qgo4fhgediee>

6. Graf, F.; Lenz, A.; & Eckhard, S. (2023). Ready, set, crisis – transitioning to crisis mode in local public administration. Taylor and Francis.

<https://www.tandfonline.com/doi/full/10.1080/14719037.2023.2242851#abstract>

Gürler, S. et al. (2024). Mobile Government Use and Crisis Management: The

THE EFFECT OF ASSET GROWTH, BUSINESS RISK, AND GOOD CORPORATE GOVERNANCE ON FINANCIAL SUSTAINABILITY; An Empirical Study on Property & Real Estate Companies Listed on the Indonesia Stock Exchange for the Period 2020 - 2024

Siti Rahma Anggraini¹, Arry Eksandy²

¹Student of Accounting Department, Pembangunan Jaya University

²Lecturer of Accounting Department, Pembangunan Jaya University

Abstract

This study aims to explore the relationship between Asset Growth, Business Risk, Good Corporate Governance on Financial Sustainability. The study was conducted using Property and Real Estate sector companies as the research population with a research period of 2020 - 2024. The population in this study amounted to 94 with a sample of 10 companies, with a total of 50 observation data obtained based on the purposive sampling method. The research data uses secondary data collected from the company's website and the Indonesia Stock Exchange. This study uses a quantitative method with data processing using Eviews12. The results of the study show that asset growth, business risk, and good corporate governance have a significant effect on financial sustainability. However, in partial testing only asset growth and business risk affect financial sustainability, while good corporate governance does not affect financial sustainability.

Keywords: Asset Growth, Business Risk, Good Corporate Governance, Financial Sustainability

1. Introduction

In recent years, companies in Indonesia, especially companies engaged in the property sector, have experienced major challenges due to the decline in people's purchasing power for property. This decrease in people's purchasing power is caused by an increase in property prices. The resulting decline in people's purchasing power has a loss impact on the company in the form of a decrease in sales figures followed by a decrease in revenue and profits. The decline in revenue and profit experienced by each property sector company has an effect on the company's

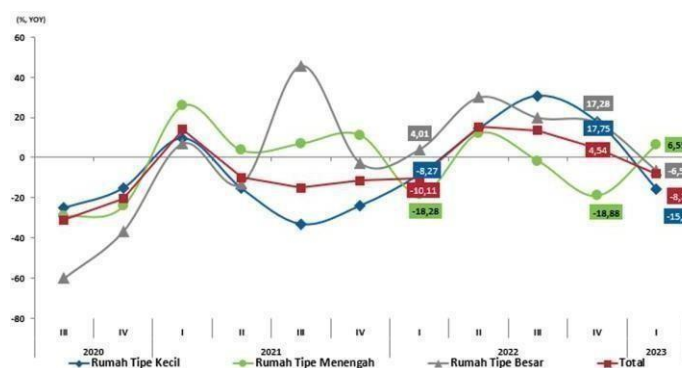


Figure 1.1 Sales Growth Chart

performance which will have an impact on the company's financial sustainability in the future.

As shown in figure 1.1, it is explained that in 2023 companies in the property sector will experience a decrease in sales. A decrease in sales is a warning for the financial health of a company where when the company experiences a decrease in sales, it will also decrease

the company's revenue and profit. In assessing *financial sustainability*, companies can do so by looking at their financial performance (Ramadhani et al., 2024).

Asset growth is an important factor in investment decision-making and improving the financial performance, which can ultimately strengthen the company's financial sustainability (Rahman, 2020). Several studies, such as those conducted by Wardaya et al. (2020), Fauzi et al. (2021), and Savitri et al. (2023), have found that asset growth has a significant positive effect on financial performance. However, Rahman (2020) argues the opposite, that asset growth actually has a negative effect on financial performance, so it does not necessarily improve the company's financial sustainability.

Business risk is uncertainty over earnings before interest and taxes due to variability in sales or asset-related income (Permana et al., 2021). Doubts about future business conditions may affect the return on operating capital opportunities. The greater the business risk, the more optimal

Good Corporate Governance (GCG) plays an important role in improving company performance by making resource management more efficient, effective, and productive and focusing on organizational and stakeholder goals (Anggara, 2023). The effective implementation of GCG has a positive impact on strategic decision-making, risk management, and operational efficiency, which increases profitability and stock value, thereby strengthening the company's ability to generate sustainable revenue. Research by Affes (2023) and Noviala et al. (2024) shows a significant positive relationship between GCG and financial performance, which means that GCG has an effect on financial sustainability. However, Dani et al. (2024) stated the opposite, that GCG has no effect on the company's financial performance and financial sustainability.

2. Literature Review and Hypothesis Development

2.1. Signalling Theory

Signalling Theory is a theory that was first introduced by Michael Spence in 1973. Spence explained that in conditions of information asymmetry, the party with better information is able to give signals to the other party to show its quality.

2.2. Stakeholder Theory

Stakeholder theory was first proposed by R. Edward Freeman in 1984 through a book he published entitled *Strategic Management: A Stakeholder Approach*. This theory explains that a company not only has responsibility to shareholders but also to all parties who have an interest in the company's activities, including employees, customers, suppliers, communities, and the government.

3. Financial Sustainability

Financial sustainability is an important indicator that shows the long-term sustainability of a company, thereby increasing investor confidence in the company's ability to survive in various situations (Oktoviyanti, 2023). Signal theory explains that companies need to provide information in annual reports related to management activities to achieve goals, so that owners can analyze the company's financial condition and financial sustainability level

(Ramadhani et al., 2024)

.

4. Asset Growth to Financial Sustainability

Asset growth includes an increase in different types of physical and financial assets that reflect the level of wealth of a company (Rahman, 2020). According to signal theory, asset growth is considered positive news that gives a good signal to the company, so good asset management will improve performance and support the company's long-term financial sustainability.

Research conducted by Wardaya, et al (2020), Fauzi, et al (2021), Savitri, et al (2023) revealed that asset growth has an influence on financial performance where this influence indicates that asset growth also affects financial sustainability

H1 : Asset Growth Has a Positive Effect on Financial Sustainability

Business Risks to Financial Sustainability

Business risk is the uncertainty of a company's future return projections (Jaribardi et al., 2022). Good risk management, supported by external stakeholder supervision, can minimize risks a**Good**

Corporate Governance for Financial Sustainability

Good Corporate Governance is a corporate governance system to run transparently, accountably, responsibly, fairly, and independently, with the aim of creating sustainable added value for stakeholders. Based on stakeholder theory, GCG directs managers to achieve high profits and long-term corporate sustainability (Ramadhan, 2020).

The role of Good Corporate Governance in the company has become an important foundation that can affect financial performance which will also affect the company's financial sustainability in the future. As research has been conducted by Malik (2022), Affes (2023), Pramanaswari (2024) and Noviala, et al (2024) states that in Good Corporate Governance there is a relationship that affects each other with financial performance where the financial performance owned by each company affects the company's financial sustainability and improve a company's finance

H2 : Good Corporate Governance Has a Positive Effect on Financial Sustainability

Asset Growth, Business Risk, and Good Corporate Governance on Financial Sustainability

Asset Growth, Risk Management, and Good Corporate Governance have a significant relationship with Financial Sustainability. Through effective asset growth with the support of proper business risk management, it is able to scale up the company's operations in a sustainable manner without having to sacrifice financial stability.

Overall, a combination of healthy asset growth, good business risk management, and strong corporate governance can create a strong foundation for corporate sustainability. This can not only improve financial performance but also strengthen the company's position in facing challenges in the future.

H3: Asset Growth, Business Risk, and Good Corporate Governance Simultaneously affects Financial Sustainability

5. Conceptual Framework

In accordance with the literature review and hypothesis development that has been described earlier, the conceptual framework that can be formed is as follows:

Figure 2 Conceptual Framework

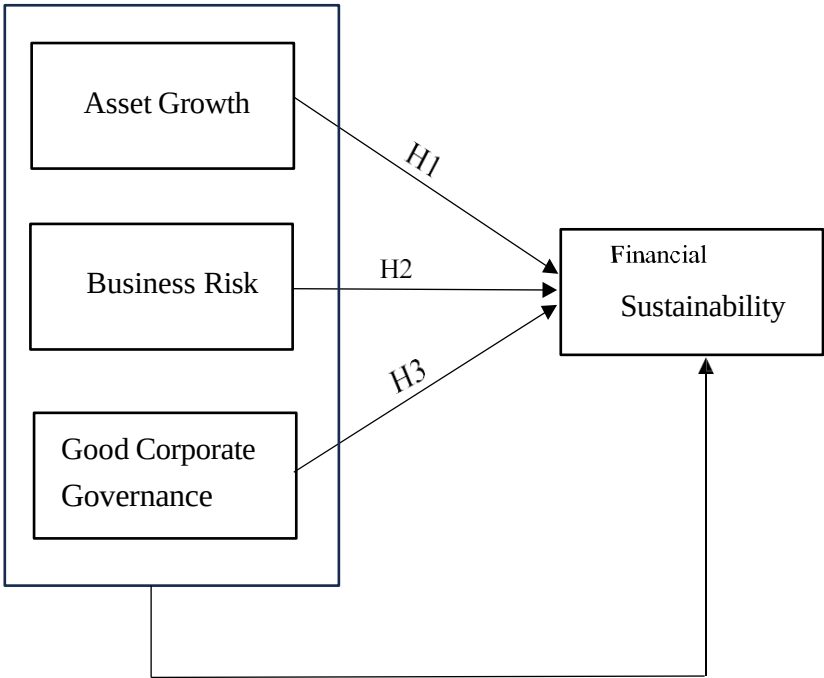


Figure 2 Conceptual Framework

6. Research Methods

The population studied is property and real estate sector companies listed on the Indonesia Stock Exchange during the period 2020 to 2024. The population consisted of 94 companies which were then eliminated so that the total sample used was 10 companies. The sampling technique is carried out using the purposive sampling method where the sample is taken based on criteria that are relevant to the research purpose.

7. Operational Definitions of Variables and Measurements Financial Sustainability

Financial sustainability is a company's ability to maintain long-term financial performance. One of the measurements that can be taken to measure financial sustainability is to use the sustainable growth rate where the sustainable growth rate is used to measure the maximum number of growth that is expected to be maintained and achieved over a long period of time. According to Dasuki (2022), the sustainable growth rate can be measured using the formula:

$$\text{SUSTAINABLE GROWTH RATE} = \text{RR} \times \text{ROE}$$

Information:

ROE : Return On Equity = Net Profit After Tax / Own Capital R : Retention

Rate = 1 – Dividend Payout Ratio Dividend Payout Ratio : Net Profit per Share / Dividend per Share

Asset Growth

Asset growth is a change in either an increase or decrease that occurs in the total assets owned by a company from one period to the next. According to Muhamad et al (2021), it was revealed that measuring the growth of assets that occur in a company can be done using the formula:

$$\text{Asset Growth : } \frac{\text{Asset}(t) - \text{Asset}(t - 1)}{\text{Asset}(t - 1)}$$

Information:

Asset (t) :

Current year's assets

Asset (t-1) :

Previous year's assets

Good Corporate Governance

Business Risk

Business risk is a possibility of losses that can affect corporate management model that the outcome of business decisions or activities. regulates the relationships that occur According to Fitriana (2024), it is revealed that between management, shareholders, business risk in a company can be measured using the board of commissioners, and the following formula:

Good Corporate Governance is a measurement of Good corporate governance in a company can be done using the Good Corporate Governance Index. In each sub-index assessment, a score of 1 will be given if the company meets the criteria and a score of 0 if the criteria are not met. In this study, the sub-index of good corporate governance consists of:

Business Risk =

$$\frac{\text{EBIT}}{\text{Total Asset}}$$

Table 1 Good Corporate Governance Index

COMPONENT	SUB INDEX
Index of Structure of the Council	There is a difference in the leadership of the board of commissioners and the board of directors
	The composition of independent commissioners is at least 1/3 of the total number of members of the board of commissioners
	The company's board meets 6 times a year
	There is a finance director
	The company has an audit committee, a nomination committee and a remuneration committee
Index of Procedures of the Council	The existence of a charter of the Board of Directors
	The existence of the charter of the Board of Commissioners
	The existence of the Audit Committee charter
	Code
	Disclosure of Audit Committee Membership
Disclosure of Information	Financial Statement Disclosure
	Disclosure of Background Information
	Disclosure of Responsibilities of the Board of Directors and Commissioners
	Audit Internal
	Publication of the Audit Committee Meeting
Share Ownership	Institutional Ownership
	Managerial Ownership
	Foreign Ownership
Shareholder Rights	The Company has disclosed information regarding related party transactions in its annual report
	The company's annual report has revealed the company's shareholding
	Company discloses dividend policy in annual report

8. Results and Discussion

9. Descriptive Statistical Analysis

Descriptive statistical analysis is a form of analysis carried out in research with the aim of providing an overview of the value of each variable being studied so that it is easy for readers to understand.

Table 2 Descriptive Statistical Analysis Results

Variable	N	Maximum	Minimum	Median	Mean	Std. Deviation
Financial Sustainability(Y)	60	6.774000	-9.350000	2.123500	2.184560	2.705141
Asset Growth (X1)	60	0.601000	-0.078000	0.034773	0.045330	0.091368
Business Risk (X2)	60	0.170000	0.000100	0.096000	0.089942	0.033751
Good Corporate Governance (X3)	60	1.000000	0.762000	0.882500	0.893680	0.077706

Source : Data processed, 2025

Based on the results of descriptive statistical analysis, the financial sustainability variable proxied using the sustainable growth rate showed a maximum value of 6,774. Meanwhile, the minimum value is -9.350000. The median number obtained in this study is 2.123500. Meanwhile, the average number (mean) obtained is 2.184560. The mean value is smaller than the standard deviation value obtained, which is 2.705141 which shows that this variable produces a wider distribution of data.

The asset growth variable in this study has a maximum value of 0.60100 obtained by. Meanwhile, the minimum value obtained in this study is -0.078000. The median value obtained in this variable is 0.034773 with a mean value obtained is 0.045330 which is smaller than the standard deviation value obtained which is 0.091368. This shows that the asset growth variable has a wider data spread.

The business risk variable in this study has a maximum value of 0.17000 obtained. Meanwhile, at the minimum value, a result of 0.000100 was obtained. The median value obtained was 0.096000. Meanwhile, the mean value obtained is 0.089942, which is greater

than the standard deviation, which is 0.033751, which means that the data owned by the business risk

The variable of good corporate governance has a maximum value of 1.000000. Meanwhile, the minimum value obtained in this study is 0.762000. The median value obtained was 0.882500. The mean value obtained is 0.893680, this figure is greater than the standard deviation value of 0.077706, which means that the data owned is not widely distributed.

Classic Assumption Test

The classical assumption test is a statistical requirement that must be met in multiple linear regression analysis using the Ordinary Least Squares (OLS) base. Based on the results of the selection of the panel regression model, it is explained that the type of regression model that will be used in this study is the Random Effect Model. The estimation technique used by the random effect model is the

General Least Squared method. According to (Sholihah, et al., 2024) reveals that regression analysis that does not use OLS is not necessary to perform classical assumption tests. This means that in this study, the classical assumption test does not need to be carried out because it uses the General sLeast Squared method.

Simultaneous Tests

Simultaneous tests in a study were performed to see if all the independent variables included in the model together exerted their influence on the bound variables.

Table 3 Simultaneous Test Results		
F-Statistic	F-Table	Prob (F-Statistic)
3.61	2.80	0.0019894

Source: Data processed, 2025

Based on the table above, it can be seen that this study has an F-Statistic probability value of 0.0019894 where this number shows a number that is below 0.05. In addition, when viewed from the comparison of F-Statistic with F-Table, the figure shows that the F-Statistic obtained in this study is larger than the F-Table.

From the description of the above results, it can be concluded that independent variables consisting of asset growth, business risk, and good corporate governance simultaneously

have an effect on the dependent variable, namely financial sustainability.

Partial Significance Test (T Test)

The partial significance test or the t test is carried out to determine the influence that occurs on each variable.

Table 4 Partial Significance Test Results

<i>Variable</i>	<i>Predictions</i>	<i>Coefficient</i>	<i>Prob.</i>	<i>Test Results</i>
Asset Growth	+	5.479862	0.0003	H1 accepted
Business Risk	+	30.10779	0.0034	H2 accepted
Good Corporate Governance	+	3.947245	0.3587	H3 rejected

Source: Data processed, 2025

Based on the results of the tests carried out, it can be concluded that:

1. The asset growth variable has a probability number of 0.0003 which means that the number is
2. less than 0.05 so it can be concluded that asset growth has an effect on financial sustainability
3. The business risk variable has a probability number of 0.0034 where the number is less than 0.05 so that it can be concluded that business risk has an influence on financial sustainability
4. Good Corporate Governance has a probability number of 0.3587 where the number is more than 0.05 so it can be concluded that good corporate governance has no influence on financial sustainability

Coefficient Determination Test

The results of the determination coefficient test explain how far a regression model can go in explaining the variation of independent variables that can affect dependent variables.

Table 5 Determination Coefficient Test Results

Variable	Coefficient	Prob.
Asset Growth	5.479862	0.0003
Business Risk	30.10779	0.0034
Good Corporate Governance	3.947245	0.3587
		R-squared0.190928
Adjusted R-Squared		0.138162
		F-Statistic3.618415
Total Observations		50

Source: Data processed, 2025

The determination of the value of the determination coefficient test is seen from the Adjusted R-Squared value obtained in the study. In this study, it has an Adjusted R-Squared value of 0.138162 which means that changes that occur in the financial sustainability variable can be explained by the variables of asset growth, business risk, and good corporate governance of 13%. While the remaining 87% is explained by other variables that were not studied in this study.

10. Conclusion

Based on the results of the tests that have been carried out and the discussion of the results of the research on the influence of asset growth, business risk, and good corporate governance on financial sustainability, the following conclusions were obtained:

1. Asset Growth (X1) shows the influence on Financial Sustainability in the Property and Real Estate sectors listed on the Indonesia Stock Exchange for the 2020-2024 period. The results of the t-test that have been carried out show that the probability number produced is $0.0003 < 0.05$ which means that asset growth has a significant effect on financial sustainability. Optimal asset growth tends to make the company have high financial sustainability because the company is considered able to balance business expansion with internal funding capacity
2. Business Risk (X2) shows the influence on Financial Sustainability in the Property and Real Estate sectors listed on the Indonesia Stock Exchange for the 2020-2024

period. The results of the t-test that have been carried out show that the probability number produced

3. by this variable is $0.0034 < 0.05$ which means that business risk can significantly affect financial sustainability. This indicates that good business risk management in the company is one of the main keys in maintaining the company's financial sustainability.

Good Corporate Governance (X3) shows that there is an influence provided by good c governance on the financial sustainability of companies in the Property and Real Estate sectors listed on the Indonesia Stock Exchange for the 2020-2024 period. From the results of the t-test that has been carried out, the variable good corporate governance has a probability number of $0.3587 > 0.05$, which means that good corporate governance significantly has no effect on the

company's financial sustainability. This happens because good corporate governance plays a more r

ole in supporting corporate transparency and accountability, not to support the company's financial sustainability.

Asset Growth, Business Risk, and Good Corporate Governance can simultaneously affect the company's financial sustainability. In this study, the simultaneous test obtained a result of $0.019894 < 0.05$ which means that asset growth, business risk, and good corporate governance can simultaneously have an effect on financial sustainability. This happens because the combination of asset growth, business risk, and good corporate governance together has a significant influence on financial sustainability because they form a solid foundation for companies to survive and thrive in a dynamic and uncertain business environment.

References

1. Affes, W., & Jarboui, A. (2023). The Impact Of Corporate Governance On Financial Performance: A Cross- Sector Study. *International Journal Of Disclosure And Governance*, 20(4), 374–394. <https://doi.org/10.1057/S41310-023-00182-8>
2. Anggara, R. (2023). Implementasi Prinsip Good Corporate Governance Pada Sekretariat Dprd Kota Banjar. *Wacana: Jurnal Ilmu Sosial Dan Ilmu Politik Interdisiplin*,

10(1), 521–531. <https://doi.org/10.37304/Wacana.V10i1.8754>

3. Dani, Rezkyta, E., Mawardah, L., Arifin, N. A., Okhtaviani, Putri, R., & Cahyati, R. D. (2024). PENGARUH GOOD CORPORATE GOVERNANCE , LEVERAGE DAN MANAGEMENT ASET TERHADAP KINERJA Abstrak. 4(2), 1–34.
4. Dasuki, R. E. (2022). Pendekatan Sustainable Growth Rate Dalam Mengukur Kinerja Keuangan Koperasi. Jurnal Co-Management, 4(3), 744–749. <https://journal.ikopin.ac.id/index.php/co-management>
5. Fauzi, A. F., & Puspitasari, E. (2021). Pengaruh Struktur Modal, Ukuran Perusahaan, Likuiditas Dan Pertumbuhan Aset Terhadap Kinerja Keuangan Perusahaan Yang Terdaftar Di Jakarta Islamic Index (Jii) Periode 2018-2020. Edunomika, 05.
6. Fitriana, G., & Arief, A. (2024). Pengaruh Risiko Bisnis , Kepemilikan Manajerial , Ukuran Perusahaan, Struktur Modal Terhadap Kinerja Keuangan Perusahaan. Jurnal Ekonomi Trisakti, 4(2), 787–796.
7. Luciana, L., Jhoansyah, D., & Danial, R. D. M. (2022). Analysis Of Business Risk, Company Size, And Sales Growth On Company Performance During Covid-19 Pandemic. Management Studies And Entrepreneurship Journal, 3(3), 1497–1504. <http://journal.yrpiiku.com/index.php/msej>
8. Malik, M. H. A. (2022). Pengaruh Good Corporate Governance Terhadap Kinerja Keuangan Sektor Aneka Industri Di Bursa Efek Indonesia. Owner: Riset& Jurnal Akuntansi, <https://doi.org/10.33395/owner.v6i3.919>
9. Muhamad, I., & Ardini, L. (2021). Pengaruh GCG, Kebijakan Dividen, FCF, Leverage, Dan Pertumbuhan Aset Terhadap Kinerja Perusahaan. Jurnal Ilmu Dan Riset Akuntansi, 10(6), 1–22.
10. Noviala, A., Dunakhir, S., & Samsinar. (2024). Pengaruh Good Corporate Governance Terhadap Profitabilitas. Jurnal Ekonomi, Bisnis Dan Manajemen, 7(1), 89–100. <https://doi.org/10.58192/ebismen.v2i2.773>
11. Permana, E., & Agustina, Y. (2021). Pengaruh Risiko Bisnis Dan Ukuran Perusahaan Terhadap Return On Asset Dengan Struktur Modal Sebagai Variabel Moderasi (Studi Pada Perusahaan Asuransi Yang Terdaftar Di Bursa Efek Indonesia Periode 2015-2018). Kompartemen : Jurnal Ilmiah Akuntansi, 19(1), 51. <https://doi.org/10.30595/kompartemen.v19i1.11224>

12. Rahman, M. A. (2020). Pengaruh Struktur Modal Dan Pertumbuhan Aset Terhadap Kinerja Keuangan Perusahaan Yang Terdaftar Dalam Jakarta Islamic Index (JII). *Akurasi : Jurnal Studi Akuntansi Dan Keuangan*, 3(1), 55–68. <https://doi.org/10.29303/Akurasi.V3i1.25>
13. Raadhani, A. P., & Sari, S. P. (2024). PERSPEKTIF PERUSAHAAN: MENILAI FINANCIAL SUSTAINABILITY MELALUI ANALISIS RASIO LIKUIDITAS, PROFITABILITAS, DAN AKTIVITAS. *Edunomika*, 08(04), 1–13.
14. Savitri, R. J., Unggul Purwohedhi, & Adam Zakaria. (2024). Pengaruh Firm Growth Dan Total Asset Turnover Terhadap Kinerja Keuangan. *Jurnal Akuntansi, Perpajakan Dan Auditing*, 4(3), 655–670. <https://doi.org/10.21009/Japa.0403.03>
15. Sesa, P. V. S., Arung Lamba, R. S., Patma, K., & Larasati, R. (2021). Pengaruh Struktur Modal, Risiko Bisnis, Pertumbuhan Penjualan Dan Ukuran Perusahaan Terhadap Kinerja Keuangan Pada Perusahaan Sub Sektor Real Estate Dan Properti Yang Terdaftar Di Bursa Efek Indonesia Tahun 2015-2019. *Jurnal Akuntansi Dan Keuangan Daerah*, 16(2), 10–20. <https://doi.org/10.52062/Jaked.V16i2.1938>
16. Wardaya, S., & Dhelo, M. M. (2020). Pengaruh Asset Growth, Arus Kas Operasi, Kepemilikan Institusional Dan Ukuran Perusahaan Terhadap Kinerja Keuangan perusahaan Manufaktur Yang Terdaftar Di Bursa Efek Indonesia Periode 2014 – 2017. *Media Akuntansi*, 32(1), 86–100.
17. Yunus, M. H., Mahfudnurnajamuddin, Semmaila, B., & Dewi, R. (2022). Pengaruh Risiko Bisnis, Literasi Keuangan Dan Inklusi Keuangan Terhadap Kinerja Keuangan Dan Keberlangsungan Usaha Pada Sektor Usaha Mikro, Kecil Dan Menengah Di Kota Palopo. *Journal Of Management Science (JMS)*, 3(2), 168– 199.

THE EFFECT OF INSTRUMENTAL BACKGROUND MUSIC TO CHOICE OF ACCOMODATION AT TRAVEL WEBSITE

Chici Putri Alifalexis Winata, Kevin Daffa Aditya Wijaya, Edelwis, Aries Yulianto

Departement of Psychology, Universitas Pembangunan Jaya, Indonesia

Abstract

Music has become a part of our lives and influences behavior in many ways. Numerous studies have proven that music can influence buying behavior. For example, slower-tempo music makes consumers decide to stay longer in restaurants and supermarkets. On the other hand, the internet has also become a part of our daily lives, especially in consumer behavior. Currently, many travel websites are widely available, which provide various needs, such as transportation and accommodation. This experimental research wanted to determine the effect of instrumental music on buying behavior at travel websites. Participants were 39 university students (16 males and 23 females), aged 18-22 years ($M = 19.385$, $SD = 0.907$). Participants were randomly assigned into three condition groups: no music, jazz music, and traditional music. Participants listened to music (or no music) while they visited a travel website for the island of Bali. At the end of the website, they were asked to choose one of their preferred accommodations. The results found that the highest frequency in resort and villa accommodations (38.46%), while in jazz music and traditional music were hostels (30.76%) and resorts (30.76%), respectively. The results of this study contribute additional information on the effect of music on consumer behavior, especially on travel websites.

Keywords: background music, travel website, buying behavior, experiment

THE EFFECT OF PROFITABILITY, LIQUIDITY, AND NON- DEBT TAX SHIELD ON CAPITAL STRUCTURE WITH FIRM SIZE AS A MODERATING VARIABLE

¹Kusumastuti, Gabriella Ageng Putri

²Dewi, Novy Silvia

¹Accounting Department, Pembangunan Jaya University

²Accounting Department, Pembangunan Jaya University

ABSTRACT

Introduction. Economic growth in a country affects the capital structure of a company. This study aims to determine the effect of Profitability, Liquidity, and Non-Debt Tax Shield on Capital Structure with Company Size as a moderator. These variables are indicators of the composition of the capital structure in a company. So it is interesting to do further analysis with the research object.

Data Collection Methods. This research using purposive sampling, where secondary data collected from financial reports on the Indonesia Stock Exchange.

Data analysis. The research object used is a property and real estate sector company will be analyzed quantitatively using SPSS.

Result and Discussion. In the research conducted, it resulted that the profitability of the company will affect the company's capital structure. This also applies to the non-debt tax shield which determines the company's capital structure. However, it is different from liquidity which does not affect the company's capital structure. In addition, all variables used cannot be moderated by company size.

Conclusion. the conclusion of this study is that hypotheses 1, 3, and 4 are accepted. While hypotheses 2, 5, 6, and 7 are rejected. Company size cannot moderate all variables used in this study. Further researchers need to add variables to deepen what factors will affect the company's capital structure.

Keywords : Liquidity, Non-Debt Tax Shield, Profitability, Capital Structure, Firm Size.

1. INTRODUCTION

There has been significant economic progress, particularly during the recovery period following the COVID-19 pandemic. This period, during which economic growth has enabled market reach to extend to the international stage (Puspitasari, 2022). has led every industry to maintain high competitiveness, resulting in expansion. Companies implementing expansion can employ various methods to ensure success, ultimately increasing profits and company value.

These sources of funds consist of short-term liabilities, long-term liabilities, preferred stock, and share capital (Martini, 2024). To meet this capital structure, a company can obtain both internal and external sources of capital. Internal funds are funds or capital derived from the company's own deposits, which can be profits from the company's operational activities or funds provided by the company's owners. In contrast, external funds are funds derived from external or foreign parties. The company obtains these funds through debt or loans from third-party banks or the sale of shares issued by the company. The funds from the available debt can be maximized, or the company can obtain more funds by conducting an Initial Public Offering (IPO) on the IDX.

A company's capital structure reflects its financial management in determining policies related to funding or capital sources. If a company lacks sufficient capital and the ability

to expand its business, it has the option of using debt (Fattimatuz et al., 2017). The ideal capital

structure is one in which the company's equity and debt have an equal composition (Geraldine et al., 2022). Furthermore, a company that heavily funds its capital structure and has a higher debt balance is more likely to rely on external funding.

Optimal funding sources or capital resources occur when a company is able to appropriately and optimally utilize long-term funds. These funds can then be reinvested to reduce capital costs, thereby achieving optimal company value. Optimal funding sources will reduce capital costs and further improve the company's financial efficiency. This efficiency will lead to a stable share price and drive share price increases. This rising share price will also build a positive corporate reputation, thereby increasing its value.

Profitability is a company's capability to generate profits or gains in each accounting period. Profitability is an indicator that demonstrates the effectiveness of a company's financial management, as measured by sales and investment profits. (Irmawati & Murtianingsih, 2024). Previous research by (Rehan et al., 2024), (Zandi et al., 2023), and (Fitri et al., 2022). However, on the contrary, research conducted by (Rákos et al., 2022), (Sutomo et al., 2020), and (Miswanto et al., 2022) which states that profitability has a negative influence on capital structure.

Liquidity is a company's ability to meet short-term obligations to investors and short-term creditors (Lilia, 2020). A company's liquidity is also a signal for investors, especially in determining the company's ability to meet these obligations in a timely manner. Research conducted by (Evinemi & Emengini, 2024), (Ghani et al., 2023), (Cahyani & Nyale, 2022) and (Putra & Deliza, 2023) states that liquidity has a negative influence on capital structure. This is in contrast to research conducted by (Rehan et al., 2024) and (Rákos et al., 2022) states that liquidity has a positive influence on capital structure. In research conducted by (Cahya et al., 2023) shows that liquidity has no influence on capital structure.

Non-debt tax shield is a form of tax saving instrument that reduces a company's tax burden. This tax arises from the depreciation of the company's fixed assets. (Wulandari & Artini, 2019). *Non-debt tax shield* can be defined as the charging of depreciation and amortization costs to profit or loss (Wulandari & Artini, 2019). Depreciation and amortization are cash flow sources of funding from within the company, so they can be a driver of reducing the debt borne by the company (Abubakar et al., 2020). In the non-debt tax shield research conducted by (Ghani et al., 2023), (Adeneye et al., 2023), (Chandra et al., 2020) shows that the non-debt tax shield has a significant positive influence on capital structure. Compared with research conducted by (Evinemi & Emengini, 2024), (Rehan et al., 2024), (Zandi et al., 2023), (Simatupang et al., 2019) which states that the non-debt tax shield has no influence on a company's capital structure.

Company size can be used to describe how a company sells the goods or services it produces. It also reflects the total assets owned by the company (Rifiana et al., 2021). Companies that are considered large companies will certainly require substantial funding in direct proportion to their growth. Research conducted by (Damina et al., 2022) and

(Zandi et al., 2023) which states that company size has a positive influence on capital structure. Compared with research conducted by (Wijandari, 2020) with the statement that company size has no impact on capital structure.

The Jakarta Composite Index (JCI) on the Indonesia Stock Exchange in the property and real estate sector fluctuated between 2019 and 2023, with a significant decline in 2020 due to the Covid-19 pandemic. Prices recorded successive increases of 12.54%, -21.23%, - 19.11%, -8.00%, and 0.41%. The current year has seen the property and real estate sector stock index rebound. Companies with declining stock values are a red flag indicating financial difficulties. In the first half of 2024, PT Modernland Realty (MLDN) was the company with the highest losses. The next highest-ranking property company was PT Agung Podomoro Land (APLN), which ranked second (Kompas.com, 2024). The phenomenon that occurred during the last 5 years is in accordance with the opinion (Meilani & Wahyudin, 2021) The magnitude of debt a company incurs can lead to financial distress, such as agency costs and bankruptcy costs.

Based on previous research gaps, including inconsistent results and phenomena in the property and real estate sector in Indonesia, researchers will examine the factors influencing capital structure. This research is expected to identify factors influencing capital structure based on current conditions and phenomena.

A. LITERATURE REVIEW

Agency Theory

Agency theory states that the working relationship between the principal and the agent, this theory was popularized by Jensen and Meckling in 1976. This working relationship explains that one or more people as principals will appoint others as agents to carry out certain work. However, in practice, sometimes agents are motivated to fulfill their personal needs, with differences in interests between agents and principals, causing agents to manipulate performance reports that show good performance for their own benefit.

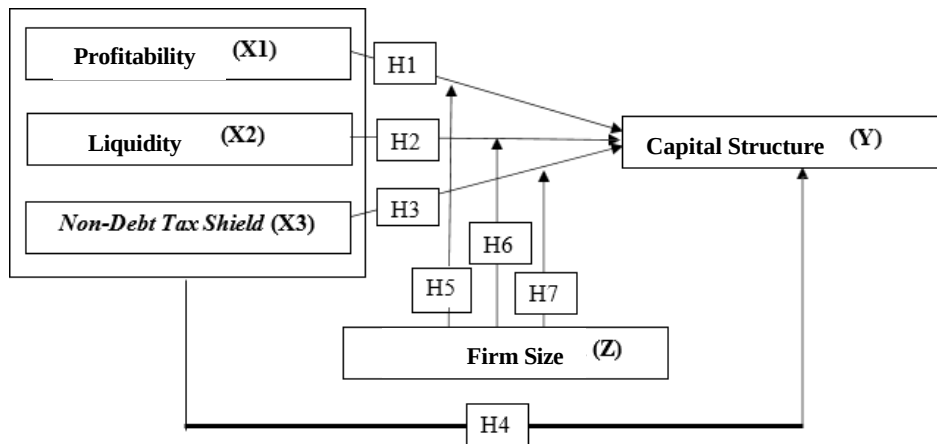
Pecking Order Theory

The theory was first put forward by Donalson in 1961 and declared as pecking order theory by Myers in 1984. (MYERS, 1984). The pecking order theory suggests that internal financing is the primary source of financing in a company's capital structure, compared to debt or external financing.

Trade-Off Theory

Trade – off Theory Contains the relationship between the risk of bankruptcy due to financial distress, taxes, and the use of debt. The optimal relationship results from a trade- off between the profits and losses that can occur from the funding sources used by the company (Adeneye et al., 2023).

Based on the theoretical explanation above, the framework used in this research is:



Picture 1. Conceptual Framework

Profitability reflects a company's ability to generate profits from its operations and effective management. With high profits, a company will reduce external funding. Research conducted by (Rehan et al., 2024) (Miswanto et al., 2022) (Irmawati & Murtianingsih, 2024) shows that profitability has a negative and significant influence on a company's capital structure.

H1 : Profitability influence Capital Structure.

Company liquidity represents a company's ability to meet short-term financial obligations as they mature. Liquidity is a risk when a company is unable to meet short-term obligations. The current ratio can be a proxy for a company's liquidity ratio. Research conducted by (Wulandari & Artini, 2019), (Geraldine et al., 2022), (Cahyani & Nyale, 2022) found that liquidity has a negative and significant effect on capital structure.

H2 : Liquidity influence Capital Structure.

A high non-debt tax shield in a company indicates a low capital structure, as the use of debt as capital is also low. High depreciation reflects the company's large fixed assets, thus increasing tax benefits and internal funding. (Rifiana et al., 2021). Research by (Irmawati & Murtianingsih, 2024) states that the non-debt tax shield has a negative and significant influence on capital structure.

H3 : Non-Debt Tax Shield influence Capital Structure

Profitability, liquidity, and the non-debt tax shield are factors that can influence financial management in determining a company's capital structure. Profitability is an aspect that can reflect the amount of retained earnings from a company's operating results.

H4 : Profitability, Liquidity, and Non-Debts Tax Shield simultaneously influence Capital Structure

Companies influenced by size can also generate greater profits. Larger companies can generate higher profits and retain earnings, which can be used as capital, in accordance with the pecking order theory.

H5 : Company Size moderates the relationship between Profitability and Capital Structure

Large companies will incur significant costs. According to the pecking order theory, greater liquidity will encourage companies to use internal funds as a source of capital rather than debt.

H6 : Company Size moderates the relationship between Liquidity and Capital Structure

A non-debt tax shield is an indicator that a company has substantial fixed assets. A large company size indicates that the company has substantial fixed assets.

H7 : Company Size moderates the relationship between Non-Debt Tax Shield and Capital Structure

B. RESEARCH METHODOLOGY

This research uses a quantitative method. Quantitative research collects and analyzes quantitative data, then tests theories and hypotheses on the relationship between independent variables with the dependent variable. The research object used in this study is the capital structure of property and real estate sector companies listed on the Indonesia Stock Exchange in the period 2019-2023. The sample selection technique used in this study is based by the purposive sampling method. Quantitative data analysis is processed using SPSS. Therefore, the company criteria used in determining the research sample are as follows:

1. Property and real estate companies listed on the Indonesia Stock Exchange for the 2019-2023 period.
2. Property and real estate companies included in the Indonesia Stock Exchange's special monitoring category for the 2019-2023 period.
3. Companies that did not report net income in their 2019-2023 financial statements.
4. Companies that experienced losses during the 2019-2023 study period. The following is a table of operational variables in the research:

Table 1. Operational Variable

No	Variable	Name	Measurement	Scale
1.	Dependent Variable	Capital Structure	$DER = \frac{Total\ Liabilities}{Total\ Equity}$ (Weygandt et al., 2019)	Ratio
2.	Independent Variable	Profitability	$NPM = \frac{Net\ Profit}{Revenue}$ (Weygandt et al., 2019)	Ratio
3.		Liquidity	$CR = \frac{Current\ assets}{Current\ Liabilities}$ (Weygandt et al., 2019)	Ratio
4.		Non-debt tax shield	$NDTS = \frac{Depreciation}{Total\ Asset}$ (Ghani, et al., 2023)	Ratio
5.	Moderating Variable	Size	$UK = Ln(Total\ Assets)$ (Ahmed & Sabah, 2021)	Ratio

C. RESULT AND DISCUSSION

Descriptive Analysis

The following table shows the results of the descriptive statistical tests in this study.:

Table 2

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
X1	85	0.00	0.99	0.2612	0.18705
X2	85	0.87	6.54	2.6229	1.41075
X3	85	0.00	0.14	0.0318	0.02935
Y	85	0.09	1.95	0.7012	0.44690
Z	85	15.60	31.83	25.9351	4.44279
Valid N (listwise)	85				

Regression moderation analysis is a statistical analysis technique used to test a moderating variable, namely company size in this study, which can weaken or strengthen the relationship between the independent variable and the dependent variable.

Table 3. Regression Analysis

Coefficients^a

Unstandardized Coefficients			Standardized Coefficients		
Model	B	Std. Error	Beta	t	Sig.
1	(Constant)	0.675		14.358	0.000
	MC_X1	-1.466	-0.614	-4.905	0.000
	MC_X2	-0.042	-0.131	-0.829	0.410
	MC_X3	-4.920	-0.323	-1.913	0.059
	MC_Z	-0.007	-0.067	-0.481	0.632
	MCX1_MCZ	0.143	0.192	1.433	0.156
	MCX2_MCZ	0.015	0.159	0.977	0.332
	MCX3_MCZ	0.545	0.140	0.688	0.493

a. Dependent Variable: Y

So the results of the moderated regression analysis in this study are as follows:

$$Y = 0.675 - 1.466NPM - 0.042CR - 4.920NDTS - 0.0007UK - 0.143NPM*Z + 0.015CR*Z + 0.545NDTS*Z + 0.714e$$

Table 4. One Sample Kolmogorov-Smirnov

One-Sample Kolmogorov-Smirnov Test

Unstandardized Residual

N		85
Normal Parameters ^{a,b}	Mean	0.0000000
	Std. Deviation	0.36155119
Most Extreme Differences	Absolute	0.084
	Positive	0.084
	Negative	-0.055
Test Statistic		0.084
Asymp. Sig. (2-tailed)		0.200 ^{c,d}

b. Calculated from data.

c. Lilliefors Significance Correction.

d. This is a lower bound of the true significance.

Based on the table, the test results obtained Asymp. Sig (2-Tailed) of 0.200 which shows a value greater than 0.05, so it can be concluded that all data in the sample studied is normally distributed. The purpose of the multicollinearity test is to determine whether all the data in the sample used does not have a very strong relationship or does not have a perfect linear relationship, in other words, there is no relationship between the independent variables with each other. (Sekaran, 2020).

Table 5. Multicollinearity Test

Coefficients ^a								
Unstandardized Coefficients				Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	0.675	0.047		14.358	0.000		
	MC_X1	-1.466	0.299	-0.614	-4.905	0.000	0.543	1.842
	MC_X2	-0.042	0.050	-0.131	-0.829	0.410	0.338	2.962
	MC_X3	-4.920	2.572	-0.323	-1.913	0.059	0.298	3.356
	MC_Z	-0.007	0.014	-0.067	-0.481	0.632	0.438	2.282
	MCX1_MCZ	0.143	0.100	0.192	1.433	0.156	0.475	2.105
	MCX2_MCZ	0.015	0.016	0.159	0.977	0.332	0.320	3.128
	MCX3_MCZ	0.545	0.791	0.140	0.688	0.493	0.207	4.832

a. Dependent Variable: Y

If the tolerance value is >0.1 or the VIF value is <10 , then the model does not show symptoms of multicollinearity or passes the multicollinearity test. The independent variables and moderating variables used in the study have tolerance values >0.1 or VIF values <10 . Therefore, it can be concluded that the variables used do not show symptoms of multicollinearity or pass the multicollinearity test.

The heteroscedasticity test is also used to examine the uneven distribution of residual data. This test uses the Spearman's Rho test, which correlates the residual values of each variable (Sekaran, 2020).

Table 6. Heteroscedasticity Test

No	Variable	Sig.	Information
1.	NPM (X1)	0.093	No heteroscedasticity
2.	CR (X2)	0.984	No heteroscedasticity
3.	NDTS (X3)	0.456	No heteroscedasticity
4.	SIZE (Z)	0.715	No heteroscedasticity
5.	NPM*SIZE	0.243	No heteroscedasticity
6.	CR*SIZE	0.700	No heteroscedasticity
7.	NDTS*SIZE	0.065	No heteroscedasticity

The heteroscedasticity test is also used to examine the uneven distribution of residual data. This test uses the Spearman's Rho test, which correlates the residual values of each variable.

The autocorrelation test is used to determine whether the regression models used are correlated with each other. According to (Sekaran, 2020), if the Durbin-Watson value is between -2 and +2 ($-2 < \text{Durbin-Watson value} < +2$), it can be concluded that the regression equation model does not experience autocorrelation or passes the autocorrelation test.

Table 7. Autocorrelation Test

Model Summary ^b					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0.588 ^a	0.345	0.286	0.37763	0.873

a. Predictors: (Constant), MCX3_MCZ, MC_X1, MC_X2, MC_Z, MCX1_MCZ, MCX2_MCZ, MC_X3

b. Dependent Variable: Y

Based on table 4.7, the Durbin-Watson value of 0.873 is between ($-2 < 0.873 < +2$). Therefore, it can be concluded that the regression equation model in this study does not experience

autocorrelation or passes the autocorrelation test.

Table 8. F-Test

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.930	3	1.643	11.236	0.000 ^b
	Residual	11.846	81	0.146		
	Total	16.776	84			

a. Dependent Variable: Y

b. Predictors: (Constant), X3, X1, X2

Based on the table, it shows that profitability (X1), liquidity (X2), non-debt tax shield (X3) simultaneously have an influence on capital structure as a dependent variable.

Table 9. Coefficient of Determination Test

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.588 ^a	0.345	0.286	0.37763

a. Predictors: (Constant), MCX3_MCZ, MC_X1, MC_X2, MC_Z, MCX1_MCZ, MCX2_MCZ, MC_X3

Based on the table, capital structure as the dependent variable accounts for 0.286, or 28.6%. The remaining 0.714, or 71.4%, is explained by other variables not included in this research model.

Table 10. T-test

No	Hypothesis	Nilai sig.	Hasil
1.	NPM - DER	0.000	H1 accepted
2.	CR – DER	0.645	H2 not accepted
3.	NDTS – DER	0.048	H3 accepted
4.	NPM*SIZE – DER	0.156	H5 not accepted
5.	CR*SIZE – DER	0.332	H6 not accepted
6.	NDTS*SIZE - DER	0.493	H7 not accepted

Discussion

The Effect of Profitability on Capital Structure

Based on the T-test, the partial test result for the profitability variable was 0.000 (<0.05). This concludes that the first hypothesis is accepted, stating that profitability influences capital structure. The results of this study align with research conducted by (Rehan et al., 2024), (Irmawati & Murtianingsih, 2024), and (Zandi et al., 2023).

The results of this study are in accordance with the Pecking Order Theory which states that companies will maximize internal funding such as retained earnings as a capital structure, compared to external funding.

The Effect of Liquidity on Capital Structure

The partial test results showed a liquidity variable of 0.645 (>0.05), indicating that the second hypothesis was rejected, as in this study, liquidity had no effect on capital structure. This result is consistent with research conducted by (Evinemi & Emengini, 2024), (Cahya et al., 2023), and (Irmawati & Murtianingsih, 2024).

This is not in accordance with the Pecking Order Theory which states that high liquidity will maximize internal funding because the company is considered capable of meeting short-term obligations.

The Effect of Non-Debt Tax Shield on Capital Structure

The T-test results in the partial testing of the non-debt tax shield variable showed a value of 0.048 (<0.05), which confirms the acceptance of the third hypothesis, where the non-debt tax shield has an influence on capital structure. This consistent result is in accordance with research conducted by (Cahya et al., 2023) and (Chandra et al., 2020).

The pecking order theory, which applies to the Non-Debt Tax Shield variable, states that companies with a high NDTs ratio will reduce external financing in the form of debt. Funds obtained from the company's fixed assets indicate sufficient internal funds. These funds will be diverted to internal financing.

The Effect of Profitability, Liquidity, and Non-Debt Tax Shield Simultaneously on Capital Structure

The F-test yielded a value of 0.000 (>0.05). This value demonstrates that the three independent variables used in this study are less than the alpha value of 0.05, or 5%.

It can be concluded that profitability, liquidity, and the non-debt tax shield simultaneously influence capital structure. These results support the fourth research hypothesis, which states that the independent variables used in this study contribute to determining a company's future capital structure.

The Effect of Profitability on Capital Structure with Company Size as a Moderating Variable

The results of equation 2 show a value of 0.156 (>0.05), indicating that company size neither strengthens nor weakens the relationship between profitability and capital structure. Therefore, the fifth hypothesis is rejected. The results of this study align with research conducted by (Vinda, 2019).

Research conducted shows that this does not align with pecking order theory. Based on this research, small companies can achieve high profitability, yet their DER values also show high values.

The Effect of Liquidity on Capital Structure with Company Size as a Moderating Variable

The results of equation 2 in the regression moderation test were 0.332 (>0.05). This value is greater than the alpha used, thus rejecting the sixth hypothesis. Company size does not weaken or strengthen liquidity through capital structure. These results support research conducted by (Raihani & Jati, 2022). This research shows that it is not in accordance with the pecking order theory which

causes differences in capital structure in companies.

The Effect of Non-Debt Tax Shield on Capital Structure with Company Size as a Moderating Variable.

In the regression moderation test of equation 2, the resulting value is 0.493 (>0.05), so the seventh hypothesis is rejected where company size does not strengthen or weaken the relationship between the non-debt tax shield and capital structure.

D. CONCLUSION

1. Profitability significantly contributes to determining capital structure.
2. Liquidity has no influence on firm structure.
3. The non-debt tax shield influences the determinants of a firm's capital structure.
4. Profitability, liquidity, and the non-debt tax shield simultaneously contribute to capital structure. A company with optimal profitability, liquidity, and the non-debt tax shield will promote an optimal capital structure.
5. Profitability influences capital structure, but the relationship between profitability and capital structure is not moderated by firm size.
6. Liquidity has no influence on capital structure, and firm size also does not influence the effect of liquidity on capital structure.
7. The non-debt tax shield has an influence on capital structure, but firm size as a moderating variable is unable to moderate the relationship between the non-debt tax shield and capital structure.

Limitation

1. The adjusted R-Square value is 28.6%, with 71.4% remaining unexplained variables in this study.
2. During the sample selection stage, there were constraints that resulted in a reduction in the sample size.

Suggestion

1. The researcher further expanded the period to 7 years to capture the five-year recovery period after the COVID-19 pandemic in order to provide a broadly representative sample. The researcher further added other variables not included in this study, such as asset tangibility, risk management, and sales growth.
2. Companies in the property and real estate sector need to utilize their debt and fixed assets as non-debt tax shields.
3. The government needs to support companies in the property and real estate sector in IPOs and strengthen the bond market to support the development of property in Indonesia.

REFERENCES

1. Abubakar, D. Y., Daat, S. C., & Sanggenafa, M. A. (2020). Pengaruh Tangibility, Profitabilitas, Growth, Risiko Bisnis Dan Likuiditas Terhadap Kebijakan Hutang (Studi pada Perusahaan Manufaktur yang Terdaftar di Bursa Efek Indonesia Periode 2015- 2017). Jurnal Akuntansi Dan Keuangan Daerah, 15(1), 113–128. <https://doi.org/10.52062/jakd.v15i1.1470>

2. Adeneye, Y. B., Kammoun, I., & Ab Wahab, S. N. A. (2023). Capital structure and speed of adjustment: the impact of environmental, social and governance (ESG) performance. *Sustainability Accounting, Management and Policy Journal*, 14(5), 945–977. <https://doi.org/10.1108/SAMPJ-01-2022-0060>
3. Ahmed, I. E., & Sabah, A. (2021). The determinants of capital structure of the GCC oil and gas companies. *International Journal of Energy Economics and Policy*, 11(2), 30–39. <https://doi.org/10.32479/ijeep.10570>
4. Ben Hamouda, R. Z., Hamzaoui, N., & Jilani, F. (2023). Capital Structure Determinants: New Evidence from the MENA Region Countries. *International Journal of Economics and Financial Issues*, 13(1), 144–163. <https://doi.org/10.32479/ijefi.13695>
4. Cahya sinta sari, Dirvi Surya Abbas, & Seleman Hardi Yahawi. (2023). Pengaruh Likuiditas, Tangibility dan Non Debt Tax Shield Terhadap Struktur Modal. *MUQADDIMAH: Jurnal Ekonomi, Manajemen, Akuntansi Dan Bisnis*, 2(1), 35–44. <https://doi.org/10.59246/muqaddimah.v2i1.572>
5. Cahyani, N., & Nyale, M. H. Y. (2022). Pengaruh Struktur Aset dan Likuiditas terhadap Struktur Modal dengan Ukuran Perusahaan sebagai Variabel Moderasi. *JiIP - Jurnal Ilmiah Ilmu Pendidikan*, 5(7), 2675–2686. <https://doi.org/10.54371/jiip.v5i7.748>
6. Chandra, T., Chandra, S., Wijaya, E., Chandra, J., & Ng, M. (2020). Co-determination of capital structure and profitability: An empirical test of Indonesia stock exchange. *Industrial Engineering and Management Systems*, 19(3), 561–575. <https://doi.org/10.7232/iems.2020.19.3.561>
7. Damina, E. K., Muritala, T., & Ibrahim, A. U. (2022). Effect of Speed of Adjustments on Capital Structure Decision: A Conceptual Analysis. *International Journal of Economics and Financial Issues*, 12(6), 15–21. <https://doi.org/10.32479/ijefi.13469>
8. Evinemi, E. A., & Emengini, S. E. (2024). The Evaluation of Liquidity and CapitalStructure of Manufacturing Firms in Nigeria. *International Journal of Economics and Financial Issues*, 14(3), 212–218. <https://doi.org/10.32479/ijefi.15946>
9. Fattimatuz, Z., Malavia, R., & Budi, W. (2017). Pengaruh Profitabilitas, Likuiditas, dan Growth Opportunity Terhadap Struktur Modal (Pada Perusahaan Food and Beverage). 218–230. www.fe.unisma.ac.id
10. Fitri Lim, Hesty Ervianni Zulaecha, Hamdani, H., & Sriyanto, S. (2022). Pengaruh Profitabilitas, Likuiditas, Struktur Aktiva, Dan Ukuran Perusahaan Terhadap Struktur Modal. *Akuntansi*, 1(4), 28–38. <https://doi.org/10.55606/jurnalrisetilmuakuntansi.v1i4.109>
11. Ghani, E. K., Adnan Hye, Q. M., Rehan, R., & Salahuddin, S. (2023). Examining Capital Structure Determinants for ASEAN Energy Firms. *International Journal of Energy Economics and Policy*, 13(3), 129–140. <https://doi.org/10.32479/ijeep.14070>
12. Irmawati, S., & Murtianingsih. (2024). Analisis Determinan Faktor Nilai Perusahaan JIMEA | *Jurnal Ilmiah MEA (Manajemen , Ekonomi , dan Akuntansi)*. 8(2), 1553–1573.
13. Kompas.com. (2024). Ini Emiten Properti yang Merugi hingga Semester Pertama 2024. <https://www.kompas.com/properti/read/2024/09/21/160000621/ini-emiten-properti-yang-merugi-hingga-semester-pertama-2024>
14. Lilia, W. (2020). 519953-None-61F197F2. 4, 627–639.
15. Martini, M. (2024). Struktur Modal Perusahaan Properti Dan Real Estate Di Bei Periode 2018-2022. *Sebatik*, 28(1), 103–111. <https://doi.org/10.46984/sebatik.v28i1.2441>

16. Meilani, U., & Wahyudin, A. (2021). Pengaruh Struktur Aset, Risiko Bisnis, dan Pertumbuhan Penjualan Terhadap Struktur Modal dengan Profitabilitas sebagai Variabel Moderasi. *Jurnal Akuntansi Bisnis* Vol. 19 No. 1.
17. Miswanto, M., Setiawan, A. Y., & Santoso, A. (2022). Analisis Pengaruh Pertumbuhan Penjualan, Struktur Aset, dan Profitabilitas terhadap Struktur Modal. *Jurnal Maksipreneur: Manajemen, Koperasi, Dan Entrepreneurship*, 11(2), 212. <https://doi.org/10.30588/jmp.v11i2.945>
18. MYERS, S. C. (1984). The Capital Structure Puzzle. *The Journal of Finance*, 39(3), 574–592. <https://doi.org/10.1111/j.1540-6261.1984.tb03646.x>
19. Puspitasari, W. A. (2022). Pengaruh likuiditas, struktur aktiva, dan profitabilitas terhadap struktur modal. *Jurnal Cendekia Keuangan*, 1(1), 42. <https://doi.org/10.32503/jck.v1i1.2258>
20. Putra Erisa, R. F., & Deliza Henny. (2023). Pengaruh Growth of Sales, Profitabilitas, Dan Likuiditas Terhadap Struktur Modal Dengan Ukuran Perusahaan Sebagai Variabel Moderasi. *Jurnal Ekonomi Trisakti*, 3(1), 1319– 1328. <https://doi.org/10.25105/jet.v3i1.16136>
21. Raihani, R., & Jati, A. K. N. (2022). Determinan struktur modal dengan firm size sebagai variabel moderasi. *Journal of Business and Banking*, 12(1), 109. <https://doi.org/10.14414/jbb.v12i1.2975>
22. Rákos, M., Szenderák, J., Erdey, L., Kőmíves, P. M., & Fenyves, V. (2022). Analysis of the Economic Situation of Energy Companies in Central and Eastern Europe. *International Journal of Energy Economics and Policy*, 12(4), 553–562. <https://doi.org/10.32479/ijeep.12957>
23. Rehan, R., Khan, M. A., Fu, G. H., Sa’ad, A. A., & Irshad, A. (2024). The Determinants of Shariah Banks’ Capital Structure. *International Journal of Economics and Financial Issues*, 14(5), 193–202. <https://doi.org/10.32479/ijefi.16923>
24. Rifiana, A. S., Febiyanti, E., & Hersugondo, H. (2021). Pengaruh Ukuran Perusahaan, Tangibility, Likuiditas dan Profitabilitas terhadap Struktur Modal pada Perusahaan Industri Pertambangan Batu Bara. *AKSES: Jurnal Ekonomi Dan Bisnis*, 16(2), 178–192. <https://doi.org/10.31942/akses.v16i2.5560>
25. Sekaran, U. (2020). *Research Methods for Business: a Skill Building Approach* 8th Edition. Jhon Wiley and Sons, Inc.
- Simatupang, H. J., Purwanti, L., & Mardiaty, E. (2019). Determinants of capital structures based on the Pecking Order Theory and Trade-off Theory. *Jurnal Keuangan Dan Perbankan*, 23(1). <https://doi.org/10.26905/jkdp.v23i1.2579>
- Sujarweni, W. (2018). *Metodologi Penelitian Bisnis dan Ekonomi Pendekatan Kuantitatif*. PUSTAKABARUPRESS.
- Sutomo, S., Wahyudi, S., Pangestuti, I. R. D., & Muharam, H. (2020). The determinants of capital structure in coal mining industry on the Indonesia Stock Exchange. *Investment Management and Financial Innovations*, 17(1), 165–174. [https://doi.org/10.21511/imfi.17\(1\).2020.15](https://doi.org/10.21511/imfi.17(1).2020.15)
26. Terdaftar, Y., Bei, D. I., & Geraldine, C. (2022). Pengaruh profitabilitas, likuiditas,. 9(2).
- Vinda Izmi, A. (2019). Pengaruh Profitabilitas dan Likuiditas Terhadap Struktur Modal dengan Ukuran Perusahaan sebagai Variabel Moderasi.
27. Wardoyo, D. U., & Andani, A. D. (2024). Analisis Faktor-Faktor Yang Mempengaruhi Struktur Modal : Studi Kasus Pada Sektor Industri Asuransi. *Jurnal Reviu Akuntansi Dan Keuangan*, 14(2), 477–491. <https://doi.org/10.22219/jrak.v14i2.33497>
28. Weygandt, J. J., Kimmel, P. D., & Kieso, D. E. (2019). *Financial Accounting with International Financial Reporting Standards* 4th Edition. Jhon Wiley and Sons, Inc.
29. Wijandari, A. (2020). Determinan Terhadap Struktur Modal Dengan Non Debt Tax Shield

Sebagai Moderasi (Studi Pada Perusahaan Properti, Real Estate dan Konstruksi Bangunan Yang Terdaftar Di BEI Periode 2015-2018). Tirtayasa Ekonomika, 15(1), 175.
<https://doi.org/10.35448/jte.v15i1.7463>

30. Wulandari, N. P. I., & Artini, L. G. S. (2019). Pengaruh Likuiditas, Non-Debt Tax Shield, Ukuran Perusahaan Dan Pertumbuhan Penjualan Terhadap Struktur Modal. E-Jurnal Manajemen Universitas Udayana, 8(6), 3560.
<https://doi.org/10.24843/ejmunud.2019.v08.i06.p10>

31. Zandi, G., Rehan, R., Hye, Q. M. A., & Choo, S. L. B. (2023). Exploring Capital Structure Determinants for OECD Energy Firms. International Journal of Energy Economics and Policy, 13(4), 338–347. <https://doi.org/10.32479/ijeep.14559>

Unveiling Decision-Making Patterns in Personal Finance: A Decade of Bibliometric Insights (2015–2025)

Abdul Rahman, Vania Rabiola Putri, Syifa Hanifa Salsabil, Cynthia Sari Dewi
Department of Management & Jaya Launch Pad, Universitas Pembangunan Jaya, Indonesia

Abstract

This study explores decision-making strategies in personal finance through a bibliometric analysis of 108 peer-reviewed articles published between 2015 and 2025 and indexed in Scopus. As individuals face increasingly complex financial choices, factors such as psychological traits, digital tools, and socioeconomic context have gained attention alongside traditional financial literacy. Using VOSviewer and Biblioshiny, this research maps key themes, trends, and collaborative networks across a decade of scholarship. Findings reveal dominant themes including financial literacy, behavioral finance, digital decision tools, and educational interventions. Particular focus is placed on young adults, students, healthcare professionals, and financially vulnerable populations. While the field has grown, gaps remain in areas like artificial intelligence applications, crisis-related financial behavior, and long-term planning (e.g., pensions and household income). The study underscores the interdisciplinary nature of personal finance research, which intersects economics, psychology, education, and technology. It also highlights the dominance of the United States and Australia in publication volume and citation impact. By providing a structured overview of the field, this study supports future academic inquiry and practical development of more inclusive and adaptive financial strategies.

Keywords: Personal Finance, Decision Making Strategies; Bibliometric Analysis; VOS Viewer; Biblioshiny

1. Introduction

In today's rapidly evolving financial environment, individuals face increasingly complex decisions—from budgeting and debt repayment to long-term investment and retirement planning. While financial literacy has long been regarded as a key determinant of sound financial behavior, emerging studies reveal that knowledge alone is insufficient. Psychological traits, digital competencies, socioeconomic contexts, and even emotional resilience significantly shape how individuals make personal financial decisions. For instance, Sewell & Rogers (2023) emphasized that integrating financial education with real-life activities enhances confidence and competence, particularly among youth. Smith et al. (2023) highlighted the persistent gap between financial knowledge and decision confidence, even among well-educated populations like pharmacy students. Other works have introduced rule-based and simulation tools Althnian (2021) to assist users in navigating financial uncertainty, especially among millennials and digitally engaged demographics.

Recent research also draws attention to financially vulnerable populations—such as low-income groups, justice-involved youth, and those lacking access to formal financial institutions. Ethridge et al. (2024) stressed the importance of life-skills-based financial training, while French et al. (2021) explored the growing role of digital platforms in offering decision support to underserved users. Despite this growing body of literature, a systematic and integrative overview of decision-making strategies in personal finance remains scarce. Most studies are fragmented across disciplines, populations, and methodologies—resulting in a lack of conceptual coherence and theoretical convergence. More importantly, few efforts have employed bibliometric mapping to uncover dominant themes, research clusters, and intellectual structures in this domain.

This study aims to fill that gap. Using bibliometric analysis of 108 peer-reviewed journal articles published between 2015 and 2025 and indexed in the Scopus database, we

systematically explore the structure, evolution, and thematic focus of scholarly work on decision-making strategies in personal finance. Leveraging tools such as VOSviewer and Biblioshiny, this paper seeks to Identify key research themes and theoretical patterns; Map the

intellectual and collaborative networks within the field; Detect emerging trends and underexplored areas to guide future research and practice. By offering a structured knowledge map, this study contributes to both academic inquiry and practical policy formulation in the realm of personal financial behavior.

2. Literature Review

The literature on personal finance decision-making has expanded beyond the traditional emphasis on financial literacy to encompass psychological, technological, and structural dimensions that influence behavior. Several studies underscore the influence of cognitive biases, emotional regulation, and behavioral tendencies on individuals' financial decisions. Baulkaran (2022) pointed out that even financially literate individuals often struggle to apply their knowledge due to systemic and psychological constraints. Riitsalu (2019) highlighted the frequent mismatch between perceived and actual financial competence, particularly among young people. Emotional fatigue and the absence of social reinforcement, according to Igu et al. (2022), diminish the sustainability of financial behavior change. Asebedo et al. (2021) further linked financial self-efficacy to long-term planning capacities.

To address the limitations of human judgment, structured decision support systems have gained traction. Ramme et al. (2021) reported increased saving consistency among healthcare workers when guided by formal prompts, while Root et al. (2017) found structured frameworks reduced impulsive spending among special-needs populations. Simulation tools, such as those developed by Kaneda et al. (2021) and Mastroeni et al. (2023), help reinforce consistent decision-making through real-time feedback and pattern repetition.

Educational interventions have shifted from content delivery to the development of problem-solving and emotional strategies. van Deventer (2020) found that problem-solving modules reduced impulsivity in money management. Ethridge et al. (2024), Woolley et al. (2024), and Didkivskyi et al. (2024) demonstrated that life-skills training, emotional regulation, and scenario-based learning significantly improve planning behavior, particularly among vulnerable youth.

The digitalization of finance has further shaped strategy development. Chin et al. (2022), French et al. (2021), and Nguyen et al. (2024) demonstrated the impact of interface design, digital nudges, and algorithmic coaching on decision outcomes. Meanwhile, Berman et al. (2016) and Maman & Rosenhek (2020, 2023) revealed how institutional barriers and digital exclusions affect financially vulnerable groups—showing that inclusive strategies must consider both psychological and systemic factors.

3. Methodology

This study employs a bibliometric analysis to systematically examine the structure and thematic evolution of research on decision strategies in personal finance. Bibliometric approaches allow researchers to trace scholarly trends by analyzing metadata such as authorship, keywords, citations, and affiliations (Ellegaard & Wallin, 2015). This method offers a replicable and objective overview of a research field's development, making it

valuable for identifying research gaps, influential works, and emerging themes (Moral- Muñoz et al., 2020). Data were sourced from Scopus (<https://www.scopus.com>), selected for its broad coverage of high-quality peer-reviewed journals in business, economics, and social sciences (Martín-Martín et al., 2018). The search query TITLE("personal finance" OR "personal finance management") yielded 312 results (1962–2025), filtered to English-language articles from 2015–2025. After removing non-article types, the final dataset comprised 108 articles.

Metadata was exported in BibTeX format. Analysis was conducted using VOSviewer and Biblioshiny. VOSviewer enabled co-authorship and keyword co- occurrence mapping (van Eck & Waltman, 2010), while Biblioshiny, part of the Bibliometrix R-package, provided strategic mapping and keyword clustering into motor, basic, niche, and emerging themes (Cobo et al., 2011), offering a multidimensional view of the conceptual landscape.

4. Result

4.1 Trends in Scholarly Publications

Understanding the evolution of research in personal finance over time provides valuable insight into the field’s intellectual development and scholarly attention. Tracking publication trends, disciplinary affiliations, and author influence enables a clearer view of how the field has grown, which themes have gained momentum, and which areas remain underexplored. This figure 1 illustrates scholarly output from 2015 to 2025, focusing on publication frequency, interdisciplinary engagement, and citation dynamics.

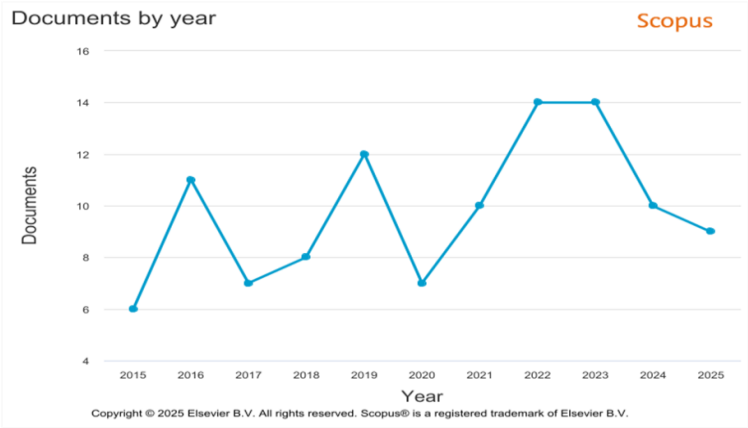


Figure 1. Publication Trend per Year

The publication trend for personal finance research, based on Scopus data from 2015 to 2025, displays noticeable fluctuations. In 2015, six relevant articles were published. This number almost doubled in 2016 with 11 publications, suggesting a spike in interest. However, 2017 witnessed a decline to 7 publications, possibly reflecting shifting funding priorities or evolving research themes. The upward trajectory resumed in 2018 (9 documents) and peaked again in 2019 with 12 publications. A drop occurred in 2020 (7 documents), potentially reflecting the early effects of the COVID-19 pandemic on research productivity. A significant rebound began in 2021 (10 documents) and culminated in sustained peaks of 14 documents in both 2022 and 2023—the highest output in the dataset. This surge may reflect increased academic attention toward financial literacy, digital finance, and behavioral finance driven by pandemic-related economic uncertainty. Publication numbers declined again in 2024 (10

documents) and 2025 (9 documents), but the latter should be cautiously interpreted due to potential mid-year data limitations. Overall, the data suggest episodic interest influenced by external events, funding availability, and emergent trends such as fintech innovation and financial resilience.

4.2 Interdisciplinary Structure of Personal Finance Research

Personal finance research is inherently interdisciplinary, as illustrated in figure 2. While traditionally rooted in Economics, Econometrics, and Finance (23.8%), it also shows substantial representation from Social Sciences (22.8%) and Business, Management, and Accounting (14.8%). Together, these fields represent nearly half of all research contributions. The overlap highlights the necessity of combining economic theory with behavioral and managerial insights to understand individual financial behavior. Additionally, the field engages deeply with disciplines beyond its core. Medical and Health Professions account for 11.1% combined, reflecting concerns about the financial well-being of healthcare professionals and patients, particularly under financial stress. Psychology contributes 7.9%, emphasizing behavioral finance themes like cognitive biases, financial anxiety, and decision-making under uncertainty. Computer Science's 6.3% share underscores the growing role of fintech, including algorithmic budgeting tools, mobile applications, robo-advisors, and financial simulation technologies. Smaller but notable contributions come from Arts and Humanities (2.6%), Decision Sciences (2.1%), and Engineering (2.1%), suggesting emerging interests in ethical finance, decision modeling, and system optimization. This wide-ranging involvement confirms that understanding personal financial behavior requires a systems approach, integrating insights from economics, psychology, technology, health, and social policy.

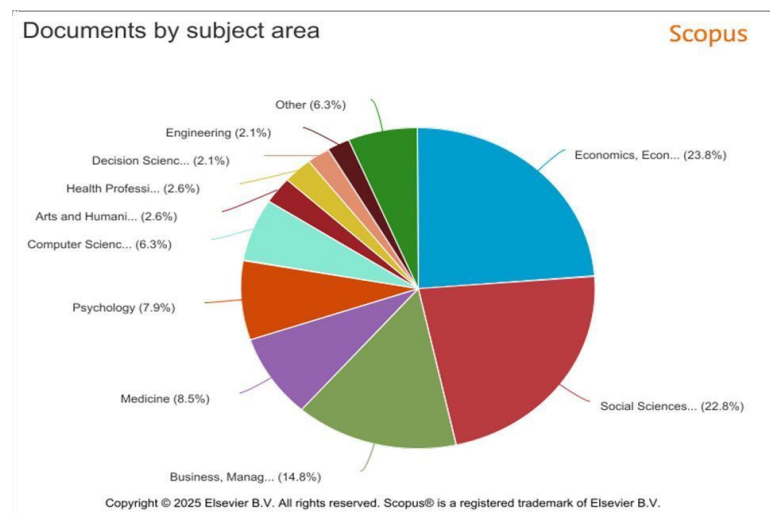


Figure 2 Subject Area

4.3 Thematic Networks and Core Research Topics

The co-occurrence of keywords in personal finance research reveals several distinct thematic clusters based on how frequently terms appear together in publications. Keywords are grouped by theme, with size indicating how often they occur and connections representing shared appearances in the same studies. These groupings highlight the main focus areas within the field and how they relate to each other.

Based on figure 3, one major cluster (yellow) centers on core personal finance topics and demographic groups. It includes terms like “personal finance,” “young adult,” “female,” “male,” “debt,” “investment,” and “self concept,” reflecting research that examines financial behavior and conditions across different population segments. Another prominent cluster (green) is focused on financial literacy and planning. Key terms include “financial literacy,” “financial self-efficacy,” “financial planning,” “stocks,” and “artificial intelligence.” This suggests attention to financial knowledge, confidence, and decision-making, as well as emerging interest in the role of technology in financial planning. The red cluster relates to financial education and behavior, with central terms such as “education,” “students, pharmacy,”

“behavior,” “learning,” “curriculum,” and “retirement.” This points to studies examining how financial education influences behavior, especially among students, and how it prepares individuals for later stages of life.

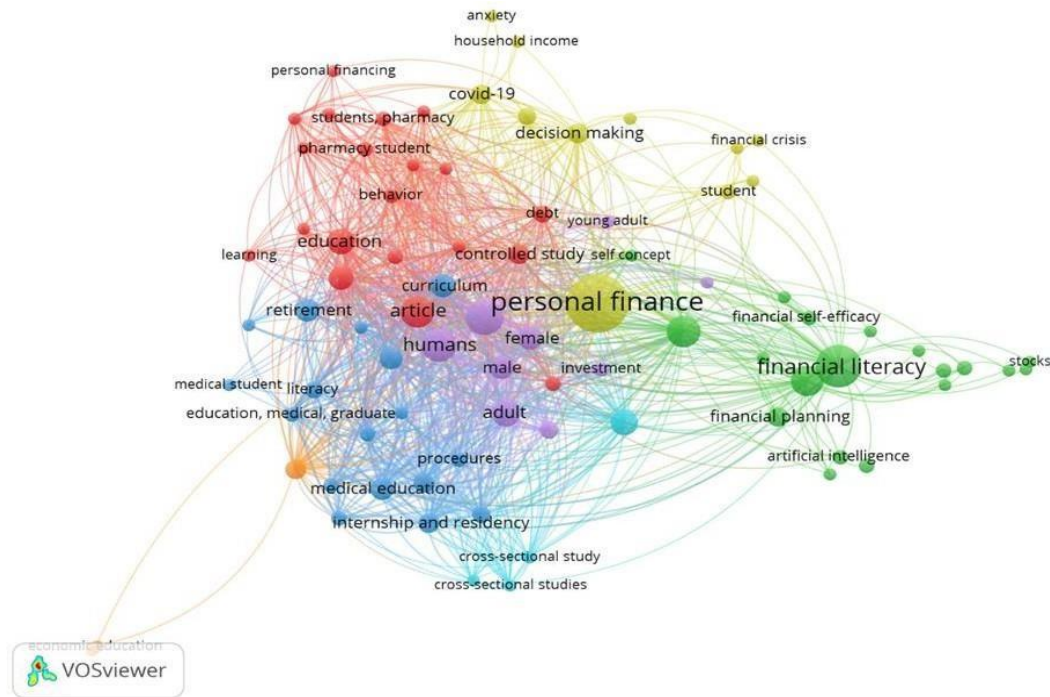


Figure 3. Keyword co-occurrence network

A separate group (blue) focuses on adult learners and general literacy, including terms like “humans,” “adult,” “literacy,” and “medical student.” These studies often explore broader literacy or educational themes, sometimes in healthcare contexts. The orange cluster is more specialized, dealing with financial aspects of medical education. Terms like “medical education,” “internship and residency,” and “procedures” suggest a niche area of research related to the economic challenges in medical training. Lastly, a light yellow or beige cluster emphasizes financial decision-making in context. It includes “decision making,” “covid-19,” “financial crisis,” “anxiety,” “household income,” and “student,” highlighting how external events and psychological factors affect financial choices. Altogether, the findings point to several key areas of personal finance research, with strong emphasis on financial literacy, education, behavior, decision-making, and context-specific challenges faced by different demographic and professional groups.

4.4. Potential research gap for future research

Based on the interpretation of the keyword co-occurrence and thematic mapping, several important gaps and opportunities emerge that warrant further exploration. The following areas are recommended as promising directions for future research in the field of personal finance:

1. Artificial Intelligence in Financial Education

Although “AI” appears as a keyword, its presence in a niche theme suggests that the intersection between AI and personal finance is still underdeveloped. Future studies could explore AI tools for teaching financial literacy, personalized financial advising, and predictive modeling for financial behavior

Financial Behavior During Crises

Keywords like “covid-19,” “pandemic,” and “financial crisis” show some engagement with crisis-based financial decision-making, but these topics appear less mature. Research could investigate how individuals and households adapt financially during various types of

crises, including health, economic, or environmental.

4. Expanding Financial Education to Diverse Professions

The focus on medical and pharmacy students is strong, but other professional groups (e.g., engineers, teachers, creative workers) are underrepresented. Broadening the scope to these fields can help make financial education more inclusive and context-specific.

5. Pensions and Long-Term Financial Planning

The keyword “pension” is isolated and underdeveloped, despite being crucial to financial well-being. There is an opportunity to deepen research on retirement planning, especially among younger populations or informal sector workers.

6. Household Income and Socioeconomic Factors

While “household income” and “income” are present, they are not central to the thematic structure. Further investigation could examine how income variability, job stability, and socioeconomic status shape financial literacy and planning behavior.

5. Discussion

This bibliometric review highlights the growing complexity of personal finance research, reflecting a shift from simple financial literacy to more nuanced, behavioral, and system-based approaches. For instance, women with higher financial self-efficacy are more likely to engage in forward-looking financial behaviors such as saving and investing, and less likely to rely on credit—indicating a strong link between psychological confidence and financial responsibility (Farrell et al., 2016). Social background also plays a critical role. A study found that students’ financial knowledge is significantly influenced by socialization patterns, migration background, and disposable income (Förster et al., 2019). This insight is important for targeted interventions, particularly for young adults from underserved backgrounds.

The need for financial education is also echoed in medical training, where improving financial literacy is linked to better stress management and job performance among physicians (Ahmad et al., 2017). Technological responses have emerged as well. A rule-based personal finance management system was designed to promote financial well-being by

offering users personalized saving and spending advice, reflecting a shift toward automation in financial decision-making (Althnian, 2021). Generational perspectives add another dimension. A study found that education level, age, and co-curricular exposure are key drivers of youth engagement with personal finance. The findings suggest a growing emphasis on integrating personal finance into school systems without sacrificing other subjects (Beck & Garriss, 2019)(Babu & G, 2024).

Psychological research indicates that procrastination in financial behavior is not due to a lack of planning, but rather to vague or poorly structured intentions. More specific and time-bound planning strategies could help close this intention-action gap, particularly among those with lower conscientiousness (Gamst-Klaussen et al., 2019). Together, these studies underscore how personal finance decisions are shaped by a mix of psychological traits, social structures, education, and technological support—suggesting a rich, interdisciplinary foundation for future research.

6. Conclusion

This bibliometric study systematically mapped the landscape of research on decision-making strategies in personal finance published between 2015 and 2025. By analyzing 108 articles from the Scopus database using VOSviewer and Biblioshiny, we unveiled the structural dynamics, thematic evolution, and intellectual core of this rapidly developing field. The analysis confirms

that personal finance research is not confined to traditional economic or financial paradigms but thrives as a highly interdisciplinary domain, drawing significant contributions from social sciences, psychology, health sciences, education, and computer science. Our findings highlight several key characteristics of the field over the past decade. Research output has shown considerable dynamism, with notable peaks suggesting responsiveness to societal events and evolving academic interests, although recent trends are subject to citation and indexing lags. The USA and Australia emerge as dominant contributors in terms of both publication volume and citation impact. Thematically, "personal finance" and "financial literacy" stand as central, well-developed motor themes, strongly interconnected with research on education, behavior, demographic factors, and planning. Distinct clusters focused on medical education finance and the contextual/psychological drivers of decision-making further illustrate the field's specialized niches and breadth.

Significantly, this study identifies critical areas for future research. While technology, particularly AI, is emerging, its integration into financial education and decision support remains underdeveloped. Similarly, research on financial behavior during crises, long-term planning horizons (including pensions and household income management), and the extension of financial education initiatives beyond specific student populations (like health professions) represent substantial opportunities for scholarly advancement. Addressing these gaps is essential for developing more comprehensive theories and effective, inclusive interventions. This research provides a structured overview of a complex and fragmented field, offering a valuable resource for researchers, educators, and policymakers. It underscores the need for continued interdisciplinary collaboration and highlights promising avenues for future inquiry aimed at enhancing understanding and improving personal financial decision-making and well-being in an increasingly intricate financial world.

References

1. Ahmad, F. A., White, A. J., Hiller, K. M., Amini, R., & Jeffe, D. B. (2017). An assessment of residents' and fellows' personal finance literacy: an unmet medical education need. *International Journal of Medical Education*, 8, 192–204. <https://doi.org/10.5116/ijme.5918.ad11>
2. Althnian, A. (2021). Design of a Rule-based Personal Finance Management System based on Financial Well-being. *International Journal of Advanced Computer Science and Applications*, 12(1). <https://doi.org/10.14569/IJACSA.2021.0120122>
3. Asebedo, S. D., Seay, M. C., Little, T. D., Enete, S., & Gray, B. (2021). Three good things or three good financial things? Applying a positive psychology intervention to the personal finance domain. *The Journal of Positive Psychology*, 16(4), 481–491. <https://doi.org/10.1080/17439760.2020.1752779>
4. Babu, J., & G, V. (2024). Tri-Decadal Review on Personal Finance and Financial Well- Being. *International Research Journal of Multidisciplinary Scope*, 05(03), 1021–1042. <https://doi.org/10.47857/irjms.2024.v05i03.0985>
5. Baulkaran, V. (2022). Personal bankruptcy and consumer credit delinquency: The case of personal finance education. *International Review of Financial Analysis*, 81, 102098. <https://doi.org/10.1016/j.irfa.2022.102098>
6. Beck, J. J., & Garris, R. O. (2019). Managing Personal Finance Literacy in the United States: A Case Study. *Education Sciences*, 9(2), 129. <https://doi.org/10.3390/educsci9020129>
7. Berman, J. Z., Tran, A. T. K., Lynch, J. G., & Zauberman, G. (2016). Expense Neglect in Forecasting Personal Finances. *Journal of Marketing Research*, 53(4), 535–550. <https://doi.org/10.1509/jmr.15.0101>
8. Chin, A., Zimmerman, D., Johnson, H., & Shu, S. B. (2022). Decisions about

overdraft coverage: Disclosure design and personal finances. *Journal of Experimental Psychology: Applied*, 28(4), 746–774. <https://doi.org/10.1037/xap0000460>

9. Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field. *Journal of Informetrics*, 5(1), 146–166. <https://doi.org/10.1016/j.joi.2010.10.002>

10. Didkivskyi, V., Antoniuk, D., Vakaliuk, T., & Ohinskyi, Y. (2024). Using the proximal policy optimization and prospect theory to train a decision-making model for managing personal finances. *Radioelectronic and Computer Systems*, 2024(4), 248–258. <https://doi.org/10.32620/reks.2024.4.20>

11. Ellegaard, O., & Wallin, J. A. (2015). The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics*, 105(3), 1809–1831. <https://doi.org/10.1007/s11192-015-1645-z>

12. Ethridge, A. N., Kemp, J. M., Edgemon, A. K., & Richling, S. M. (2024). Teaching basic personal finance to justice-involved youth. *Behavioral Interventions*, 39(4). <https://doi.org/10.1002/bin.2035>

13. Farrell, L., Fry, T. R. L., & Risse, L. (2016). The significance of financial self-efficacy in explaining women's personal finance behaviour. *Journal of Economic Psychology*, 54, 85–99. <https://doi.org/10.1016/j.joep.2015.07.001>

14. Förster, M., Happ, R., & Walstad, W. B. (2019). Relations between young adults' knowledge and understanding, experiences, and information behavior in personal finance matters. *Empirical Research in Vocational Education and Training*, 11(1), 2. <https://doi.org/10.1186/s40461-019-0077-z>

15. French, D., McKillop, D., & Stewart, E. (2021). Personal finance apps and low-income households. *Strategic Change*, 30(4), 367–375. <https://doi.org/10.1002/jsc.2430>

16. Gamst-Klaussen, T., Steel, P., & Svartdal, F. (2019). Procrastination and Personal Finances: Exploring the Roles of Planning and Financial Self-Efficacy. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00775>

17. Igu, J. A., Zakaria, S., & Bar-Or, Y. D. (2022). Systematic review of personal finance training for physicians and a proposed curriculum. *BMJ Open*, 12(12), e064733. <https://doi.org/10.1136/bmjopen-2022-064733>

18. Kaneda, M., Kubota, S., & Tanaka, S. (2021). Who spent their COVID-19 stimulus payment? Evidence from personal finance software in Japan. *The Japanese Economic Review*, 72(3), 409–437. <https://doi.org/10.1007/s42973-021-00080-0>

19. Maman, D., & Rosenhek, Z. (2020). Facing future uncertainties and risks through personal finance: conventions in financial education. *Journal of Cultural Economy*, 13(3), 303–317. <https://doi.org/10.1080/17530350.2019.1574865>

20. Maman, D., & Rosenhek, Z. (2023). Governing individuals' imaginaries and conduct in personal finance: The mobilization of emotions in financial education. *Journal of Consumer Culture*, 23(1), 188–208. <https://doi.org/10.1177/14695405211069952>

21. Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, 12(4), 1160–1177. <https://doi.org/10.1016/j.joi.2018.09.002>

22. Mastroeni, L., Naldi, M., & Vellucci, P. (2023). Personal Finance Decisions with Untruthful Advisors: An Agent-Based Model. *Computational Economics*, 61(4), 1477–1522. <https://doi.org/10.1007/s10614-022-10256-4>

23. Moral-Muñoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A., & Cobo, M. J. (2020). Software tools for conducting bibliometric analysis in science: An up-to-date review. *El Profesional de La Información*, 29(1). <https://doi.org/10.3145/epi.2020.ene.03>

24. Nguyen, M. H., Sun Hosoya, L., & Guyon, I. (2024). Meta-learning from learning curves for budget-limited algorithm selection. *Pattern Recognition Letters*, 185, 225–231. <https://doi.org/10.1016/j.patrec.2024.08.010>
25. Olafsson, A., & Pagel, M. (2018). The Liquid Hand-to-Mouth: Evidence from Personal Finance Management Software. *The Review of Financial Studies*, 31(11), 4398–4446. <https://doi.org/10.1093/rfs/hhy055>
26. Pafmj. (2024). author index. *Pakistan Armed Forces Medical Journal*, 74(6). <https://doi.org/10.51253/pafmj.v74i6.13071>
27. Ramme, A. J., Patel, M., Patel, K. A., Montag, W. H., Schau, A. J., Sabo, S. I., & Bedi, A. (2021). Personal Finance Primer for the Future Orthopaedic Surgeon: A Starting Point. *JBJS Open Access*, 6(1). <https://doi.org/10.2106/JBJS.OA.20.00006>
28. Riitsalu, L. (2019). Disentangling financial literacy: three-dimensional approach to analysing management of personal finances in Estonia. In *Article in Journal of Social Research and Policy*. <https://www.researchgate.net/publication/332267196>
29. Root, J., Saunders, A., Spooner, F., & Brosh, C. (2017). Teaching Personal Finance Mathematical Problem Solving to Individuals With Moderate Intellectual Disability. *Career Development and Transition for Exceptional Individuals*, 40(1), 5–14. <https://doi.org/10.1177/2165143416681288>
30. Sewell, J., & Rogers, S. (2023). Assessing the Impact of a Personal Finance Elective Course on Student Attitudes and Intentions. *American Journal of Pharmaceutical Education*, 87(2), ajpe8942. <https://doi.org/10.5688/ajpe8942>
31. Shappell, E., Ahn, J., Ahmed, N., Harris, I., Park, Y. S., & Tekian, A. (2018). Personal Finance Education for Residents: A Qualitative Study of Resident Perspectives. *AEM Education and Training*, 2(3), 195–203. <https://doi.org/10.1002/aet2.10090>
32. Smith, S. E., Smith, M. D., & Bland, C. M. (2023). Perceptions and knowledge of personal finance in doctor of pharmacy versus undergraduate students: More wisdom or just more debt? *Currents in Pharmacy Teaching and Learning*, 15(5), 473–482. <https://doi.org/10.1016/j.cptl.2023.04.020>
33. Soulier, L., Jabeur, L. Ben, Tamine, L., & Bahsoun, W. (2013). On ranking relevant entities in heterogeneous networks using a language-based model. *Journal of the American Society for Information Science and Technology*, 64(3), 500–515. <https://doi.org/10.1002/asi.22762>
34. van Deventer, M. (2020). African Generation Y students' personal finance behavior and knowledge. *Investment Management and Financial Innovations*, 17(4), 136–144. [https://doi.org/10.21511/imfi.17\(4\).2020.13](https://doi.org/10.21511/imfi.17(4).2020.13)
35. van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
36. White, M. D. (2017). Nudging Debt: On the Ethics of Behavioral Paternalism in Personal Finance. *Journal of Financial Counseling and Planning*, 28(2), 225–234. <https://doi.org/10.1891/1052-3073.28.2.225>
37. Woolley, P., Kendall, M., Tacvorian, S., & Asher, S. (2024). Integrating a Personal Finance Workshop to Enhance Financial Literacy Among Senior Medical Students: A Single Institution's Experience. *Advances in Medical Education and Practice*, Volume 15, 885–891. <https://doi.org/10.2147/AMEP.S474002>

THE EFFECT OF JOB SEARCH SELF-EFFICACY ON CAREER ADAPTABILITY OF INDONESIAN FRESH GRADUATES

Adinda Anggun Pramesti

Veronica Anastasia Melany Kaihatu (corresponding author)

Universitas Pembangunan Jaya

Abstract

New graduates from higher education often experienced difficulties in finding and securing a job due to the rapidly changing work environments. They need career adaptability to adjust to the changing labor market. One factor that can affect career adaptability is job search self-efficacy. The purpose of this study is to determine the effect of job search self-efficacy on career adaptability in Indonesian fresh graduates. The instruments used in this study are the Career Adapt-Abilities Scale-Short Form and Job Search Self-Efficacy Scale developed by Saks et al. (2015). Online data collected showed that job search behavior ($X^2(415)=5,017$, $p=0,025$, $R^2= 0,203$) and job search outcome ($X^2(415)=7.215$, $p= 0.007$, $R^2= 0.291$) influenced the career adaptability of fresh graduates. This indicates that the higher the level of job search self-efficacy, the higher the career adaptability. Conversely, the lower the job search self-efficacy, the lower the career adaptability. Other findings were also discussed.

Keywords: career adaptability; fresh graduate; job search efficacy

1. INTRODUCTION

Fresh graduates with bachelor's degrees will look for work so that they can earn money so that they can support themselves. However, finding a job is actually a difficult challenge for them (Satriatama, 2023). A survey by Monster (2021) of 1,000 fresh graduates revealed that 450 fresh graduates were still looking for work even though they had graduated from university a year ago. In addition, it was found that the average fresh graduate needed to apply to 23 workplaces before successfully obtaining a permanent job (Indeed, 2023). One of the reasons fresh graduates find it difficult to get a job is because companies often consider them to be unprepared for work. They are considered incapable of performing tasks and unprepared to pursue their chosen careers.

Fresh graduates adapt to their careers when they manage successful transitions amid major organizational and job market changes. A high level of career adaptability helps them cope with these transitions. Individuals with a high career adaptability are concerned about their careers (career concern) and seek jobs and careers that match their interests (career curiosity). When fresh graduates are concerned about their careers, they explore career-related knowledge related to desired jobs. Explorations that can be done by fresh graduates include improving their CVs to attract the attention of recruiters and participating in internships that are in line with their career interests. In addition, fresh graduates seek information about job vacancies and opportunities that

interest them and engage in discussions with people around them who are more knowledgeable about the world of work (career curiosity). Through these efforts, fresh graduates can gain control over their careers (career control), meaning that they themselves can decide on their preferred career choices and take responsibility for building their careers. Thus, when fresh graduates gain extensive career knowledge and readiness for job transition changes, it enhances their confidence in their desired career (career confidence), enabling them to confidently navigate a series of successful transitions in the job market and secure employment.

A factor that can increase career adaptability in fresh graduates is job search self-efficacy. Job search behaviour encourages fresh graduates to have confidence that they are capable of performing job search-related behaviours. Job search behaviour makes fresh graduates more exploratory about their careers and better prepared to face the recruitment process in their field of interest. These efforts enable fresh graduates to obtain a wealth of information about their desired field and job vacancies, giving them confidence when looking for work. After engaging in job search behaviour, fresh graduates need to have confidence in the job search outcome. Confidence in the results of their job search efforts will enable fresh graduates to better cope with obstacles and transitions to the workplace.

Fresh graduates with high job search self-efficacy are likely to have high career adaptability, making them better prepared to face transitional changes in their work. On the other hand, several previous studies by Hartono and Gunawan (2017) and Majitas and Sersic (2021) found that job search self-efficacy is the foundation for developing career adaptability skills in fresh graduates. Based on these premises, the researchers hypothesized that job search self-efficacy has an influence on career adaptability in fresh graduates.

2. METHOD

2.1. Participants

Participants of this study were individuals aged 20–26 years who had graduated from their undergraduate education no more than 12 months ago. In addition, they were not yet in permanent employment, but they could be interns, part-time employees, or freelancers. There were 417 participants, comprising 187 men and 230 women.

2.2. Instruments

This study utilised two instruments. The first instrument was the Career Adapt-Abilities Scale-Short Form (CAAS-SF) by Maggiori et al (2015), which consists of 12 items divided evenly into four dimensions and produces a total score. The items in the CAAS-SF describe an

individual's career adaptability to handle work transitions and tasks in career development. The CAAS-SF measurement tool has a reliability coefficient of 0.88 (Yu et al., 2019) and is considered valid for testing career adaptability levels. The CAAS-SF measurement tool uses a 4-point Likert scale ranging from 'strongly disagree' (1) to 'strongly agree' (4). The instrument score in this measurement tool is calculated based on the scores on each sub-scale and then added together to obtain a total score (total mean).

The second instrument is the Job Search Self-Efficacy Scale (JSSE), developed by Saks et al. (2015). The JSSE consists of 20 items comprising two dimensions, job search behaviour and job search outcome. Both dimensions have an equal number of items and produce two different scores. All items on this measurement tool will reveal a person's confidence in finding and obtaining employment. The reliability of the JSSE measurement tool is measured based on each dimension, with the job search behaviour dimension having a reliability of 0.89 and the job search outcome dimension having a reliability of 0.96 (Saks et al., 2015). The JSSE measurement tool is considered valid and can be used to test an individual's level of confidence in seeking and obtaining employment. The JSSE measurement tool uses a Likert scale with four points, ranging from 'strongly disagree' (1) to 'strongly agree' (4).

This study conducted validity and reliability tests on both instruments. The validity test used content validity through expert judgement, while the reliability test used Cronbach's alpha calculation. The test was conducted online with 32 participants, 29 women and 11 men. The validity test for the Career Adapt-Abilities Scale-Short Form (CAAS-SF) was conducted to make adjustments related to the existing context. The results of the expert judgement were then followed by a readability test with three participants who had the characteristics required for the study. The test showed that all participants understood each item of the career adaptability measurement tool (CAAS-SF), indicating that its validity was quite good and that it could be used in measuring a person's career to cope with job transitions and career development. Next, the researcher used Cronbach's alpha to test the reliability of the CAAS-SF instrument using the JASP 0.18.3.0 software application. The CASS-SF measurement tool obtained satisfactory results, namely a reliability coefficient of 0.796 with an item-rest correlation range of 0.339-0.650.

The validity test for the Job Search Self-Efficacy Scale (JSSE) measurement tool used the same procedure as the CAAS-SF and resulted in the decision that all items were well understood and no statements needed to be changed. Reliability testing was conducted on each dimension and showed that the job search behaviour dimension had an alpha coefficient of 0.731 with an internal consistency error coefficient of 0.269, while the job search outcome dimension

had an alpha

coefficient of 0.904 with an internal consistency error coefficient of 0.096. Furthermore, it was found that the range of item-rest correlation was between 0.323 and 0.569, so the JSSE instrument was declared to have good reliability and no items needed to be eliminated from this measuring tool.

2.3. Analytical Techniques

This is a quantitative approach study that aims to examine the effect of career adaptability on job search self-efficacy among fresh graduates and to identify the strength and significance of this effect. The researchers first conducted assumption tests, namely normality tests, linearity tests, error independence tests, and homoscedasticity tests, as prerequisites for simple linear regression tests. However, Shapiro-Wilk calculations found that the three data groups were not normally distributed. The job search behaviour dimension showed ($W = 0.852$, $p < 0.001$), the job search outcome dimension ($W = 0.880$, $p < 0.001$) and career adaptability ($W = 0.896$, $p < 0.001$). Based on this, simple linear regression tests cannot be used, and the analysis will use logistic regression tests.

3. RESULTS

3.1. Descriptive Overview

Descriptive overviews of the research results shows that the empirical mean for career adaptability ($M = 42.091$) is higher than the theoretical mean ($M = 30$), with a difference of 12.091 and a standard deviation of 3.033. This means that the participants in the study tend to have high career adaptability so they are prepared to face changes in job transitions.

Similar results were obtained in the measurement of job search self-efficacy. The empirical mean for the job search behaviour dimension ($M = 32.787$) was higher than the theoretical mean (M

$= 25$), with a difference of 7.787 and a standard deviation of 4.465. This means that the participants in this study have a tendency towards high job search behaviour, namely they have self-confidence when engaging in job search activities. In line with this, the empirical mean on the job search outcome dimension ($M = 33.002$) is higher than the theoretical mean ($M = 25$) with a difference of

8.002 and a standard deviation of 3.909. This means that the research participants have a tendency towards high job search outcomes, namely that they have confidence that their efforts in seeking employment will be successful.

3.2. Hypothesis Testing

Researchers grouped career adaptability scores into two categories (binary), namely low and high, in order to perform logistic regression tests (Goss-Sampson, 2019). High scores were categorised as 1 and low scores as 0. The reference for high and low categories was based on the theoretical mean of career adaptability. If the participant's score was >30 , it was categorised as (1), while if the participant's score was <30 , it was categorised as (0). A total of 415 participants had high career adaptability scores (99.52%), and 2 participants had low career adaptability scores (0.48%). The researchers also tested the hypothesis for each dimension based on the categories of participants classified as having high job search behaviour and high job search outcomes.

Table 1. Logistic Regression Test of Job Search Behaviour on Career Adaptability

X₂	df	p	Sensitivity	Nagelkerke R₂	Odds Ratio
5.017	415	0.025	1	0.203	1.339

Table 1 shows the results of the logistic regression test, namely that job search behaviour has a significant positive effect on career adaptability ($X_2(415) = 5.017$, $p = 0.025$, $R^2 = 0.203$) with data sensitivity of 100% and an odds ratio of 1.339. The odds ratio value provides an explanation of a variable's magnitude. An odds ratio value >1 indicates that a variable has a positive effect, while an odds ratio value <1 indicates that a variable has a negative effect (Goss-Sampson, 2019). The odds ratio result for job search behaviour is 1.339, meaning that the relationship between job search behaviour and career adaptability is positive. Based on the results of the logistic regression test that has been conducted, it can be said that there is a significant positive influence of job search behaviour on career adaptability among fresh graduates.

The extent of the influence of job search behaviour on career adaptability can be seen through the Nagelkerke R^2 coefficient. The Nagelkerke R^2 coefficient value shows significant results similar to R^2 in linear regression analysis (Goss-Sampson, 2019). The job search behaviour coefficient shows 0.203, $p = 0.025$. This means that job search behaviour affects career adaptability by 20.3%. These results indicate that the influence of job search outcome on career adaptability among fresh graduates is in the moderate category.

Table 2. Logistic Regression Test of Job Search Results on Career Adaptability

X₂	df	p	Sensitivity	Nagelkerke R₂	Odds Ratio
7.215	415	0.007	1	0.291	1.556

Table 2 shows that job search outcome has a significant positive effect on career adaptability ($X^2(415) = 7.215$, $p = 0.007$, $R^2 = 0.291$) with data sensitivity of 100% and odds ratio of 1.556. The odds ratio result for job search outcome is 1.556, meaning that the relationship between job search

outcome and career adaptability is positive. This means that there is a significant positive effect of job search outcome on career adaptability among fresh graduates. The job search outcome coefficient is 0.291, $p = 0.007$. This means that job search outcome has a 29.1% effect on career adaptability. These results indicate that the effect of job search outcome on career adaptability among fresh graduates is significant.

4. DISCUSSION

This study found that job search behaviour has a significant positive effect on career adaptability among fresh graduates. In the first dimension, job search behaviour affects career adaptability among fresh graduates by 20.3%. According to the researchers, this significant influence may occur because both variables in this study contain issues related to career development among fresh graduates and self-confidence in their job search abilities, which can be seen from how well fresh graduates are able to take action in situations that may arise when experiencing a career transition. Confidence in job search abilities supports fresh graduates to be more concerned about preparing for their future careers, exploring and carrying out a series of activities that interest them in the world of work.

The results of Guan et al's (2013) research show that fresh graduates with a high level of career awareness will develop self-confidence in their job search behaviour, enabling them to view job positions positively. In addition, fresh graduates with high career control will be more thorough and careful in making decisions during their career transition, enabling them to find jobs that match their values, needs and abilities. This results in job search behaviour having a significant effect on career adaptability.

The researcher interviewed two subjects to reinforce the results of the analysis in this study. Subject One said that he was confident in his ability to find a job, so after graduating, he tried to find information about job vacancies or career opportunities and find out about the background of companies he was interested in. Subject One said that after graduating, he did not waste any time and continued to apply for jobs at eighty to one hundred companies before successfully obtaining a permanent job. Subject Two, on the other hand, showed less activity in their job search and delayed looking for work for a long time. They stated that they needed peace of mind and to do fun activities with friends after graduating because they had worked

hard to complete their final assignments.

Both interviews showed that job hunting is the main thing individuals must do to transition from university to the world of work so that they can find opportunities and succeed in the job market. Behaviours that fresh graduates can adopt when looking for work include not procrastinating and continuing to apply for jobs at companies until they find permanent employment, as well as being able to seek out career opportunities and explore careers in their field of interest. This finding is also in line with a Monster survey (2021) which states that fresh graduates need to apply to 23 companies before successfully securing a permanent job.

The second dimension, job search outcome, affects career adaptability in fresh graduates by 29.1%. This may occur because both variables in this study contain issues of career development in fresh graduates and self-confidence in their job search abilities, which can be seen from how well fresh graduates are able to take action in situations that may occur when experiencing a career transition. Self-confidence gained from their efforts in seeking employment supports fresh graduates' belief in their ability to successfully implement a series of activities and make appropriate career choices in order to have good career adaptability.

Previous research by Bashori and Hidayat (as cited in Putri, 2020) stated that having career confidence in seeking employment is important for individuals because confidence enables individuals to be objective, optimistic, and responsible in their efforts to build a career. Research by Majitas and Sersic (2021) reveals that individuals who have self-confidence in seeking and obtaining employment tend to obtain information and apply for jobs well, and they even tend to perform well during the interview process. This is in line with research by Guan et al (2013) that job search outcomes have a positive contribution to career adaptability.

The researchers also conducted interviews to reinforce the results of the analysis in this study. The results of the interview with Subject One showed that he was confident that his efforts to find work would be successful, so during his studies he took the initiative to undertake internships that were relevant to his future career interests. He also said that after graduating, he actively participated in certified training courses to develop his skills and add value to himself in order to compete in the job market. Meanwhile, Subject Two did not have the confidence to be accepted for a job in line with his major, so during college he was less proactive in honing his skills. He only participated in internships when required by his university.

The results of the interviews show that having confidence in the results of job search efforts is necessary for individuals to be able to successfully implement a series of activities and make appropriate career choices. A series of activities can include planning one's future career,

showing concern for one's career by preparing for one's future career, and having the confidence to make choices in one's desired career because career selection is a process of solving various complex problems so that individuals who have confidence will be able to solve these problems (Savickas, 2005). Conversely, as was the case with subject Two, he showed little concern for his career and

preparing for his future career, and he lacked the confidence to be accepted for a job in line with his field of study, so he felt he did not perform well during the interview process. This is in line with Majitas and Sersic (2021), who state that individuals need to have self-confidence in seeking and obtaining job information and applying for jobs properly in order to be able to perform optimally during the interview process.

1) REFERENCES

- 2) Agustini, D. W. (2022). Faktor-faktor yang mempengaruhi career adaptability: Personality, emotional intelegence dan work value (suatu kajian literature review manajemen sumber daya manusia). *Jurnal Ekonomi Dan Manajemen Sistem Informasi*, 3(6), 613–620. <https://doi.org/https://doi.org/10.31933/jemsi.v3i6.1100>
- 3) Anggaraini, G. A., Djudiyah, & Syakarofath, N. A. (2023). Bagaimana peran adaptabilitas karir terhadap work engagement karyawan milenial di provinsi jakarta? *Jurnal Ilmiah Psikomuda Connectedness*, 3(1), 12–25. [https://eprints.umm.ac.id/id/eprint/878/2/Anggraini Djudiyah Syarofah-adaptabilitas karir, work engagement, karyawan milenial.pdf](https://eprints.umm.ac.id/id/eprint/878/2/Anggraini%20Djudiyah%20Syarofah-adaptabilitas%20karir,%20work%20engagement,%20karyawan%20milenial.pdf)
- 4) Artahayest, N. Y. (2021). Self-efficacy with career adaptability for final student the SWCU psychology faculty. *Jurnal Ilmiah Bimbingan Konseling Undiksha*, 12(2), 176–185. <https://doi.org/10.23887/jibk.v12i2.34018>
- 5) Azwar, S. (2012). *Penyusunan skala psikologi* (2nd ed.). Pustaka Pelajar.
- 6) Coolican, H. (2019). *Research methods and statistics in psychology* (7 edition). Routledge.
- 7) Creed, P. A., Fallon, T., & Hood, M. (2009). The relationship between career adaptability, person and situation variables, and career concern in young adults. *Journal of Vocational Behavior*, 74(2), 219–229. <https://doi.org/doi:10.1016/j.jvb.2008.12.004>
- 8) Creed, P. A., & Sawitri, D. R. (2021). Adolescent-parent career congruence as a predictor of job search preparatory behaviors: The role of proactivity. *Journal of Career Development*, 49(1)s, 60–75. <https://doi.org/10.1177/0894845321992548>

The Influence of Live Shopping on Impulsive Purchasing Through FOMO and Hedonic Shopping Motivation as Mediating Variabel for Shopee Costomers with Fashion Products

¹Andhika Yulianto, ²Windarko, ³Fazya Nabila

^{1,2,3}Department of Management & Jaya Launch Pad, Universitas Pembangunan Jaya, Indonesia

Email: win.darko@upj.ac.id

ABSTRACT

This study aims to analyze the effect of Live Shopping on Impulsive Buying, with Fear of Missing Out (FOMO) and Hedonic Shopping Motivation as mediating variables, on Shopee customers who have an interest in fashion products. The focus of this research is directed at consumers in the South Tangerang area. This research uses a quantitative approach with data collection techniques through online questionnaires, which are distributed to respondents who meet the criteria, namely active Shopee users, have participated in Live Shopping sessions, have an interest in fashion products, and live in South Tangerang. The population in this study were all people in South Tangerang, while the sample was determined as 160 respondents. Data were analyzed using the Structural Equation Modeling (SEM) method with the help of the SmartPLS 4.0 application, to test direct and indirect relationships between variables. The results of this study indicate that Live Shopping has a significant effect on Impulsive Buying. Live Shopping has a significant effect on FOMO. FOMO has a significant effect on Impulsive Buying. Live Shopping has a significant effect on Hedonic Shopping Motivation. Hedonic Shopping Motivation has a significant effect on Impulsive Buying. Live Shopping has a significant effect on Impulsive Buying through FOMO mediation and Live Shopping has a significant effect on Buyer Impulsivity through the mediation of Hedonic Shopping Motivation.

Keywords : Live Shopping, FOMO, Hedonic Shopping Motivation, Impulsive Buying, Fashion

1. INTRODUCTION

The rapid digitalization of commerce has fundamentally reshaped consumer behavior and marketing strategies, particularly in the fashion industry—one of the sectors most responsive to shifts in economic conditions and evolving consumer preferences. Amidst ongoing uncertainties marked by weakened purchasing power, stagnant household consumption, and market volatility, businesses have increasingly turned to digital platforms to sustain competitiveness and strengthen consumer engagement (Agustinus Rangga Respati, 2025). Indonesia, recognized as one of Southeast Asia's most dynamic digital economies, has experienced significant growth in e-commerce adoption. Platforms such as Shopee and TikTok Shop have led this transformation by incorporating interactive and real-time features that enhance customer experience and promote engagement, especially in visually driven product categories like fashion.

Shopee, in particular, has emerged as the dominant player in Southeast Asia's e-commerce landscape. Data from Momentum Works (2023) places Shopee at the forefront with a Gross Merchandise Value (GMV) of US\$55.1 billion—surpassing competitors such as Lazada (US\$18.8 billion), TikTok Shop (US\$16.3 billion), and Tokopedia (US\$16.3 billion). This dominance underscores a regional shift in consumer purchasing behavior from conventional retail models to digital transactions. Such trends are especially prominent in Indonesia, where Shopee is the preferred platform for online shopping, reflecting its strong resonance with local consumer needs and digital culture.



Figure 1.1 Gross Merchandise Value in Southeast Asia Source: Momentum Works

Among Shopee's most impactful innovations is live shopping, a real-time broadcast format that allows sellers to showcase products and engage directly with customers through streaming sessions (Hasibuan & Najmudin, 2024). This evolution in digital marketing allows consumers to ask questions, receive immediate responses, and make purchases on the spot, thereby transforming them from passive viewers into active participants in the sales process. Shopee, through its Shopee Live feature, has become a key proponent of this strategy, particularly in the fashion segment, where visual appeal and instant gratification are central to consumer decision-making (Refiyahya & Azhar, 2025)

The immersive and interactive nature of live shopping is particularly effective in stimulating Impulsive Buying, a spontaneous, emotionally driven form of consumer behavior. Traditionally viewed as irrational and often regretful (Fumar et al., 2023), impulse buying is increasingly understood as a strategic behavioral outcome in digital retail environments, particularly in the fashion category, where products are often bought based on emotion, style, and limited-time offers (Li & Zeng, 2025). Features such as urgency-driven promotions, exclusive discounts, and real-time social interaction heighten emotional arousal, prompting unplanned purchases that are made quickly and with little deliberation.

Central to this behavioral response are two psychological constructs: Fear of Missing Out (FOMO) and Hedonic Shopping Motivation. FOMO is the anxiety of being excluded from rewarding or time-sensitive experiences, often triggered in live shopping by scarcity cues, countdowns, and social proof elements like live viewership and user testimonials (Refiyahya & Azhar, 2025). Meanwhile, Hedonic Shopping Motivation refers to the pursuit of pleasure, excitement, and social enjoyment during the shopping process itself, especially relevant in live shopping contexts where entertaining hosts, interactive content, and real-time feedback create a sense of fun and engagement (Dewi & Adi, 2023). In fashion shopping, where aesthetic pleasure and identity expression are key, these motivations become even more prominent.

While previous studies have validated the role of live shopping in enhancing consumer engagement and trust (Huo et al., 2023), its influence on impulsive purchasing behavior, particularly through the mediating roles of FOMO and hedonic motivation, remains an area of academic contention. For instance, (Dinova & Suharyati, 2023) found a significant relationship between live shopping participation and impulse buying, whereas (Huo et al., 2023) argue that the effects are more nuanced and depend on psychological mediators. Similarly, research into the role of FOMO and hedonic motivation presents mixed findings; some support their significant influence (Safaroh, 2023), while others suggest limited or context-dependent effects (Mariyana et al., 2023).

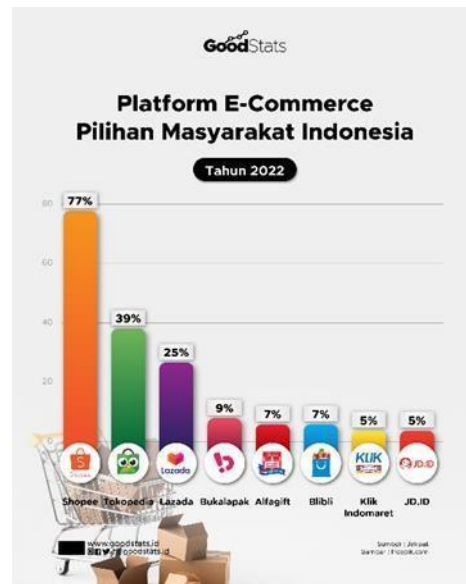


Figure 1.2 Indonesian Consumers' Preferred E-Commerce Platforms Source: Jakpat

Survey data from Jakpat (2022), involving 1,420 respondents across Indonesia, reinforces the relevance of Shopee in this research context. Not only was Shopee named the most preferred e-commerce platform by Indonesian consumers, but the fashion category also dominated online purchases at 58%, ahead of electronics (47%) and beauty products (43%). Interestingly, categories like health, hygiene (33%), fresh food (33%), and home furnishings (24%) are also gaining traction, yet fashion remains the most emotionally and visually driven segment, making it particularly susceptible to impulse buying behaviors in live shopping contexts.



Figure 1.3 Most Purchased Product Categories Source: Jakpat

Given these dynamics and the increasing importance of live shopping in digital fashion retail, this study aims to investigate the influence of live shopping on Impulsive Buying among Shopee users, focusing on fashion products, and examining the mediating effects of FOMO and Hedonic Shopping Motivation. By integrating behavioral, psychological, and technological dimensions, this research seeks to fill existing gaps in the literature and provide actionable insights for marketers and platform developers seeking to optimize consumer engagement in the evolving landscape of digital commerce.

2. LITERATURE REVIEW

2.1. Impulsive Purchasing (Y)

According to (Li & Zeng, 2025), Impulse buying occurs when a person buys goods without an urgent need, but is encouraged by attractive marketing strategies. The factors that trigger this behavior are diverse, ranging from the store environment (such as product displays and promotions), individual psychological conditions (life satisfaction, self-esteem, anxiety, and mood), to persuasive communication carried out through advertising (Paradiba et al., 2024).

(Nurhaliza, 2024) added that Impulsive Purchasing is the act of buying spontaneously, without planning, and without careful consideration of the need or value of the product being purchased with variable indicators in the form of:

1. Spontaneity,
2. Strength and Steadiness,
3. Stimulation,
4. Indifference to Consequences,

2.2. Live Shopping (X1)

According to (Saputra & Fadhilah, 2022), Live Shopping is a broadcast that is broadcast live in real time via the internet network, allowing videos to be accessed by a large audience. This feature is believed to be able to increase awareness and engagement effectively. In the context of e-commerce, live shopping allows buyers as well as sellers to communicate 2-way. During the broadcast, buyers can ask questions, provide criticism, or make suggestions, making the shopping experience more interactive (Ramadhani & Nugroho, 2024). In general, Live Shopping is often described as a live broadcast that can be accessed by many viewers simultaneously with the original broadcast time. As an interesting and interactive form of media, Live Shopping provides a unique experience for users by presenting direct communication between sellers and buyers (Yulius & Aprillia, 2023). According to research conducted by (Ho et al., 2022), there are a number of key indicators used to measure variables in Live Shopping selling. These indicators include

1. Attractiveness,
2. Trustworthiness,
3. Expertise,
4. Purchase Convenience.

2.3. Fear Of Missing Out (FOMO) (Z1)

Fear of Missing Out (FOMO) is anxiety or fear of missing out on other people's experiences. Individuals who experience FOMO often feel worried or anxious for fear of missing out on important moments experienced by friends or people around them (Nurhaliza, 2024). FOMO is a more complex psychological phenomenon, where an individual not only exhibits certain personal tendencies, but also experiences a fear of missing out on something that happens online. This condition often causes individuals to become obsessive about the activities of others, especially about trends or items on social.

2.4. Hedonic Shopping Motivation (Z2)

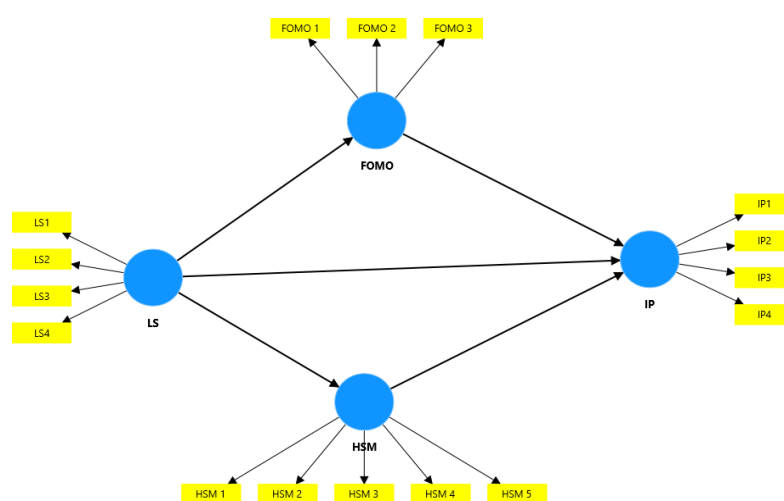
According to research (Paradiba et al., 2024), Hedonism is a concept that is the basis of a hedonic lifestyle, where a person makes worldly pleasure the main goal in his life. Individuals who embrace this lifestyle are often referred to as hedonists. Hedonic shopping Motivation or hedonic shopping motivation refers to the emotional benefits that a person gets when shopping, which reflects feelings of excitement and emotional satisfaction (Li & Zeng, 2025). This concept is also closely related to Impulse Buying behavior, where individuals who have a hedonic shopping impulse are more easily influenced to make purchases without planning

(Rabiah et al., 2024). According to (Mariyana et al., 2023), there are several main indicators in hedonic shopping motivation, namely:

1. Adventure Shopping,
2. Social Shopping,
3. Gratification Shopping,
4. Idea Shopping,
5. Role Shopping.

3. METHODS

This study applies Structural Equation Modeling (SEM) as an inferential statistical method that allows the analysis of complex relationships between independent, mediating, and dependent variables simultaneously (Rahadi, 2023). explains that in SEM there are two main models: a measurement model, which is used to assess Harahap & Nasution (2023)discriminant validity , and a structural model, which describes the relationships between hypothesized variables . This study uses SmartPLS 4.0, a statistical software widely used in SEM analysis based on Partial Least Squares (PLS-SEM).



Source: Author's Processing

Figure 3. Research Model Framework

The hypothesis in the research that refers to the research model framework is as follows.

1. H1: There is The Effect of Live Shopping on Impulsive Buying.
2. H2: There is The Effect of Live Shopping on FOMO.
3. H3: There is The Effect of FOMO on Impulsive Buying.
4. H4: There is The Effect of Live Shopping on Hedonic Shopping Motivation.
5. H5: There is The Effect of Hedonic Shopping Motivation on Impulsive Buying.
6. H6: There is The Effect of Live Shopping on Impulsive Buying through FOMO.
7. H7: There is The Effect of Live Shopping on Impulsive Buying through Hedonic Shopping Motivation.

This research uses a series of Testing Methods, namely:

- 1) Evaluation of the measurement model (Outer Model) is carried out using the Convergent Validity method, which measures the extent to which indicators in a construct have a strong correlation with latent variables. The criteria used to indicate validity are loading factor values above 0.7 and Average Variance Extracted (AVE) more than 0.50 (Hair et al., 2021).
- 2) Composite Reliability to ensure that the research instrument has high precision, consistency, and accuracy in measuring latent variables (Ghozali & Latan, 2020). This test was conducted

using SmartPLS 4.0, with Composite Reliability and Cronbach's Alpha as the main indicators to assess the extent to which indicators in a construct have strong internal consistency. A construct is considered reliable if the Composite Reliability and Cronbach's Alpha values exceed 0.70 Emta et al. (2022), which indicates that the indicators in the model have a good level of reliability and can be used for further analysis. However, according to Sugiyono (2022) A variable is said to be reliable if it provides a Cronbach Alpha value > 0.60.

- 3) Discriminant Validity to ensure that each latent variable has a clear distinction and does not overlap. Discriminant Validity tests whether the indicators of a variable are more strongly correlated with its own variable compared to other variables in the research model. The cross-loading factor test is used in this study to assess the extent to which a variable differs from other variables Tri Rachmawan & Nita Aryani (2020) adding that this validity ensures the suitability of the variable with its theoretical concept. Meanwhile, Ghazali & Latan (2020) it states that the cross-loading value must be more than 0.70 for the construct to have adequate discrimination.
- 4) Inner Model is used to analyze the causal relationship between latent variables, with the aim of knowing how much influence the independent variable has on the dependent variable, either directly or through the role of mediating variables. According to Musyaffi et al. (2022), the evaluation of the structural model aims to identify the relationship between exogenous and endogenous variables and assess the causal relationship in the research model.
- 5) Square (R^2) value is used to assess the quality of the structural model (Inner Model) by looking at the extent to which the independent variables are able to explain the dependent variable. According to Musyaffi et al. (2022), R-Square is a coefficient of determination that shows how much variation in the dependent variable can be explained by the independent variable. The R-Square value indicates the level of influence, namely 0.75 for strong influence, 0.50 for moderate influence, and 0.25 for weak influence (Ghozali & Latan, 2020).
- 6) F-Square method, according to (Ayatulloh Michael Musyaffi, 2021), the F-Square value is categorized based on the level of influence, namely 0.02 (small), 0.15 (medium), and 0.35 (large) (Ghozali & Latan, 2020). Thus, this analysis provides a deeper picture of the contribution of each exogenous variable to the endogenous variables in this study.

	FOMOHSM	IP	LS
FOMO 1	0.794		
FOMO 2	0.792		
FOMO 3	0.852		
HSM 1	0.810		
HSM 2	0.812		
HSM 3	0.798		
HSM 4	0.768		
HSM 5	0.765		
IP1		0.840	
IP2		0.820	
IP3		0.838	
IP4		0.788	
LS1			0.800
LS2			0.828
LS3			0.734
LS4			0.822

1. Path Coefficients to measure the strength and direction of the relationship between latent variables. Path Coefficients indicate how much influence the independent variable has on the dependent variable, either directly or through the mediating variable. Correlation values range from -1 to 1, with values closer to these extremes indicating the strength of the relationship and its direction, denoted by a positive (+) or negative (-) sign (Windarko et al., 2024).

2. Hypothesis testing in this study uses the bootstrap method in PLS-SEM via SmartPLS 4.0, the hypothesis is declared significant if the P-Values <0.05 and T- Statistics> 1.96 Windarko et al. (2023), which indicates that the relationship between variables is statistically reliable. The level of significance, determined using the two- tailed t-value, is 1.65 (significance level = 10%), 1.96 (significance level = 5%), and 2.58 (significance level = 1%)(Subagyo & Kristian, 2023)

4. RESULTS

1. Outer Model

Table 1. Outer Model Results

Average variance extracted (AVE)	
FOMO	0.662
HSM	0.625
IP	0.675
LS	0.635

Source: Author's Data Processing Results Using SmartPLS 4.0

The outer loading results presented in Table 1 demonstrate that all indicator values exceed the minimum threshold of 0.70, indicating adequate indicator reliability. Therefore, each indicator can be considered valid and reliable for measuring its respective construct. Specifically, items under FOMO (ranging from 0.792 to 0.852), HSM (ranging from 0.765 to 0.812), Impulse Purchase (ranging from 0.788 to 0.840), and Live Shopping (ranging from 0.734 to 0.828)

meet the standard requirements. Furthermore, the

Average Variance Extracted (AVE) values for all latent variables—FOMO (0.662), HSM (0.625), IP (Impulse Purchase) (0.675), and LS (Live Shopping) (0.635)—all exceed the 0.50 threshold, indicating sufficient convergent validity. These values suggest that more than 50% of the variance in each construct is explained by its respective indicators.

6. Composite Reliability

Table 2. Reliability Value Results

	Composite Reliability (rho_a)
FOMO	0.854
Hedonic Shopping Motivation	0.893
Impulsive Purchasing	0.893
Live Shopping	0.874

Source: Author's Data Processing Results Using SmartPLS 4.0

The results of all Cronbach Alpha in this study show all variables > 0.60. Thus, this study has indicators that are able to explain all Latent Variables of this study.

7. Discriminant Validity

Table 3. Discriminant Value Results

Discriminant Validity				
	FOMO	HSM	IP	LS
FOMO	0.813			
HSM	0.801	0.791		
IP	0.820	0.823	0.822	
LS	0.856	0.850	0.835	0.797

Source: Author's Data Processing Results Using SmartPLS 4.0

All latent variables in this study demonstrate adequate discriminant validity, as indicated by the fact that each construct's square root of the AVE (diagonal values) is greater than its correlations with other constructs (off-diagonal values). This confirms that each construct is distinct from the others. Therefore, there are no significant issues related to multicollinearity, and all variables can be retained for further analysis.

8. Inner Model (R-Square)

Table 4. R-Square Value Results

	R-square	R-square adjusted
FOMO	0.732	0.731
HSM	0.723	0.721
IP	0.767	0.762

Source: Author's Data Processing Results Using SmartPLS 4.0

Based on the test results, the R-square values for the endogenous variables indicate a strong

explanatory power of the model. Specifically, the R-square for Impulsive Buying (IP) is 0.767, suggesting that approximately 76.7% of the variance in Impulse Purchase can be explained by the predictor variables in the model. Similarly, FOMO and HSM exhibit R-square values of 0.732 and 0.723, respectively, which also reflect substantial predictive relevance. These results demonstrate that the independent variables included in the model account for a significant proportion of the variance in the respective dependent variables, indicating the model's robustness in explaining the observed relationships.

9. F-Square

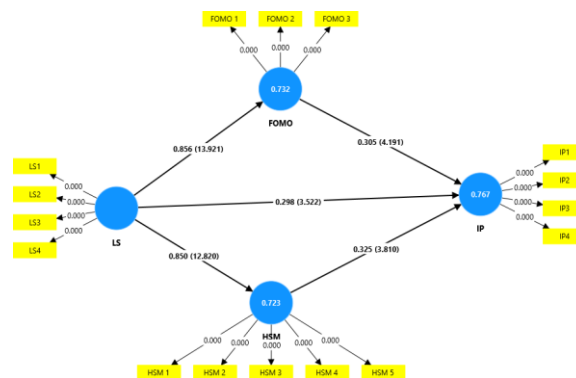
Table 5. F-Square Results

	f-square
Live Shopping -> FOMO	2.736
Live Shopping -> HSM	2.606
Live Shopping -> IP	0.073
HSM -> IP	0.116
FOMO -> IP	0.099

Source: Author's Data Processing Results Using SmartPLS 4.0

Based on the results shown in Table 5, the influence of Live Shopping on FOMO and HSM demonstrates strong effect sizes, with F-square values of 2.736 and 2.606, respectively both significantly above the 0.35 threshold for a strong effect. In contrast, the effects of Live Shopping on Impulsive Buying (IP), HSM on IP, and FOMO on IP show weak effect sizes, with F-square values of 0.073, 0.116, and 0.099 respectively each falling below the threshold for a moderate effect. These findings indicate that while Live Shopping strongly affects both FOMO and HSM, its influence as well as that of HSM and FOMO on Impulse Purchase is limited.

10. Path Coefficients Test



Source: Author's Data Processing Results Using SmartPLS 4.0

Figure 3. Bootstrap Method Data Processing Results

The results of the direct path analysis between variables in Figure 3 can be explained as follows.

- Live Shopping -> FOMO, has a path value of 0.856 which indicates that there is a strong and positive influence between Live Shopping and FOMO directly.
- Live Shopping -> Impulsive Buying, has a path value of 0.298, meaning that the direct relationship between the Live Shopping variable and the Impulsive Buying is

Moderate and positive.

c. Live Shopping -> Hedonic Shopping Motivation, with a correlation value of 0.850, indicates a strong and positive direct relationship between the Live Shopping variable and the Hedonic Shopping Motivation.

d. FOMO -> Impulsive Buying, has a path value of 0.305 which indicates that there is a moderate and positive influence between FOMO and Impulsive Buying directly.

e. Hedonic Shopping Motivation -> Impulsive Buying, has a path value of 0.325, meaning that the direct relationship between the Hedonic Shopping Motivation variable and the Impulsive Buying is moderate and positive.

Table 6. Results of Path Coefficients Values - Direct and Indirect Effects

<u>Path coefficients</u>					
<u>Mean, STDEV, T values, p values</u>					
Direct effects					
	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (O/STDEV)	P values
LS -> IP	0.298	0.297	0.085	3.522	0.000
LS -> FOMO	0.856	0.843	0.061	13.921	0.000

FOMO -> IP	0.305	0.301	0.073	4.191	0.000
LS -> HSM	0.850	0.836	0.066	12.820	0.000
HSM -> IP	0.325	0.320	0.085	3.810	0.000
<u>Specific indirect effects</u>					
	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (O/STDEV)	P values
LS -> FOMO -> IP	0.261	0.254	0.064	4.072	0.000
LS -> HSM -> IP	0.276	0.269	0.077	3.571	0.000

Source: Author's Data Processing Results Using SmartPLS 4.0

The results of the indirect path analysis between the independent variables and the dependent variables through the mediating variables in Table 6 can be explained as follows.

a. Live Shopping -> Hedonic Shopping Motivation-> Impulsive Buying, has a path value of 0.276 which shows a moderate and positive influence between Live Shopping and Impulsive Buying through Hedonic Shopping Motivation.

b. Live Shopping -> FOMO -> Impulsive Buying, has a path value of 0.261 which indicates a moderate and positive influence between Live Shopping and Impulsive Buying through FOMO.

11. Hypothesis testing

From table 6, we can see the results of the t-statistic and p-value values which show the results of the hypothesis test with the following details:

- a. H1: There is an influence between Live Shopping Variables on Impulsive Buying **accepted**. This decision is due to the t-statistic value < from the t-table ($3.522 > 1.96$) and has a p-value > 0.05 ($0.000 < 0.05$).
- b. H2: There is an influence between the Live Shopping Variable on FOMO **accepted**. This decision is due to the t-statistic value $>$ from the t-table ($13.921 > 1.96$) and has a p-value < 0.05 ($0.000 < 0.05$).
- c. H3: There is an influence between the FOMO Variable and the Impulsive Buying **accepted**. This decision is caused by the t-statistic value $>$ from the t-table ($4.191 > 1.96$) and has a p-value < 0.05 ($0.000 < 0.05$).
- d. H4: There is an influence between Live Shopping Variables on Hedonic Shopping Motivation **accepted**. This decision is caused by the t-statistic value $>$ from the t-table ($12.820 > 1.96$) and has a p-value < 0.05 ($0.000 < 0.05$).
- e. H5: There is an influence between the Hedonic Shipping Motivation Variable on Impulsive Buying **accepted**. This decision is caused by the t-statistic value $>$ from the t-table ($3.810 > 1.96$) and has a p-value < 0.05 ($0.000 < 0.05$).
- f. H6: There is an influence between Live Shopping Variables on Impulsive Buying through FOMO **accepted**. This decision is caused by the t-statistic value $<$ from the t-table ($4.072 > 1.96$) and has a p-value > 0.05 ($0.000 < 0.05$).
- g. H7: There is an influence between the Live Shopping Variable on Impulsive Buying through Hedonic Shopping Motivation **accepted**. This decision is caused by the t-statistic value $>$ from the t-table ($3.571 > 1.96$) and has a p-value < 0.05 ($0.000 < 0.05$).

5. DISCUSSION

Live Shopping has a significant impact on impulse purchasing, indicating that interactive features such as live chat, limited-time discounts, and real-time testimonials create emotional pressure that drives unplanned purchases. This finding is consistent with Wongkitrungrueng and Assarut (2020), who assert that live interaction enhances urgency and exclusivity, thereby triggering impulsive behavior. Similarly, Live Shopping has a significant effect on Fear of Missing Out (FOMO). Przybylski et al. (2013) explain that FOMO arises from the need to stay connected with ongoing social activities, which is highly relevant in the context of live commerce. During live sessions, consumers observe other buyers' comments and real-time purchase notifications, generating a fear of being left out.

FOMO also significantly influences impulse buying. It shortens the decision-making time and intensifies emotional urgency, consistent with the findings of Yoon et al. (2021), who suggest that FOMO increases emotional engagement, often leading to spontaneous purchase decisions. Additionally, Live Shopping enhances hedonic shopping motivation. This occurs because application features such as entertainment by hosts and aesthetically appealing product presentations create an enjoyable shopping experience. This is supported by Arnold and Reynolds (2003), who emphasize that hedonic motivation arises from the pleasure of product exploration and the emotional experience during shopping. Hedonic shopping motivation significantly affects impulse buying, as Hausman (2000) indicates that consumers driven by hedonic motives are more susceptible to impulse purchases due to decisions being influenced more by immediate pleasure than by rationality.

The mediation effect of FOMO is proven to be significant, indicating that social effects of live commerce, such as social proof and urgency, stimulate FOMO, which in turn triggers impulse buying. Balakrishnan et al. (2021) found that FOMO is a crucial mediator in the relationship

between digital shopping experiences and impulsive consumer behavior. Likewise, hedonic motivation also significantly mediates the effect of Live Shopping on impulse purchasing. This finding reinforces the assertion by Jiang and Benbasat (2007) that interactive and engaging experiences within digital platforms foster emotional involvement and increase the likelihood of impulse purchases.

6. CONCLUSION

Overall, the findings of this study highlight the significant role of Live Shopping as an interactive platform in driving impulse purchasing behavior. The results demonstrate that Live Shopping exerts both direct and indirect effects on impulse buying through the mediation of Fear of Missing Out (FOMO) and hedonic shopping motivation. These mediating variables serve as critical psychological mechanisms by which live interactive features, such as real-time engagement and social proof, intensify emotional urgency and reduce consumers' deliberation time, thereby fostering unplanned purchase decisions.

The significant mediating role of FOMO underscores the importance of social influence and the perception of exclusivity within live commerce environments, while the impact of hedonic motivation reflects the emotional gratification consumers derive from engaging shopping experiences. These findings contribute to a deeper understanding of consumer behavior in digital commerce and emphasize the need for platforms to integrate emotional and social elements effectively to enhance purchasing outcomes.

As a limitation, this study primarily focuses on the immediate effects of live shopping features and psychological mediators without exploring potential moderating factors such

as individual differences or cultural context. Future research is encouraged to examine these dimensions to provide a more comprehensive understanding of impulse purchasing in live commerce settings. For practitioners, it is imperative to recognize that creating an engaging, socially interactive, and emotionally stimulating live shopping experience can significantly influence consumer behavior and drive sales. An integrated approach that balances interactive features with psychological triggers will be essential for optimizing the effectiveness of live commerce strategies.

REFERENCES

- 1) Agustinus Ranga Respati, T. M. V. A. (2025). Daya Beli Masyarakat Lesu, Ekonomi Kuartal I 2025 Diproyeksi Maksimal 5,03 Persen. Kompas. <https://money.kompas.com/read/2025/04/03/145118626/daya-beli-masyarakat-lesu-ekonomi-kuartal-i-2025-diproyeksi-maksimal-503#:~:text=JAKARTA%2C KOMPAS.com – Daya beli masyarakat yang melemah,%28PHK%29 masih masif dalam dua bulan pertama 2025>
- 2) Arnold, M. J., & Reynolds, K. E. (2003). Hedonic shopping motivations. *Journal of Retailing*, 79(2), 77–95. [https://doi.org/10.1016/S0022-4359\(03\)00007-1](https://doi.org/10.1016/S0022-4359(03)00007-1)
- 3) Balakrishnan, J., Griffiths, M. D., & Shukla, P. (2021). Fear of missing out, impulsive buying, and online consumer behavior during the COVID-19 pandemic: A moderated mediation model. *International Journal of Consumer Studies*, 45(4), 1397–1410. <https://doi.org/10.1111/ijcs.12657>
- 4) Dedi Andriansah. (2023). PENINGKATAN IMPULSIVE BUYING EVENT HARBOLNAS SHOPEE MELALUI FoMO DAN HEDONIC SHOPPING. *Ekonomi Dan Bisnis: Berkala Publikasi Gagasan Konseptual, Hasil Penelitian, Kajian, Dan Terapan Teori*, 27(2), 83–90. <https://doi.org/10.24123/jeb.v27i2.5810>
- 6) Dewi, N. W. R. R., & Adi, I. N. R. (2023). Shopping Lifestyle, Fashion Involvement,

- Hedonic Shopping Motivation, Positive Emotion, Impulse Buying E-Commerce. Quantitative Economics and Management Studies, 4(6), 1082–1092. <https://doi.org/10.35877/454ri.qems2032>
- 7) Dinova, S., & Suharyati. (2023). Pengaruh Live Streaming Shopping dan Flash Sale Terhadap Pembelian Implusif Pengguna Shopee di Kecamatan Jatinegara. In Journal Of Young Entrepreneurs (Vol. 2, Issue 4). <https://ejournal.upnvj.ac.id/index.php/jye>
 - 8) Fahri, M. A., Widodo, A., & Rubiyanti, N. (2025). Locus of Control Mediates Impulsive Buying at Coffee Shops in Kendari City A Conceptual Paper. 4(1), 15–28.
 - 9) Fumar, M., Setiadi, A., Harijanto, S., & Tan, C. (2023). The Influence of Fear of Missing Out (FOMO), Sales Promotion, and Emotional Motive Mediated Self- Control on Impulsive Buying for Hypebeast Products. Riwayat: Educational Journal of History and Humanities, 6(3), 1363–1375. <http://jurnal.unsyiah.ac.id/riwayat/>
 - 10) Hasibuan, R. R., & Najmudin, N. (2024). Content Marketing, Customer Engagement On Marketing Performance Mediated By Digital Marketing In Batik Msmes In Banyumas Regency. International Journal of Economics, Business and Innovation Research, 3(03), 287–304. <https://www.e-journal.citakonsultindo.or.id/index.php/IJEBIR/article/view/859>
 - 11) Hausman, A. (2000). A multi-method investigation of consumer motivations in impulse buying behavior. Journal of Consumer Marketing, 17(5), 403–426. <https://doi.org/10.1108/07363760010341045>
 - 12) Ho, C. I., Liu, Y., & Chen, M. C. (2022). Antecedents and consequences of consumers' attitudes toward live streaming shopping: an application of the stimulus– organism– response paradigm. Cogent Business and Management, 9(1). <https://doi.org/10.1080/23311975.2022.2145673>
 - 13) Huo, C., Wang, X., Sadiq, M. W., & Pang, M. (2023). Exploring Factors Affecting Consumer's Impulse Buying Behavior in Live-Streaming Shopping: An Interactive Research Based Upon SOR Model. SAGE Open, 13(2), 1–15. <https://doi.org/10.1177/21582440231172678>
 - 14) Jiang, Z., & Benbasat, I. (2007). The effects of presentation formats and task complexity on online consumers' product understanding. MIS Quarterly, 31(3), 475–500. <https://doi.org/10.2307/25148804>
 - 16) Li, H., & Zeng, W. (2025). Consumption, relative deprivation and mental health: evidence from hedonic consumption. Frontiers in Public Health, 13. <https://doi.org/10.3389/fpubh.2025.1511547>
 - 17) Mariyana, M., Djawoto, D., & Suhermin, S. (2023). The influence of hedonic shopping motivation and shopping lifestyle on impulse buying through positive emotion in e-commerce. Manajemen Dan Bisnis, 22(2), 25. <https://doi.org/10.24123/jmb.v22i2.691>
 - 18) Nurhaliza. (2024). PENGARUH FOMO , SHOPPING LIFESTYLE , DAN HEDONIC SHOPPING MOTIVATION TERHADAP IMPULSE BUYING DECISION PADA BEAUTY E-COMMERCE BEAUTY HAUL (STUDI KASUS MAHASISWA FAKULTAS EKONOMI DAN BISNIS UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA) SKRIPSI OLEH : NURHALIZA.
 - 19) Paradiba, A. R., Palandeng, I. D., Poluan, J. G., Manajemen, J., Ekonomi, F., Ratulangi, U. S., Paradiba, A. R., Palandeng, I. D., & Poluan, J. G. (2024). PENGARUH HEDONIC LIFESTYLE DAN HEDONIC SHOPPING MOTIVATION TERHADAP PEMBELIAN IMPULSIF KONSUMEN (STUDI PADA MAHASISWA S1 MANAJEMEN UNIVERSITAS SAM RATULANGI MANADO PENGGUNA TIKTOK SHOP) THE INFLUENCE OF HEDONIC LIFESTYLE AND HEDONIC SHOPPING MOTIVATION . 12(3), 868–879.
 - 20) Przybylski, A. K., Murayama, K., DeHaan, C. R., & Gladwell, V. (2013). Motivational,

- emotional, and behavioral correlates of fear of missing out. *Computers in Human Behavior*, 29(4), 1841–1848. <https://doi.org/10.1016/j.chb.2013.02.014>
- 21) Rabiah, A. S., Setiawan, M., Rohman, F., & Mugiono. (2024). Assess the Impact of Hedonic Motivation, Social Influence, Gamification, and Education Level on Behavioral Intention To Adopt Mobile Shopping in Jakarta, Indonesia. *Revista de Gestao Social e Ambiental*, 18(7), 1–24. <https://doi.org/10.24857/RGSA.V18N7-008>
 - 22) Ramadhani, L. N., & Nugroho, D. A. (2024). Pengaruh Live Streaming, Flash Sale, Dan Hedonic Shopping Motivation Terhadap Impulsive Buying. *Jurnal Manajemen Pemasaran Dan Perilaku Konsumen*, 3(1), 207–215. <https://doi.org/10.21776/jmppk.2024.03.1.21>
 - 23) Refiyahya, A. H., & Azhar, A. (2025). DARI LIVE STREAMING HINGGA CHECKOUT : MEMAHAMI IMPULSE BUYING DI SHOPEE. 9(1), 210–216.
 - 24) Safaroh, N. (2023). Pengaruh Fear of Missing Out (FoMO), Shopping Enjoyment, dan Hedonic Shopping Motivation Terhadap Impulse Buying di E-Commerce Shopee Pada Waktu Flash Sale. 5, 34–38. <https://doi.org/10.29303/alexandria.v5iSpecialIssue.604>
 - 25) Saputra, G. G., & Fadhilah, F. (2022). Pengaruh Live Streaming Shopping Instagram Terhadap Kepercayaan Konsumen Online dan Dampaknya pada Keputusan Pembelian. *Ekonomi, Keuangan, Investasi Dan Syariah (EKUITAS)*, 4(2), 442–452. <https://doi.org/10.47065/ekuitas.v4i2.2353>
 - 26) Wongkitrungrueng, A., & Assarut, N. (2020). The role of live streaming in building consumer trust and engagement with social commerce sellers. *Journal of Business Research*, 117, 543–556. <https://doi.org/10.1016/j.jbusres.2018.08.032>
 - 27) Yoon, S., Narasimhan, B., & Kim, K. (2021). Fear of missing out (FOMO) and impulse buying: A moderated mediation model. *Journal of Retailing and Consumer Services*, 63, 102682. <https://doi.org/10.1016/j.jretconser.2021.102682>
 - 28) Yulius, A., & Aprillia, A. (2023). Pengaruh Live Streaming Terhadap Pembelian Impulsif pada Platform Tiktok. In *Jurnal Ilmiah Mahasiswa Ekonomi Manajemen* (Vol. 8, Issue 4). https://www.researchgate.net/profile/Ariesya-Aprillia/publication/375991994_PENGARUH_LIVE_STREAMING_TERHADAP_PEMBELIAN_IMPULSIF_PADA_PLATFORM_TIKTOK/links/656632303fa26f66f4356a8e/PENGARUH-LIVE-STREAMING-TERHADAP-PEMBELIAN-IMPULSIF-PADA-PLATFORM-TIKTOK.

Social Entrepreneurship in Indonesia: Social Capital, Financial Capital, Market Orientation on Social Entrepreneurship Performance and its Impact of Economic Empowerment and Social Impact of Local Communities

Lutfi Nurizki, Yusuf Iskandar, Dede Suleman, Andini Putri Erianti Department of Management & Jaya Launch Pad, Universitas Pembangunan Jaya, Indonesia
lutfi.nurizki@student.upj.ac.id; yusuf.iskandar@upj.ac.id dede.suleman@upj.ac.id;
andini.putrierianti@student.upj.ac.id

ABSTRACT

Social entrepreneurship has become a potent tool for promoting economic empowerment and good change in local communities. In the framework of social entrepreneurship in Indonesia, this quantitative study investigates the connections between important drivers. It specifically looks into how social capital, financial capital, and market orientation influence local communities' social impact, economic empowerment, and social entrepreneurial success. The study uses partial least squares (PLS) path modeling in conjunction with structural equation modeling (SEM) to examine data gathered from 473 participants in social entrepreneurship across a range of Indonesian industries. According to the findings, social capital, financial capital, market orientation, social entrepreneurship success, economic empowerment, and social impact are all significantly positively correlated. The results highlight the significance of fostering both financial and social capital, implementing market-driven strategies, and striking a balance between the social effect and the performance of social entrepreneurship. These results highlight the critical elements that propel positive change and sustainable development in the Indonesian setting, with practical implications for policymakers, social entrepreneurs, and impact investors.

Keywords: Social Capital; Financial Capital; Market Orientation; Social Entrepreneurship; Empowerment; Social Impact; Indonesia

1. INTRODUCTION

Social and environmental issues are addressed, and sustainable economic growth is encouraged (Bidet & Defourny, 2019; Qastharin, 2016). A flat organizational structure that promotes participation and ownership among all stakeholders, a focused organizational goal (social purpose), and an active interaction between ambivalent positive and negative feelings associated with the social mission pursued by the social entrepreneur are all key components of the practice of social entrepreneurship (Costa & Andreus, 2021). One method for addressing the problems faced by disenfranchised and underprivileged kids is social entrepreneurship (Hausberg & Korreck, 2020; Ozkazanc-Pan & Clark Muntean, 2018). Research on social entrepreneurship has improved our comprehension of the dynamics and procedures that underlie constructive social transformation. However, the location of that alteration has received little consideration (Klapper & Farber, 2016; Ozkazanc-Pan & Clark Muntean, 2018; Sabeti, 2011). Understanding the extra-organizational settings represented by the word "social" in the phrase "social entrepreneurship" requires community-level examination.

The goal of social entrepreneurship is to address difficult social issues while promoting economic empowerment and constructive change in local communities. It stands at the nexus of social and economic growth (Murniati et al., 2021; Segun, 2021). Indonesia, with its distinctive socioeconomic landscape and cultural variety, is not an exception to the field's substantial global growth (Iskandar & Kaltum, 2021). The advent of social entrepreneurship has provided creative solutions to a range of urgent problems, from poverty alleviation and education access to environmental sustainability and community development, in a nation known for its lively traditions and dynamic economy (Sari & Kusumawati, 2022; Tabasum & Shaikh, 2022; Tabuena et al., 2022).

An archipelago of thousands of islands called Indonesia is home to a diversified population that

struggles with a number of socioeconomic inequalities (Anggahegari et al., 2021; Awaluddin et al., 2021; Margiono & Feranita, 2021). One method of addressing these gaps is through social entrepreneurship, which has been a catalyst for change. Its influence can be seen in the creation of programs that close gaps in access to economic, medical, and educational possibilities, as well as in measures to safeguard the environment and strengthen neglected groups (Bacq & Eddleston, 2018; Eikenberry & Kluver, 2004).

Indonesia provides an intriguing setting for researching social entrepreneurship efforts that place a priority on social and environmental results because it is one of the most populated nations in the world and

is home to many different cultures and communities (Iskandar et al., 2021). Social entrepreneurs in Indonesia have led a number of initiatives in recent years that not only produce financial value but also advance the country's sustainable development goals (Calvo et al., 2020; Murniati et al., 2021). With a special emphasis on the critical facets of social capital, financial capital, market orientation, and the connections between them, this study begins an in-depth investigation of social entrepreneurship in Indonesia. Additionally, it aims to determine the extent to which these variables affect social business success before evaluating how they affect local communities' social welfare and economic empowerment.

The ties, networks, and conventions that promote collaboration and coordination between individuals and groups are referred to as social capital (Andriani & Christoforou, 2016). Social capital can assist social entrepreneurs in gaining access to the finance, knowledge, and information they need to make their businesses successful (Sari & Kusumawati, 2022; Smith et al., 2016; Yohanes et al., 2017). Social capital can also assist social entrepreneurs in developing credibility and trust with partners, clients, and investors, which can improve their standing and raise their chances of success (Lang et al., 2022; Osei & Zhuang, 2020).

Financial capital is the cash and other resources that social entrepreneurs require to launch and expand their businesses (Hidayat et al., 2022; Ngeek, 2016; Yusuff et al., 2018). Financial resources can assist social entrepreneurs in investing in the facilities, tools, and personnel necessary for their ventures to be successful (Berno, 2017; Bontis et al., 2018; Fatoki, 2011). Additionally, it can encourage social entrepreneurs to take chances and pursue novel concepts with the potential to have a positive social and economic impact (Dar & Mishra, 2020; Mansour et al., 2018; Yusuff et al., 2018).

The degree to which social entrepreneurs prioritize gratifying the needs and desires of clients and other stakeholders is referred to as market orientation (Evers & Schulze-Böing, 2001). Market opportunities, such as unmet needs or neglected groups, that have the potential to generate social and economic value can be identified and taken advantage of by social entrepreneurs with the aid of market orientation (Glaveli & Geormas, 2018; Karnowati & Handayani, 2022). Market orientation can also assist social entrepreneurs develop bonds of trust, mutual benefit, and shared values with clients, investors, and partners, which can improve their reputation and boost their chances of success (Melly et al., 2022).

In conclusion, market orientation, social capital, and financial capital all play significant roles in boosting the success of social entrepreneurship, which can have a good effect on local communities' social and economic empowerment. By making use of these elements, social entrepreneurs can gain access to the tools, knowledge, and networks they require to develop original responses to societal and environmental problems and establish long-lasting, significant businesses (Fulgence & Mori, 2009; Halder, 2019; Murniati et al., 2021).

The need for this research is driven by concerns about entrepreneurship in Indonesia, particularly social entrepreneurship, which is expanding and playing a bigger role in social missions. In Indonesia, social entrepreneurship—which employs cutting-edge business concepts to produce social and environmental impact—has grown in popularity as a strategy for tackling societal issues (Anggahegari et al., 2021; Margiono & Feranita, 2021). While there

is no doubt about the potential for positive change, there is an urgent need for a thorough quantitative analysis to identify and quantify the critical elements that influence the performance of social enterprises and their effects on local economic empowerment and community development.

Indonesia is interested in fostering social entrepreneurship as a means of achieving the UN Sustainable Development Goals (SDGs), as it is a signatory to the SDGs. A key component of the SDGs is sustainable development, which involves social welfare and economic empowerment. Measuring the effect of social entrepreneurship on these facets of development in the Indonesian context is essential in order to effectively contribute to the SDGs.

In Indonesia, local populations are frequently neglected and marginalized. These problems may be resolved by social entrepreneurship by promoting neighborhood growth, generating job opportunities, and enhancing general wellbeing (Herlina et al., 2021; Zid et al., 2020). However, it is still mostly unclear and requires quantifying how social entrepreneurship can effectively empower local communities. Governmental organizations, decision-makers, and impact investors are becoming more and more interested in promoting social entrepreneurship as a way to meet larger development objectives (Sari & Kusumawati, 2022). Given Indonesia's expanding social entrepreneurship environment, this research is especially important.

Social entrepreneurship operates within complex social, economic, and cultural circumstances; it does not exist in a vacuum. It's crucial to comprehend how market orientation, social capital, and financial capital interact and affect entrepreneurial performance. This study tries to clarify these intricate connections. This study is urgent because we must hasten the good social and economic effects. A quantitative comprehension of the variables that influence performance can help social entrepreneurs in their active efforts to tackle urgent concerns. The urgency of this research is highlighted by its alignment with the UN SDGs and Indonesia's national development ambitions. The results of this study will assist in achieving these international and national objectives by providing knowledge and guidance.

2. LITERATURE REVIEW

2.1 Social Entrepreneurship in Indonesia

Indonesia, with its vast population and diverse culture, provides a unique backdrop for social entrepreneurship. The country has a long history of community-based initiatives aimed at addressing societal challenges (Singgalen et al., 2019). Social entrepreneurship in Indonesia often combines commercial strategies with a social mission, creating innovative solutions to local problems (Anggahegari et al., 2021; Margiono & Feranita, 2021). These initiatives have the potential to drive economic empowerment and social impact by targeting issues such as poverty, environmental sustainability, and access to education (Hermanto, 2017; Idris & Hijrah Hati, 2013).

Indonesia offers a distinctive environment for social entrepreneurship because of its lengthy history of community-based programs that seek to address social issues (Utomo et al., 2019). In Indonesia, social entrepreneurship frequently blends business tactics with charitable endeavors, resulting in creative answers to regional issues (Gunawan & Fraser, 2016). By focusing on problems like poverty, environmental sustainability, and educational access, these programs have the potential to promote social impact and economic empowerment (Herlina et al., 2021; Zid et al., 2020).

Numerous current research works have investigated the function of social entrepreneurship within Indonesia. For instance, one study looked at university students' risk perception, intellectual capital, and cultural intelligence in relation to the emergence and growth of social entrepreneurship (Iskandar & Kaltum, 2021; Sari & Kusumawati, 2022). A different study

looked at the relationship between social entrepreneurship and economic sustainability, specifically as it relates to information and communication technology (Frączkiewicz-Wronka & ..., 2014; Mulgan et al., 2007). A survey of the literature on women's social entrepreneurship in Indonesia revealed gaps in the knowledge about women's participation in social entrepreneurship (Anggadwita et al., 2017; Gah et al., 2020; Hani et al., 2012). Lastly, a case study looked at how social entrepreneurship functions in Pesantren, Indonesia's Islamic solidarity economy (Abdullah & Hoetoro, 2020; Saputra et al., 2021).

2.2. Social Capital in Social Entrepreneurship

In Indonesia, social capital is essential to the success of social entrepreneurship programs because it enables social entrepreneurs to effectively collaborate, mobilize community resources, and establish trust (Sari & Kusumawati, 2022; Yohanes et al., 2017). (Sari & Kusumawati, 2022) performed research in Indonesia and highlighted the role that social capital networks have in supporting micro and small social companies' operations. Strong social capital ties help entrepreneurs locate possible partners, obtain money, and build credibility—all of which are critical to the success of social enterprises (Estrin et al., 2016; Guliyeva et al., 2021; Robinson, 2006).

The connection between social capital and entrepreneurship in Indonesia has also been examined in other research. As an illustration, a study conducted in (Sari & Kusumawati, 2022). discovered that social capital significantly influences a rural community's intention to become an entrepreneur (Osei & Zhuang, 2020; Pickernell et al., 2007). Another study (Iskandar & Kaltum, 2021) looked into students' risk perception, intellectual capital, and cultural intelligence in relation to the emergence and growth of social entrepreneurship. The findings demonstrated that while risk perception had a negative impact on social entrepreneurship, cultural intelligence and intellectual capital had a favorable impact (Guliyeva et al., 2021; Hota, 2023).

All things considered, social capital plays a critical role in the accomplishment of social entrepreneurship projects in Indonesia. Strong social capital ties give entrepreneurs access to information, resources, and assistance—all of which are critical for the expansion and long-term viability of social companies.

H1 : Social capital is suspected to be a positive and significant factor in the performance of social entrepreneurs.

H2 : Social capital is suspected to be a positive and significant factor in Economic Empowerment

2.3. Financial Capital in Social Entrepreneurship

Financial capital is a crucial element in social entrepreneurship, enabling social entrepreneurs to sustain their initiatives and drive growth. Access to funding and financial resources has a significant impact on the performance and success of social enterprises (Dar & Mishra, 2020; Mansour et al., 2018). Social enterprises require financial sustainability to maintain operations, expand their reach, and maximize social impact (Widjjaningsih et al., 2022; Yusuff et al., 2018). The growth of social enterprises in Indonesia underscores the pivotal role of financial capital in their development.

Research has shown that financial capital has an independent effect on the income of MSMEs actors (Hidayat et al., 2022). Intellectual capital management can also lead to long-term profits and improve organizational performance, creating a competitive advantage for companies (Buallay et al., 2019; Diugwu, 2011). Social capital, which is a resource embedded in social relationships, can promote rural entrepreneurship by sharing financial literacy (Osei & Zhuang, 2020; Pickernell et al., 2007). Microfinance services, such as micro-credit, micro-savings, micro-insurance, training, and social networking, are essential drivers for SMEs' survival and can enhance the productivity of microfinance services for MSEs (Anwar et al., 2023;

Ardiansyah, 2016; Lawhaishy & Othman, 2023).

In conclusion, financial capital is a crucial element in social entrepreneurship, enabling social entrepreneurs to sustain their initiatives and drive growth. Access to funding and financial resources has a significant impact on the performance and success of social enterprises in Indonesia.

H3 : Financial capital is suspected to be a positive and significant factor in the performance of social entrepreneurs.

H4 : Financial capital is suspected to be a positive and significant factor in Economic Empowerment

2.4. Market Orientation in Social Entrepreneurship

Market orientation is a crucial aspect of social entrepreneurship, as it involves aligning business strategies with market dynamics and consumer needs to drive success and sustainability (Evers & Schulze- Böing, 2001; Karnowati & Handayani, 2022). Studies have explored the significance of market orientation in the context of social entrepreneurship in Indonesia and have emphasized the value of market-oriented strategies in achieving both financial sustainability and societal impact (Sari & Kusumawati, 2022). Social entrepreneurs who effectively adapt to market dynamics tend to perform better and create more significant economic empowerment and social impact (Glaveli & Geormas, 2018; Karnowati & Handayani, 2022).

Market-driven social entrepreneurship emphasizes the development of innovative solutions and the identification of market opportunities to maximize social and economic outcomes (Diochon & Ghore, 2016; Karnowati & Handayani, 2022). In the Indonesian market, external pressure can positively enhance the social performance of NGOs, and the application of the principles of public entrepreneurship can be quite effective in the management of public assets (Abidin et al., 2022; A Octavia & Ali, 2017).

Education attainment levels can reduce the effect of market orientation on entrepreneurship orientation and increase the effect of technology orientation on entrepreneurship orientation (Evers & Schulze-Böing, 2001; Karnowati & Handayani, 2022; Nursal et al., 2022). The ability to negotiate skill and risk-taking has a significant effect on both entrepreneurial orientation and market orientation, but optimism and locus of control have no effect on market orientation and orientation (Baker et al., 2020; Ade Octavia et al., 2020). Sustainability performance is influenced by market orientation and entrepreneurship, which are capability dimensions (Narver & Slater, 1990).

In conclusion, market orientation is a crucial aspect of social entrepreneurship in Indonesia, and social entrepreneurs who effectively adapt to market dynamics tend to perform better and create more significant economic empowerment and social impact.

H5 : Market Orientation is suspected to be a positive and significant factor in the performance of social entrepreneurs.

H6 : Market Orientation is suspected to be a positive and significant factor in Economic Empowerment

3. Relationships with Economic Empowerment and Social Impact

The idea of "social entrepreneurship" attempts to promote social impact and economic development in local communities. By addressing social issues and coming up with creative solutions that benefit the community, social entrepreneurship seeks to improve society. The social impact and economic empowerment of local communities can be greatly impacted by effective social entrepreneurship performance. Social investment initiatives have been put into

place in Nigeria in an effort to combat poverty, generate employment, and foster social and economic advancement (Yarima et al., 2023). According to a study, there is a favorable and noteworthy correlation between these programs' recipients' development of entrepreneurship and economic empowerment.

In nations like Namibia, the advancement of women's entrepreneurship and human capital depends on their economic empowerment (Bobek et al., 2022). However, sociocultural norms, a sizable and unprotected informal economy, and lack of access to resources and education provide obstacles for female entrepreneurs. Supporting the unorganized sector, upholding gender-responsive laws and policies, and granting access to resources and knowledge are all ways that governments and other institutions can encourage women's empowerment.

Youth economic empowerment can also benefit from social entrepreneurship. Due to its promotion of entrepreneurial endeavors as the primary means of addressing societal issues, social entrepreneurship has become a prominent topic in Kaduna, Nigeria's social arena (Adomdza et al., 2017). Women in rural India can use social entrepreneurship as a vehicle for sustainable development that leads to economic empowerment (Samineni, 2018). In order for women to economically empower themselves and support the development of their country, they must work together to create an ecosystem. Women in rural India can have better prospects for a living by participating in social projects carried out at different levels by community-based organizations founded by mutual assistance groups.

Through social impact bonds, social entrepreneurship can also promote economic empowerment for reentering citizens (Toussaint, 2016). The stigma of having a criminal record follows a person who has served time in prison for a long time. This stigma affects the person's civil rights and decreases their chances of being successful in the job market. The impact investing strategy known as "social impact bond" has the potential to lower the recidivism rate among ex-offenders.

In Indonesia, social entrepreneurs operate in a variety of ways. Social effect is another important metric used to assess social entrepreneurship performance, in addition to typical company performance measures. One particular problem these entrepreneurs face is striking a balance between social effects and financial profits. Studies reveal that social firms that proficiently oversee this equilibrium typically exhibit superior performance in both aspects (e.g., Rachmat et al., 2021).

One of the main objectives of social entrepreneurship is the economic empowerment of local communities and people. Social entrepreneurs play a major role in economic development through expanding access to sustainable livelihoods, improving education and healthcare, and establishing income-producing opportunities (Javed et al., 2019; Mahato et al., 2022). Furthermore, the programs' social impact encompasses domains including education, healthcare, and environmental sustainability, ultimately enhancing individuals' overall well-being (Gigauri, 2022).

The importance of assessing the influence of social entrepreneurship on the Sustainable Development Goals (SDGs) is highlighted by Indonesia's commitment to the SDGs. The potential for accomplishing more comprehensive societal goals is shown by the alignment between social entrepreneurship and the SDGs in

areas like gender equality, high-quality education, and poverty reduction (Lang et al., 2022; Mahato et al., 2022).

To sum up, social entrepreneurship can promote social impact and economic empowerment in nearby areas. The social impact and economic empowerment of local communities can be positively impacted by effective social entrepreneurship performance.

In Indonesia, the ultimate goals of social entrepreneurship are to promote social impact and economic emancipation. Both the social impact of local communities and economic empowerment are positively impacted by effective social entrepreneurship performance.

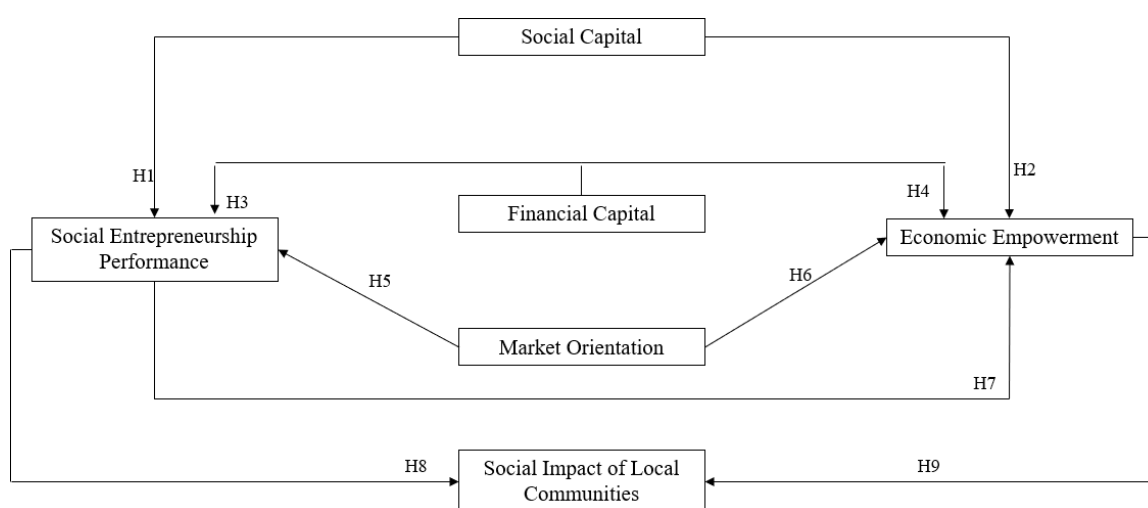
According to (Iskandar & Kaltum, 2021), social entrepreneurs who achieve success are more likely to foster beneficial social consequences and contribute to economic development. In (Anggahegari et al., 2021; Margiono & Feranita, 2021; Singgalen et al., 2019) carried out a study on the effects of social entrepreneurship in Indonesian communities, emphasizing the contribution that social entrepreneurs make to the improvement of community well-being. This study highlighted the multifaceted effectiveness of these efforts in Indonesia by presenting evidence that social entrepreneurship is not only an economic accelerator but also a driver of social transformation.

H7 : Social Entrepreneurship is suspected to be a positive and significant factor in the Economic Empowerment

H8 : Social Entrepreneurship is suspected to be a positive and significant factor in the Social Impact of Local Communities

H8 : Economic Empowerment is suspected to be a positive and significant factor in the Social Impact of Local Communities

4. Research Gaps



The literature currently in publication highlights significant research gaps, despite the growing interest in social entrepreneurship in Indonesia as a catalyst for economic empowerment and positive community impact. These gaps are especially evident in the lack of thorough and connected studies on the joint effects of market orientation, financial capital, and social capital on social entrepreneurship performance and on social impact and economic empowerment for local communities. Furthermore, in order to improve policy decision-making and the efficacy and durability of social entrepreneurship initiatives, focused academic exploration and practical insights are needed to address the dearth of research on Indonesia's particular challenges, opportunities, and cultural factors affecting social entrepreneurship, as well as the need for culturally appropriate impact assessment tools and the understudied area of long-term sustainability in such initiatives. Figure 1 displays the conceptual framework and research hypotheses.

Figure 1. Conceptual Framework and Hypotesis

5. METHOD

5.1 Design & Sample

The purpose of this questionnaire survey study was to get opinions from Indonesian social enterprise owners. Respondents helped collect data for the study. This study encompasses all provinces that have social enterprises, defined as businesses that are social in nature and

contribute to the empowerment of local communities. The questionnaire was first created with input from two professors and ten students from the entrepreneurship science clusters at three Indonesian state institutions before it was distributed. The questionnaire was tested initially, following conversations with instructors and students. In order to eliminate confusion, the cover letter for this survey includes definitions and a brief history of social entrepreneurship as well as information on how it affects local community empowerment in line with the goals of the study. Due to the geographical breadth of Indonesia, the questionnaire distribution process lasted seven months, starting on February 29 and ending on August 11, 2023. There were two parts to the research survey: an offline survey that focused on the provinces of West Java, Banten, DKI Jakarta, East Java, Central Java, Yogyakarta, and South Kalimantan, and an online survey that was conducted throughout the province using Google but did not list the provinces that were surveyed offline. After distributing 500 questionnaires to Indonesian social entrepreneurs, the author received 473 valid and accurately completed questionnaires back. Factor loadings, variable validity and reliability, correlations, multicollinearity assumptions, and structural equation modeling of effect sizes were all part of the data analysis process. Using Smart-PLS 4, SmartPLS GmbH, Oststeinbek, Germany, the study was assessed. Results from the measured questionnaire items were doable.

The variables that are covered by the expanding body of scientific literature (Andriani & Christoforou, 2016; Hausberg & Korreck, 2020; Karnowati & Handayani, 2022; Melly et al., 2022; Rianto et al., 2022; Widjianingsih et al., 2022; Yao & Meng, 2022) were social capital, financial capital, market orientation, social entrepreneurship performance, economic empowerment, and social impact of local communities. These variables were covered by the self-administered questionnaire used for data collection. Because the researchers' focus was on social enterprises in Indonesia, they utilized a non-probabilistic random sampling technique to choose samples for their study. If it were okay with them, the researchers requested the owners of social enterprises in Indonesia to fill out the questionnaire and gave their permission to gather data. By withholding their identities from the responders and visitors, the researchers allayed their anxieties regarding anonymity.

Table 1. Respondent Demographics

Age	N	%	Business Experience	N	%
18 – 25 years	78	16.01%	< 5 years	172	35.31%
26 – 35 years		22145.36%	5 – 10 years	137	28.12%
36 – 45 years		10120.74%	11 – 15 years	128	26.26%
46 – 55 years	42	8.64%	15 – 20 years	35	7.20%
> 55 years	45	9.24%	> 20 years	15	3.08%
Education	N	%	Business Scale	N	%
Bachelor's		13227.09%	Micro (< 10 employee)	121	24.85%
Master's		19740.45%	Small (10 – 49 employee)	197	40.45%
Ph.D		13828.30%	Medium (50 – 249 employee)	101	20.74%
Others	20	4.10%	Large (> 250 employee)	68	13.95%
Business income (per year/Rupiah)	N	%	Total Asset (Rupiah)	N	%
2.000.000.000.00	227	46.56%	1.000.000.000.00	189	38.79%
2.000.000.000.00-15.000.000.000.00	132	27.09%	1.000.000.000.00-5.000.000.000.00	201	41.27%
15.000.000.000.00-50.000.000.00	128	26.26%	5.000.000.000.00-10.000.000.000.00	88	18.06%

Province	N	%	Province Geographic	N	%
Geographic					
West Java	109	22.37%	Jawa Tengah	43	8.84%
Banten	32	6.57%	DIY Yogyakarta	32	6.57%
DKI Jakarta	60	12.31%	Kalimantan Selatan	22	4.52%
Jawa Timur	70	14.39%	Other's (Online survey)	119	24.44%

Source: The results of the author's data analysis (2023)

The demographics of this research are shown in Table 1 above. A total of 487 social enterprise owners in Indonesia were surveyed, and the age range of 26 to 35 years old accounted for the majority of respondents (45.36%). This suggests that a higher percentage of respondents from this age group participated in the research because they were highly active and involved in the social world. It is evident that greater business experience correlates with older age groups. The idea that elder respondents have had more time to gain business expertise helps to explain this. A Master's degree is held by 40.45% of respondents, showing that respondents typically have a higher educational background. This could indicate that social entrepreneurs typically hold more advanced degrees. 40.45% and 24.85% of the respondents, respectively, have medium- and small-sized firms. This is in line with the features of social companies in Indonesia, which tend to function on a more controllable scale. Most respondents fell into the medium group when it came to overall firm assets and annual income. our could suggest that the majority of social enterprises in our study are located in the economy's middle sector.

Most of the responders were from Java, specifically from East and West Java. This has to do with Indonesia's commercial and economic hubs, many of which are found on the island of Java. A sizable portion of respondents (24.44%) completed the poll online. This indicates that a variety of places, even those not directly covered by the in-person survey, contributed to the participation.

The demographic distribution, with the majority in the productive age range, well educated, and mostly employed by medium-sized businesses, may mirror the features of social enterprises in Indonesia. The results, which show respondents' predilection toward digital devices and geographic variety, may be influenced by the high participation rate from Java and the use of online surveys. The age of the respondents and their degree of education can be correlated. Higher education levels may be more common among older responders. Local business and economic conditions mirror geographic distribution. There might be an increase in social entrepreneurs on Java, the island with a more developed economy. The degree of technology access and proficiency among respondents may be reflected in the high uptake of online surveys.

5.2 Data Analysis

The research data was analyzed using PLS-SEM, a method that combines structural equation modeling and partial least squares. PLS-SEM analysis was performed using the SMARTPLS version 4 program. The investigation was strengthened by the application of the Confirmatory Factor Analysis (CFA) technique. This approach is founded on a solid theoretical framework that has been proven in other research to guarantee the robustness of the model design and latent variable indicators employed in this investigation. One of the crucial steps in the two-stage analysis process that the PLS-SEM approach mandates is examining the outer and inner models. The constructs that were utilized to develop the various survey instrument indicators are evaluated for validity and coherence using a variety of statistical analyses that make up the outer model.

The data received was thoroughly checked and verified by the study experts. In order to assess validity and reliability further, the study filtered the feedback forms. In addition to assessing

discriminant and convergent validity, the study investigated average variance extracted (AVE), the heterotrait-to-monotrait ratio (HTMT) and the variance inflation factor (VIF). Additionally, using composite reliability and Cronbach alpha to measure the reliability of the data. The measurement scale used in this investigation has a reliability that falls within reasonable bounds. The reliability level of 0.7 has been referenced in the literature by (Hair et al., 2017; Ringle et al., 2012). Every item reliability score, according to the current study, is higher than the threshold value.

The findings demonstrated that the internal variables in this investigation are sufficient since Cronbach's Alpha, or Composite Reliability, consistently has a value above 0.70. A sufficient AVE value, according to the literature, ought to be higher than 50%. The results show that the modeling is at a good level because the AVE result is more than the threshold value of 0.5, which establishes the measurement scale acceptability minimum value (Hair et al., 2019). According to (Jarvis et al., 2003), satisfactory findings are produced when the outer VIF statistics derived from the measurement model fall below the cutoff value of 5. There are no significant multicollinearity problems indicated by the VIF results. Thus, the findings for every item were satisfactory. Table 2 displays the AVE values, reliability, outer VIF, and item factor loadings.

Table 2. Validity, Reliability and VIF of Questionnaire

Variabel	Items	Loadin g Factor	Outer VIF
Social (SCP)	Capital CA = 0.846, CR = 0.896, AVE = 0.684.		
	1. I have good ties to support business with government and non-government institutions	0.791	2.493
	2. My company interacts or engages in activities that involve professional or social networks	0.864	2.213
	3. Values and goals are aligned with peers and business partners within the organization or community	0.876	1.918
	4. In social activities or joint activities, the company participates to support or empower others around it.	0.772	2.106
Financial (FCP)	Capital CA = 0.810, CR = 0.886, AVE = 0.722.		
	1. I have invested your financial capital to develop or strengthen my social business	0.864	2.020
	2. I rate the level of financial sustainability of my social business as impactful in the long term	0.843	1.515
	3. My social business has diverse sources of income to reduce financial risk	0.842	2.750
Market Orientation (MAO)	CA = 0.907, CR = 0.935, AVE = 0.782.		
	1. Focus my social business on customer needs and preferences	0.884	2.709

		2. I conduct an analysis of competitors or similar organizations in the market.	0.862	2.811
		3. I often create innovations or adopt new practices in my social business	0.928	2.481
		4. I respond quickly and effectively to market changes	0.860	1.584
Social Entrepreneurship Performance (SES)	CA = 0.832, CR = 0.889, AVE = 0.667.			
		1. My social business achieves its set goals and mission	0.765	2.414
		2. My social business adds value or makes a positive contribution to society	0.850	2.498
		3. Do the innovations implemented in your social business have a positive impact on performance outcomes	0.882	1.730
Economic Empowerment (EMP)	CA = 0.856, CR = 0.912, AVE = 0.777.			
		1. Your social business empowers individuals or communities by creating income opportunities	0.911	2.564
		2. I have helped many people to be empowered or given job opportunities through your social business.	0.885	2.843
		3. Your social business supports skills development and individual capacity building	0.847	1.551
Social Local Communities (SLC)	CA = 0.885, CR = 0.913, AVE = 0.636.			
		1. Your social business contributes to the well-being and happiness of the local community	0.847	2.478
		2. Your social business strives to preserve and promote local cultural heritage	0.758	2.156
		3. your social business takes environmental sustainability into account in its activities and operations	0.766	1.858
		4. your social business supports social inclusion, especially for vulnerable or marginalized groups	0.843	2.682
		5. your social business contributes to education and awareness raising in the local community	0.755	1.797
		6. your social business has an impact on improving the health and well-being of local communities	0.810	2.064

Source: The results of the author's data analysis (2023)

In order to solve the multicollinearity issue, this study assessed the internal VIF value. The

multicollinearity issue between the chosen variables was therefore substantiated by this investigation. Social capital, financial capital, and market orientation to social entrepreneurial performance, economic empowerment, and local social communities are the research factors that have a substantial association, which is explained by the satisfactory internal VIF value. This suggests that the presence of multicollinearity in variables is not significant. Consequently, there is sufficient convergent validity in the model presented in this study. The internal VIF values are reported in Table 3.

Table 3. Internal VIF

Variable	Social Entrepreneurship Performance	Economic Empowerment	Social Impact of Local Communities
Social Capital	1.436	2.181	
Financial Capital	2.456	1.482	
Market Orientation	1.963	2.021	
Social Entrepreneurship Performance		2.212	2.134
Economic Empowerment			1.732

Source: The results of the author's data analysis (2023)

The most popular tests, including the HTMT ratio and the Fornell and Larcker Criteria, are utilized in this study to assess the discriminant validity of the data. According to (Hair et al., 2019), discriminant validity is the presence of poor correlations between indicators of distinct variables, which leads the reader to believe that the indicators measure the same thing. According to (Hair et al., 2019; Joe F Hair Jr et al., 2017; Sarstedt et al., 2017), the pairwise squared correlation indicates that the largest pairwise correlation is located at the top of each column, supporting the Fornell and Larcker criteria's considerable discriminant validity. Table 4 presents the findings of the investigation with respect to the Fornell and Larcker criterion.

Table 4. Fornell and Larcker criterion

	SCP	FCP	MAO	SES	EMP	SLC
Social Capital	0.764					
Financial Capital	0.632	0.782				
Market Orientation	0.287	0.551	0.772			
Social Entrepreneurship Performance	0.439	0.742	0.748	0.812		
Economic Empowerment	0.584	0.493	0.432	0.791	0.789	
Social Impact of Local Communities	0.762	0.532	0.754	0.692	0.374	0.821

Source: The results of the author's data analysis (2023)

The approved data's discriminant validity is supported by the Heterotrait-Monotrait Ratio of Correlations (HTMT) results, which show an appropriate level. The HTMT ratio-based approval requirements state that the final value must be less than 0.85 (Sarstedt et al., 2021). All of the resultant numbers, as indicated by Table 5, are less than the 0.85 criterion. It thus exhibits selective and courteous legitimacy.

Table 5. HMTHT Test

	SCP	FCP	MAO	SES	EMP	SLC

Social Capital	1					
Financial Capital	0.331	1				
Market Orientation	0.532	0.691	1			
Social Entrepreneurship Performance	0.314	0.442	0.431	1		
Economic Empowerment	0.143	0.493	0.547	0.531	1	
Social Impact of Local Communities	0.582	0.232	0.429	0.492	0.441	1

Source: The results of the author's data analysis (2023)

6. RESULTS AND DISCUSSION

6.1. Goodness of Fit Model

In order to confirm the validity of the analysis procedure's output and the variables' dependability, this research study assesses the advantages and high quality of the measurement model using the PLS-SEM approach. In light of earlier arguments, the purpose of this study is to evaluate each item of the chosen variables for discriminant validity, concurrent validity, and reliability in order to test and validate the suggested hypothesis statements (Hair et al., 2019; Joe F Hair Jr et al., 2017; Joseph F Hair Jr et al., 2017; Sarstedt et al., 2021). A component of the research framework, the measurement model examines the connection between the measurements and the research latent variables. The structural equation modeling, or SEM, approach presents a model that quantifies the relationship between the structure or elements supporting the hypothesis and observations (indicators) collected during research surveys and studies. This study uses confirmatory factor analysis (CFA) to assess the efficacy and competency of the suggested measuring model. According to revised research, factor loadings should have a minimum threshold value of 0.7 (Hair et al., 2019). The remaining values for every item in this study satisfied the acceptance criteria for factor loading at 0.05; yet, all parameter values obtained in this study were over the threshold value, with a minimum value of 0.638. The Standardized Root Mean Square Residual for Structural Equation Modeling (SRMR) of the model fit index is 0.07, which is within the acceptable range according to other studies (Hair et al., 2019).

The predictive relevance of the model (Q2) was examined in a subsequent stage of the study, and the findings of the variables that were chosen demonstrated a value greater than zero, confirming the predictive relevance of the model (Sarstedt et al., 2021). The dependent variables, social entrepreneurship performance (Q2 = 0.535), economic empowerment (Q2 = 0.493), and the social effect generated in the local community (Q2 = 0.481), all had respectable Q2 values in the results. The outcomes are noteworthy and pleasing. R2 is a measure of the amount of variance that can be accounted for by the endogenous variables of social entrepreneurship success, economic empowerment, and social impact in the local community. R2 hence aids in assessing the study model's prediction accuracy. The social entrepreneurship performance variable fits the data set with an R2 of 0.643, or 64.3%, the economic empowerment variable fits the data set with 54.7%, and the local community social impact variable fits the data set with 49.4%. As a result, the sample size, the study approach, and the outcomes of every item measured appear appropriate.

6.2. Hypotesis Test

PLS-SEM techniques are commonly used in research to measure path coefficients for standardized beta (β) coefficients (Hair et al., 2019; Joseph F Hair Jr et al., 2017). Any unit of fluctuation, change, or variation in an external variable is equivalent to the probable interpretation or variation that could arise from an endogenous variable. The path coefficients in the analysis results describe these potential variants. Based on the final values of each path, specialists estimate and generalize the conceptual model. In the end, routes with bigger values show a highly significant influence on the endogenous variables in the model, whereas paths

with lower values show a weaker or insufficient influence.

The hypothesis is deemed significant when the t-statistic value, as determined at the 95% confidence level, exceeds the t-statistic (>1.96). The results displayed here were obtained using the SmartPLS bootstrap software. Table 6 presents the construct hypothesis analysis in addition to the beta values, means, standard deviations, t-values, and p-values. Consequently, a choice was made using a p-value of 0.05.

Table 6. Results of Hypothesis Tets

Hypotesis	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T-statistic	p- Values
Social Capital -> Social Entrepreneurship Performance	0.693	0.647	0.242	2.478	0.013
Social Capital -> Economic Empowerment	0.359	0.329	0.140	2.575	0.010
Financial Capital -> Social Entrepreneurship Performance	0.929	0.941	0.161	5.769	0.000
Financial Capital -> Economic Empowerment	0.501	0.492	0.106	4.737	0.000
Market Orientation -> Social Entrepreneurship Performance	0.596	0.624	0.201	2.964	0.003
Market Orientation -> Economic Empowerment	0.823	0.803	0.104	7.947	0.001
Social Entrepreneurship Performance -> Economic Empowerment	0.420	0.425	0.048	7.969	0.000
Social Entrepreneurship Performance -> Social Impact of Local Communities	0.748	0.751	0.046	16.128	0.000
Economic Empowerment -> Social Impact of Local Communities	0.268	0.265	0.052	5.165	0.000

Source: The results of the author's data analysis (2023)

By demonstrating how effectively the variables correlate with one another, hypothesis testing elucidates the link suggested by the research variables. According to (source), researchers can test hypotheses that elucidate the link between the variables chosen for the study by using the PLS-SEM (structural equation modeling) approach (source). The theoretical pathways proposed in this work are depicted in Table 6, which demonstrates the relationship between market orientation, financial capital, and social capital with social entrepreneurship performance, economic empowerment, and local community social impact.

In the first discussion, H1 demonstrates that social capital and social entrepreneurship performance have a statistically significant positive association. P-value of 0.013 indicates that social entrepreneurs with higher social capital likely to do better. This shows that social networks and linkages may have a beneficial impact on the prosperity of social companies in Indonesia. H2 shows a statistically significant positive correlation between economic empowerment and social capital. Economic empowerment within communities is more likely to be attained by social companies with greater social capital. The low p-value of 0.010 suggests that social capital is crucial for the growth of the economy.

With a p-value of 0.000, the financial capital variable in H3 shows a highly significant positive association with social entrepreneurship performance. The fact that social entrepreneurs with more financial resources typically outperform others in their field emphasizes how crucial financial resource availability is to the success of social entrepreneurship. H4 presents statistical evidence of a statistically significant positive correlation between economic empowerment and financial capital. The low p-value of 0.000 suggests that people with more financial resources are linked to better levels of economic empowerment in their communities. Furthermore, with a p-value of 0.003, the final independent variable—where H5 is answered—is approved since it demonstrates a strong positive correlation between market orientation and social entrepreneurship performance. Strong market orientation boosts performance for social entrepreneurs, underscoring the need of comprehending and satisfying market demands for social firms. With a p-value of 0.001, H6 in the statistical results demonstrates a highly significant positive link between market orientation and economic empowerment. This suggests that a market-oriented strategy plays a major role in local communities' economic empowerment.

With a p-value of 0.000, H7 in the results demonstrated a highly significant positive association between social entrepreneurship performance and economic empowerment. Effective social entrepreneurship is critical for economic growth, as evidenced by the higher likelihood that high-performing social firms will support economic empowerment in their communities. Additionally, with a p-value of 0.000 indicating that H8 is genuine, the data demonstrates a highly significant positive association between social entrepreneurship performance and local community social effect. Proficient social entrepreneurs play a crucial role in promoting constructive transformation in their communities by making a substantial beneficial social impact.

At a p-value of 0.000, the final set of results showed a strong positive correlation between social effect in local communities and economic empowerment. The relationship between economic development and social well-being was highlighted by the observation that achieving economic empowerment was linked to beneficial social effects on local communities.

The study's findings offer insightful information about the connections between different variables and results in the context of social entrepreneurship in Indonesia. The findings imply that market orientation, social and financial capital, and other factors are critical to the success of social entrepreneurship and its

7. DISCUSSION

The findings indicated that social capital and social entrepreneurship performance in Indonesia were significantly positively correlated. The success of social companies is significantly influenced by social capital, which encompasses social networks, relationships, and connections. This is consistent with social capital theory, which highlights the importance of social networks in promoting the sharing of information and the mobilization of resources (Bourdieu & Richardson, 1986; Coleman, 1988). Strong social ties improve the performance of social enterprises, as demonstrated by the study's findings, which support this theoretical framework. According to these MSME entrepreneurs, having a strong sense of social capital is crucial for any MSME to succeed. In certain of the provinces examined in this study, there are numerous communities of entrepreneurs who collaborate to establish standards and mutual trust, which in turn affects the entrepreneurship performance of their companies.

The importance of social capital in social entrepreneurship has also been noted by earlier studies. According to (Febrian & Maulina, 2018), social capital has a favorable impact on how well micro and small social enterprises in Indonesia perform. Expanding upon previous research, this study provides a more comprehensive examination that highlights the significance of social capital for the prosperity of social entrepreneurs operating in Indonesia.

Additionally, a strong correlation between social capital and economic empowerment is revealed by the research. This connection emphasizes the role that social networks play in promoting local economic growth. The present discovery aligns with prior investigations that emphasize the significance of social connections and community involvement in promoting economic empowerment (Guliyeva et al., 2021; Hausberg & Korreck, 2020; Yohanes et al., 2017). Social businesses may obtain important resources, win over supporters, and establish trust thanks to social capital, which takes the shape of community networks and relationships. In the end, social capital enhances the financial health of nearby communities. A number of provinces have shown that the existence of these social enterprises is highly beneficial to local economic empowerment and has even formed an entrepreneurial ecosystem. These businesses claim that their primary objective in establishing a social enterprise is to effect social change, particularly in the local communities surrounding their establishment. In the framework of community-based organizations in Indonesia, (Andriani & Christoforou, 2016) investigated the connection between social capital and economic empowerment in a prior study. The results of our investigation support the conclusions of that study and offer more proof of the critical role that social capital plays in promoting economic empowerment in Indonesia's social entrepreneurship scene.

The study's findings firmly support the theory that financial capital improves the performance of social entrepreneurship. Social enterprises may grow their operations, create novel solutions, and make a lasting effect when they have access to capital and financial resources. The impact investing literature and financial capital theory are consistent with this conclusion (Mansour et al., 2018; Widjiansih et al., 2022; Yusuff et al., 2018). The study's findings highlight how crucial impact investors and funding options are to the expansion and prosperity of social enterprises. It's interesting to note that 34% of the social entrepreneurial endeavors in this study were not even connected to loans from financial institutions, but 60% of them had significant money. The significance of financial capital for social companies has also been emphasized by earlier studies. Financial capital was proven to be a key predictor of social enterprise growth and performance (Bontis et al., 2018; Young, 2009). The present study bolsters the previous research outcomes by offering supplementary empirical data that validates the notion that financial resources significantly contribute to the prosperity of social entrepreneurs in Indonesia.

The results of this investigation lend more credence to the theory that financial capital and economic empowerment are positively correlated. This implies that financial resources support both economic development and the improvement of social entrepreneurship performance. The findings are consistent with earlier studies that highlight the contribution of financial capital to the growth of local economies (Acs et al., 2018).

Research on Indonesian microfinance institutions and their effects on economic empowerment was done by (Febriansyah et al., 2020). The results of our study provide support to earlier research by demonstrating that financial resources, regardless of source, can have a major impact on economic empowerment. In an environment where social companies seek to alleviate poverty and economic inequality, the connection between financial resources and economic development is especially pertinent.

The findings indicate a favorable correlation between market orientation and the success of social entrepreneurship. This emphasizes how crucial it is to match business plans with the characteristics of the marketplace. According to the research on market-driven social transformation (Pinheiro et al., 2021; Sari & Kusumawati, 2022), social businesses that modify their strategy to satisfy market needs typically have greater performance. These results underscore the significance of creativity and flexibility in social entrepreneurship. The study's socially conscious business owners stated that their ability to survive to this day is a result of

their constant efforts to adapt, particularly to the demands of the market. Innovation and market orientation are two continuous concepts; in the tabulation of study data, the market orientation variable has the highest degree of strongly agreed responses.

The notion of market orientation has received a lot of attention in the literature on marketing and entrepreneurship. The significance of market-oriented techniques in the context of social entrepreneurship has been studied in the past, for example, by (Pinheiro et al., 2021). By offering empirical proof of the favorable correlation between market orientation and social entrepreneurship performance, our research adds to the body of knowledge by highlighting the significance of adaptability and market-based strategies in social entrepreneurship initiatives. The findings indicate that market-based strategies have the ability to stimulate local economic development since there is a substantial correlation between market orientation and economic empowerment. This finding aligns with other research that highlights the market-oriented tactics used by prosperous social entrepreneurs (Evers & Schulze-Böing, 2001; Glaveli & Geormas, 2018; Karnowati & Handayani, 2022). Market orientation's impact on the growth of social companies in Indonesia. The results of our investigation are consistent with those of other studies, offering more proof that market-oriented tactics can result in economic empowerment. Not only can social companies operate better when they can adjust to changing market conditions, but local communities also benefit economically from this flexibility.

This study demonstrates that the social effect and economic empowerment of local communities are positively impacted by successful social entrepreneurship. Its multifaceted success highlights the wider effects of social entrepreneurship on economic growth and community well-being. Efficient social firms have a higher probability of fostering economic growth and generating favorable social consequences. These results are consistent with the primary objective of social entrepreneurship, which is to alleviate economic inequality and advance the welfare of communities (Konda et al., 2015; Yarima et al., 2023).

Research on the effects of social entrepreneurship in Indonesian communities was done by (Anggahegari et al., 2021; Idris & Hijrah Hati, 2013; Iskandar & Kaltum, 2021; Margiono & Feranita, 2021), who discovered that social enterprises are crucial to enhancing community wellbeing. Our study supports these studies, demonstrating that successful social entrepreneurship can promote both beneficial social effects and economic development. The respondents' answers indicate that social enterprises have a dual role in bringing about social and economic transformation in their communities.

The study's conclusions, taken together, are consistent with a number of important theories and frameworks that offer a conceptual basis for comprehending social entrepreneurship, its bearing on economic empowerment, and its impact on local communities. A thorough comprehension of the research findings is made possible by the Resource-Based View (RBV), Institutional Theory, Social Capital Theory, Triple Bottom Line (TBL) Theory, and Collective Impact framework.

The intricate interplay among these variables in the larger framework of social entrepreneurship in Indonesia is highlighted by the positive correlations between social capital, financial capital, market orientation, social entrepreneurship performance, economic empowerment, and social impact. The results validate the notion that social entrepreneurial endeavors can promote economic empowerment and generate favorable social change, hence aiding in Indonesia's sustainable development, through the efficient utilization of resources and active engagement with external institutions and communities.

8. Limitations

Although this study offers insightful information, it is crucial to recognize its limits. The cross-sectional data employed in this study made it more difficult to determine causes and effects. Longitudinal data might be used in future research to study the temporal correlations between

variables. Additionally, sectoral disparities in social entrepreneurship were not investigated because the focus of this study was on general linkages. Subsequent investigations may go deeper into the subtleties of social entrepreneurship in various Indonesian sectors and areas.

9. Implications

The following stakeholders will be significantly impacted by the study's findings:

1. The significant influence that market orientation, financial capital, and social capital have on the development of social entrepreneurship in Indonesia should be taken into account by policy makers. The growth of social entrepreneurship can be supported by fostering market-based approaches, fostering collaboration, and facilitating financial access.
2. These insights can help social entrepreneurs concentrate on building robust networks, pursue financial sustainability, and modify their plans in response to changing market conditions. Their social effect and entrepreneurial performance may both be enhanced by these actions.
3. It is imperative for investors and funders to acknowledge the significance of financial capital in enabling social enterprises to achieve success and accomplish desired social impact. Sustained assistance in the form of capital and coaching can enable social entrepreneurs to propel constructive transformation.

10. Suggestions and Future Research Directions

To help social entrepreneurs become more proficient in areas like impact measurement, market research, and financial management, policymakers and support organizations can fund capacity building initiatives. It is important to work toward giving social companies better access to funding. This can be accomplished by providing new finance channels, facilitating the connection between impact investors and social entrepreneurs, or setting up local funds specifically designated for social enterprises. In order to match the dynamics of the market with their endeavors, social entrepreneurs should be encouraged to carry out in- depth market research and given training on market-oriented methods. Help social entrepreneurs make connections with possible partners, mentors, and other stakeholders by organizing conferences, workshops, and networking events. Success requires establishing a robust social capital network.

Areas for Future Research

1. Examine more sophisticated impact measuring techniques in the context of Indonesian social entrepreneurship, going beyond conventional measures to capture comprehensive social and economic results.
2. Examine how social entrepreneurs in Indonesia are affecting the environment, especially in areas like waste management, renewable energy, and sustainable agriculture.
3. Analyze government laws and rules that impact social entrepreneurship in detail, evaluate their efficacy, and make recommendations for changes to establish a more encouraging environment.
4. Analyze the impact of gender on social entrepreneurship in Indonesia by investigating the opportunities and difficulties encountered by female social entrepreneurs as well as the role that gender can play in these efforts.

11. CONCLUSION

This study clarifies the complex dynamics of social entrepreneurship in Indonesia and offers insightful information on the variables influencing social impact, economic empowerment, and social entrepreneurial performance. Many important conclusions emerged from the analysis: First, it has been discovered that social

capital—including networks and relationships—strongly influences economic empowerment as well as social entrepreneurial performance. These findings highlight how important social networks are for exchanging knowledge and mobilizing resources in the context of social entrepreneurship. Second, it became clear that financial capital was a key factor in both the success of social entrepreneurship and economic empowerment. The availability of capital and financial resources greatly improved social entrepreneurs' capacity to operate efficiently and support regional economic growth. The study also emphasized the significance of market orientation. Social entrepreneurs that matched their approaches to the needs of the market typically had higher success rates and more successful programs promoting economic empowerment. This emphasizes how important market-driven strategies are for attaining both societal benefit and financial sustainability.

Additionally, the study showed that effective social entrepreneurship performance has a good impact on local communities' social impact as well as economic empowerment. This multifaceted achievement highlights the wider consequences of social entrepreneurship for the prosperity of communities and the growth of the economy.

REFERENCE

1. Abdullah, M. A., & Hoetoro, A. (2020). Social Entrepreneurship as an Instrument to Empowering Small and Medium Enterprises: An Islamic Perspective. *Scholars Journal of Science and Technology*, 1(5), 147– 155.
2. Abidin, Z. Z., Rahayu, S. M., Pamungkas, M. G. W. E. N., Handayani, S. R., & Utami, R. B. (2022). Entrepreneurial Knowledge, Market Orientation, Digitalization, and Entrepreneurial Competencies:
3. Evidence from SMEs in Indonesia. *Journal of Economics, Finance and Accounting Studies*, 4(2), 91–103. <https://doi.org/https://doi.org/10.32996/jefas.2022.4.2.8>
4. Acs, Z. J., Estrin, S., Mickiewicz, T., & Szerb, L. (2018). Entrepreneurship, institutional economics, and economic growth: an ecosystem perspective. *Small Business Economics*, 51, 501–514.
5. Adomdza, G., Michael, A., Lartey, A. S. D., & Lawal, A. Ben. (2017). Exploring the entrepreneurship landscape in Ghana to support a new social enterprise policy.
6. Andriani, L., & Christoforou, A. (2016). Social capital: a roadmap of theoretical and empirical contributions and limitations. *Journal of Economic Issues*, 50(1), 4–22. <https://doi.org/https://doi.org/10.1080/00213624.2016.1147296>
7. Anggadwita, G., Lutarlean, B. S., Ramadani, V., & Ratten, V. (2017). Socio-cultural environments and emerging economy entrepreneurship: Women entrepreneurs in Indonesia. *Journal of Entrepreneurship in Emerging Economies*.
8. Anggahegari, P., Yudoko, G., Rudito, B., Mulyaningsih, H. D., & Ramadani, V. (2021). Female social entrepreneurship in Indonesia: a critical literature review. *International Journal of Entrepreneurship and Small Business*, 43(3), 315–331.
9. Anwar, M. K., Ridlwan, A. A., & Laili, W. N. R. (2023). The Role of Baitul Maal wat Tamwil in Empowering Msmes in Indonesia: a Study of Indonesian Islamic Microfinance Institutions. *International Journal of Professional Business Review*, 8(4), e0913–e0913.
10. Ardiansyah, M. (2016). Development models and strategy of micro, small and medium enterprises (MSMEs) in south Sulawesi. *International Conference, Integrated Microfinance Management for Sustainable Community Development (IMM 2016)*, 175–180.
11. Awaluddin, R., Suhardi, D., & Djuniardi, D. (2021). Company Position Analysis Using Porter's Value Chain and PEST (Political, Economic, Socio-Cultural, Technological) In PT. West Java International Airport. *UNISSET 2020: Proceedings of the 1st*

- Universitas Kuningan International Conference on Social Science, Environment and Technology, UNiSET 2020, 12 December 2020, Kuningan, West Java, Indonesia, 159.
12. Bacq, S., & Eddleston, K. A. (2018). A resource-based view of social entrepreneurship: How stewardship culture benefits scale of social impact. *Journal of Business Ethics*, 152, 589–611.
 13. Baker, W. E., Grinstein, A., & Perin, M. G. (2020). The impact of entrepreneurial orientation on foreign market entry: the roles of marketing program adaptation, cultural distance, and unanticipated events. *Journal of International Entrepreneurship*, 18, 63–91.
 14. Berno, T. (2017). Social enterprise, sustainability and community in post-earthquake Christchurch: Exploring the role of local food systems in building resilience. *Journal of Enterprising Communities: People and ...* <https://doi.org/10.1108/JEC-01-2015-0013>
 15. Bidet, E., & Defourny, J. (2019). Social enterprise in Asia: Theory, models and practice. *library.oapen.org*. <https://library.oapen.org/handle/20.500.12657/58593>
 16. Bobek, V., Schachl, A., & Horvat, T. (2022). The Impact of Women's Economic Empowerment on the Rise of Female Entrepreneurship in Namibia. *Naše Gospodarstvo/Our Economy*, 68(4), 53–65.
 17. Bontis, N., Ciambotti, M., Palazzi, F., & Sgro, F. (2018). Intellectual capital and financial performance in social cooperative enterprises. *Journal of Intellectual Capital*.
 18. Bourdieu, P., & Richardson, J. G. (1986). *Handbook of Theory and Research for the Sociology of Education*. Greenwood Press New York.
 19. Buallay, A., Cummings, R., & Hamdan, A. (2019). Intellectual capital efficiency and bank's performance: A comparative study after the global financial crisis. *Pacific Accounting Review*.
 20. Calvo, S., Lyon, F., Morales, A., & Wade, J. (2020). Educating at scale for sustainable development and social enterprise growth: The impact of online learning and a massive open online course (MOOC). *Sustainability*. <https://www.mdpi.com/2071-1050/12/8/3247>
 21. Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, S95–S120.
 22. Costa, E., & Andreus, M. (2021). Social impact and performance measurement systems in an Italian social enterprise: a participatory action research project. In *Journal of Public Budgeting, Accounting &*
 23. *emerald.com*. <https://doi.org/10.1108/JPBAFM-02-2020-0012>
 24. Dar, I. A., & Mishra, M. (2020). Dimensional impact of social capital on financial performance of SMEs. *The Journal of Entrepreneurship*, 29(1), 38–52.
 25. Diochon, M., & Ghore, Y. (2016). Contextualizing a social enterprise opportunity process in an emerging market. *Social Enterprise Journal*. <https://doi.org/10.1108/SEJ-11-2015-0032>
 26. Diugwu, I. (2011). Knowledge acquisition and sharing: a sustainable source of competitive advantage in supply chains. *Proceedings of the International Conference on Intellectual Capital, Knowledge Management & Organizational Learning*, 157–163.
 27. Eikenberry, A. M., & Kluver, J. D. (2004). The marketization of the nonprofit sector: Civil society at risk? *Public Administration Review*. <https://doi.org/10.1111/j.1540-6210.2004.00355.x>
 28. Estrin, S., Mickiewicz, T., & Stephan, U. (2016). Human capital in social and commercial entrepreneurship. In *Journal of business venturing*. Elsevier. <https://www.sciencedirect.com/science/article/pii/S0883902616300192>

29. Evers, A., & Schulze-Böing, M. (2001). Social enterprises and transitional employment. The Emergence of Social Enterprise. https://books.google.com/books?hl=en&lr=&id=3c2AAgAAQBAJ&oi=fnd&pg=PA120&dq=social+enterprise&ots=iD10oJucqC&sig=hnPDZcV_35gh7wSlsXDZSHUEbkU
30. Fatoki, O. O. (2011). The impact of human, social and financial capital on the performance of small and medium-sized enterprises (SMEs) in South Africa. *Journal of Social Sciences*, 29(3), 193–204.
31. Febrian, A. F., & Maulina, E. (2018). The influence of social capital and financial capability on sustainable competitive advantage through entrepreneurial orientation : Empirical evidence from Small and Medium Industries in Indonesia using PLS-SEM. 5(12), 218–232. <https://doi.org/10.14738/assrj.512.5720>.
32. Febriansyah, H., Watson, C. W., & Gkikas, A. (2020). The entrepreneurial role of Indonesian universities in the economic development of rural communities: in search of empowerment. *Research Handbook on Entrepreneurship in Emerging Economies: A Contextualized Approach*, 160.
33. Frączkiewicz-Wronka, A., & ... (2014). The use of ICT for achieving the objectives of the business model: Social enterprise perspective. *Polish Journal of* <https://www.infona.pl/resource/bwmeta1.element.baztech-ec322abe8dfc-436f-a095-888f15c98590>
34. Fulgence, K., & Mori, N. (2009). Social entrepreneurship in Tanzania: assessment of enabling environment. ... *International Conference on Social Enterprise* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1549849
35. Gah, Y., Nugroho, A. A., & Arif, M. (2020). Factors Influencing Women Entrepreneurship in Rural Area, an Exploratory Study in Nusa Tenggara Timur Province of Indonesia. In ... of the 2nd International Conference on [scitepress.org. https://www.scitepress.org/PublishedPapers/2019/85473/85473.pdf](https://www.scitepress.org/PublishedPapers/2019/85473/85473.pdf)
36. Gigauri, I. (2022). The landscape of social entrepreneurship opportunities in Georgia. *International Journal of Sustainable Entrepreneurship and Corporate Social Responsibility (IJSECSR)*, 7(1), 1–17.
37. Glaveli, N., & Geormas, K. (2018). Doing well and doing good: Exploring how strategic and market orientation impacts social enterprise performance. *International Journal of Entrepreneurial* <https://doi.org/10.1108/IJEBr-04-2017-0132>
38. Guliyeva, A., Rzayeva, U., & Huseynova, R. (2021). INFLUENCE OF INTELLECTUAL CAPITAL ON SME'S EFFICIENCY IN THE TRANSITION ECONOMY OF AZERBAIJAN. *Economic and Social*
39. Development: Book of Proceedings, 493–499.
40. Gunawan, J., & Fraser, K. (2016). Exploring young and green entrepreneurship in Indonesia: An introduction. *Journal of Asian Business Strategy*, 6(9), 185–194.
41. Hair, J. F., Ringle, C. M., & Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modelling (PLS-SEM)*. 2e Edition. SAGE Publications.
42. Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24. <https://doi.org/https://doi.org/10.1108/EBR-11-2018-0203>
43. Hair Jr, Joe F, Matthews, L. M., Matthews, R. L., & Sarstedt, M. (2017). PLS-SEM or CB-SEM: updated guidelines on which method to use. *International Journal of Multivariate Data Analysis*, 1(2), 107–123.
44. Hair Jr, Joseph F, Babin, B. J., & Krey, N. (2017). Covariance-based structural equation

- modeling in the Journal of Advertising: Review and recommendations. *Journal of Advertising*, 46(1), 163–177.
45. Haldar, S. (2019). Towards a conceptual understanding of sustainability-driven entrepreneurship. *Corporate Social Responsibility and Environmental Management*, 26(6), 1157–1170.
46. Hani, U., Rachmania, I. N., Setyaningsih, S., & Putri, R. C. (2012). Patterns of Indonesian women entrepreneurship. *Procedia Economics and Finance*, 4, 274–285.
47. Hausberg, J. P., & Korreck, S. (2020). Business incubators and accelerators: a co-citation analysis-based, systematic literature review. *Journal of Technology Transfer*, 45(1), 151–176. <https://doi.org/10.1007/s10961-018-9651-y>
48. Herlina, H., Disman, D., Sapriya, S., & Supriatna, N. (2021). Factors that influence the formation of Indonesian SMEs' social entrepreneurship: a case study of West Java. *Entrepreneurship and Sustainability Issues*, 9(2), 65.
49. Hermanto, B. (2017). Entrepreneurship ecosystem policy in Indonesia. *Mediterranean Journal of Social Sciences*, 8(1), 110.
50. Hidayat, R., Alliyah, S., & Dewi, N. G. (2022). Financial Inclusion, Intellectual Capital, and MSMEs Performance with Business Age as Moderating Variable.
51. Hota, P. K. (2023). Tracing the intellectual evolution of social entrepreneurship research: Past advances, current trends, and future directions. *Journal of Business Ethics*, 182(3), 637–659.
52. Idris, A., & Hijrah Hati, R. (2013). Social entrepreneurship in Indonesia: Lessons from the past. *Journal of Social Entrepreneurship*, 4(3), 277–301.
53. Iskandar, Y., Joeliaty, J., Kaltum, U., & Hilmiana, H. (2021). Bibliometric Analysis on Social Entrepreneurship Specialized Journals. *Journal: WSEAS TRANSACTIONS ON ENVIRONMENT AND DEVELOPMENT*, 941–951. <https://doi.org/10.37394/232015.2021.17.87>
54. Iskandar, Y., & Kaltum, U. (2021). The Relationship Between Intellectual Capital and Performance of Social Enterprises: A Literature Review. <https://doi.org/https://doi.org/10.36941/ajis-2021-0141>
55. Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, 30(2), 199–218.
56. Javed, A., Yasir, M., & Majid, A. (2019). Is social entrepreneurship a panacea for sustainable enterprise development? *Pakistan Journal of Commerce and Social* <https://www.econstor.eu/handle/10419/196185>
57. Karnowati, N., & Handayani, E. (2022). Mediation role of business performance on entrepreneurship orientation and market orientation to create MSME competitiveness in pandemic times. *International Journal of Research in Business and Social Science* (2147-4478), 11(6), 138–147.
58. Klapper, R. G., & Farber, V. A. (2016). In Alain Gibb's footsteps: Evaluating alternative approaches to sustainable enterprise education (SEE). *The International Journal of Management* <https://www.sciencedirect.com/science/article/pii/S1472811716300556>
59. Konda, I., Starc, J., & Rodica, B. (2015). Social challenges are opportunities for sustainable development: tracing impacts of social entrepreneurship through innovations and value creation. *Economic Themes*, 53(2), 211–229.
60. Lang, L. D., Dong, N. T., Ferreira, J. J. M., Behl, A., & Dao, L. T. (2022). Sustainable agribusiness entrepreneurship during the Covid-19 crisis: the role of social capital. *Management Decision*, 60(9), 2593–2614.

61. Lawhaishy, Z. B., & Othman, A. H. A. (2023). Introducing an Islamic equity-based microfinance models for MSMEs in the State of Libya. *Qualitative Research in Financial Markets*, 15(1), 1–28.
62. Mahato, J., Jha, M. K., & Verma, S. (2022). The role of social capital in developing sustainable micro- entrepreneurship among rural women in India: a theoretical framework. *International Journal of Innovation*, 10(3), 504–526.
63. Mansour, D. M., Sedita, S. R., & Apa, R. (2018). Dynamics of Entrepreneurship in Egypt: Assessing the Entrepreneurial Ecosystem: Can Entrepreneurship Contribute to the Economic Development in Egypt?
64. ... Ecosystem in the Middle East and North https://doi.org/10.1007/978-3-319-75913-5_19
65. Margiono, A., & Feranita, F. (2021). The past, present, and future of social entrepreneurship in Indonesia: A strategy to move the ecosystem forward. *Entrepreneurial Connectivity: Network, Innovation and Strategy Perspectives*, 29–40.
66. Melly, L. T., Cheboi, J., & Keter, C. K. S. (2022). Social Entrepreneurial, Market Orientation, and Learning Orientation: Evidence from Small and Medium-sized Enterprises in Kenya. *The International Journal of Business & Management*, 10(10).
67. Mulgan, G., Tucker, S., Ali, R., & Sanders, B. (2007). Social Innovation: what it is, why it matters, how it can be accelerated. [ictlogy.net. https://ictlogy.net/bibliography/reports/projects.php?idp=4189&lang=es](https://ictlogy.net/bibliography/reports/projects.php?idp=4189&lang=es)
68. Murniati, M., Maski, G., Noor, I., & Ekawaty, M. (2021). Entrepreneurship in the tourism industry: Implication on sustainable economic development. In *Environmental, social, and governance perspectives on economic development in Asia* (Vol. 29, pp. 137–156). Emerald Publishing Limited.
69. Narver, J. C., & Slater, S. F. (1990). The effect of a market orientation on business profitability. *Journal of Marketing*, 54(4), 20–35.
70. Ngeke, N. B. (2016). Performance implications of financial capital availability on the financial literacy– performance nexus in South Africa. *Investment Management and Financial Innovations*, 13, Iss. 2 (contin. 2), 354–362.
71. Nursal, M. F., Rianto, M. R., & Bukhari, E. (2022). The Influence of Market Orientation, Entrepreneurial Orientation, Knowledge Management and Learning Organization on Performance Mediated by Innovation in Culinary SME's in Bekasi. *East Asian Journal of Multidisciplinary Research*, 1(8), 1691– 1702.
72. Octavia, A, & Ali, H. (2017). The model of market orientation, entrepreneurial orientation and business performance of small and medium enterprises. *International Review of Management and ...* <https://dergipark.org.tr/en/pub/irmm/issue/32110/356034?publisher=http-www-cag-edu-tr-ilhan-ozturk>
73. Octavia, Ade, Indrawijaya, S., Sriayudha, Y., & Hasbullah, H. (2020). Impact on E-commerce adoption on entrepreneurial orientation and market orientation in business performance of SMEs. *Asian Economic and Financial Review*, 10(5), 516–525.
74. Osei, C. D., & Zhuang, J. (2020). Rural poverty alleviation strategies and social capital link: The mediation role of women entrepreneurship and social innovation. *Sage Open*, 10(2), 2158244020925504.
75. Ozkazanc-Pan, B., & Clark Muntean, S. (2018). Networking towards (in)equality: Women entrepreneurs in technology. *Gender, Work and Organization*, 25(4), 379–400. <https://doi.org/10.1111/gwao.12225>

76. Pickernell, D., Sullivan, D., Senyard, J., & ... (2007). Social capital and network building for enterprise in rural areas: can festivals and special events contribute? ... and Entrepreneurship <https://eprints.qut.edu.au/13201/>
77. Pinheiro, P., Daniel, A., & Moreira, A. (2021). Social enterprise performance: The role of market and social entrepreneurship orientations. VOLUNTAS: International Journal of <https://doi.org/10.1007/s11266-020-00266-x>
78. Qastharin, A. R. (2016). Business model canvas for social enterprise. In Journal of Business and Economics. researchgate.net. https://www.researchgate.net/profile/Annisa-Qastharin/publication/323393037_Business_Model_Canvas_for_Social_Enterprise/links/5aa8e20d0f7e9b0ea308294a/Business-Model-Canvas-for-Social-Enterprise.pdf
79. Rianto, M. R., Woestho, C., & Fikri, A. W. N. (2022). The Role of Mediating Innovation and Social Media: Market Orientation and Entrepreneurial Orientation on the Performance of MSME's Processed by Sea Products in Labuan Village, Banten. East Asian Journal of Multidisciplinary Research, 1(8), 1703–1714.
80. Ringle, C. M., Sarstedt, M., & Straub, D. W. (2012). Editor's comments: a critical look at the use of PLS- SEM in" MIS Quarterly". MIS Quarterly, iii–xiv.
81. Robinson, J. (2006). Navigating social and institutional barriers to markets: How social entrepreneurs identify and evaluate opportunities. Social Entrepreneurship. https://doi.org/10.1057/9780230625655_7
82. Sabeti, H. (2011). The for-benefit enterprise. In Harvard Business Review. nnphi.org. <https://nnphi.org/wp-content/uploads/2015/08/The-For-Benefit-Enterprise.pdf>
83. Samineni, S. (2018). Social entrepreneurship as a tool for sustainable development of women in rural India leading to economic empowerment. IOSR Journal of Business and Management (IOSR-JBM) e- ISSN.
84. Saputra, A. D., Rahmatia, A., & Muslimah, M. (2021). How far are the benefits of the Islamic Philanthropy and Social Entrepreneurship movement? El-Qish: Journal of Islamic Economics, 1(1), 11– 24.
85. Sari, N. T. P., & Kusumawati, A. (2022). Literature Review : The Efforts To Strengthening of Micro, Small and Medium-Sized Enterprises (MSME) in Indonesia. Asian Journal of Management, Entrepreneurship and Social Science, 2(01 SE-Articles), 98–115.
86. Sarstedt, M., Ringle, C. M., & Hair, J. F. (2017). Treating unobserved heterogeneity in PLS-SEM: A multi-method approach. Partial Least Squares Path Modeling: Basic Concepts, Methodological Issues and Applications, 197–217.
87. Sarstedt, M., Ringle, C. M., & Hair, J. F. (2021). Partial least squares structural equation modeling. In Handbook of market research (pp. 587–632). Springer.
88. Segun, P. A. A. (2021). THE STRATEGIC IMPORTANCE OF ENTREPRENEURSHIP AS A MECHANISM FOR SUSTAINABLE ECONOMIC GROWTH IN NIGERIA. HUMANITIES AND SOCIAL SCIENCES LATVIA, 90.
89. Singgalen, Y. A., Sasongko, G., & Wiloso, P. G. (2019). Social Capital and Livelihood Diversification: Tourism Entrepreneurship in a Remote Area of North Halmahera, Indonesia. Jurnal Kawistara, 9(3), 286–298.
90. Smith, D. H., Smith, D. H., Smith, D. H., Smith, D. H., & ... (2016). Participation in worker cooperatives. The Palgrave Handbook https://doi.org/10.1007/978-1-137-26317-9_21
91. Tabasum, S., & Shaikh, A. R. (2022). Impact of employee engagement on employee performance with mediating effect of employee retention. International Journal of Social Science & Entrepreneurship, 2(2), 1–20.

92. Tabuena, A. C., Necio, S. M. L., Macaspac, K. K., Bernardo, M. P. E., Domingo, D. I., & De Leon, P. D.
93. M. (2022). A Literature Review on Digital Marketing Strategies and Its Impact on Online Business Sellers During the COVID-19 Crisis. *Asian Journal of Management, Entrepreneurship and Social Science*, 2(01), 141–153.
94. Toussaint, E. C. (2016). Incarceration to incorporation: Economic empowerment for returning citizens through social impact bonds. *J. Affordable Hous. & Cmty. Dev. L.*, 25, 61.
95. Utomo, H., Priyanto, S. H., Suharti, L., & Sasongko, G. (2019). Developing social entrepreneurship: A study of community perception in Indonesia. *Entrepreneurship and Sustainability Issues*, 7(1), 233.
96. Widjianingsih, R. W. R., Bakri, B. B. B., & Suratmi, T. S. T. (2022). The Effect of Financial Capital, Labor and Social Capital on The Income of Culinary Smes in Jagakarsa District. *Journal of Entrepreneur and Business*, 1(1), 1–13.
97. Yao, L., & Meng, D. (2022). Role of social capital and financial wellbeing in reaching successful entrepreneurial financial performance: a moderated-mediated model of financial intelligence. *Frontiers in Psychology*, 13, 420.
98. Yarima, Y. A., Gamawa, A. I., & Ladan, A. (2023). SURVEY OF THE IMPACT OF SOCIAL INVESTMENT PROGRAMS TOWARD BENEFICIARIES'ECONOMIC EMPOWERMENT AND ENTREPRENEURSHIP DEVELOPMENT IN NIGERIA.
99. Yohanes, F., Zainul, A., & Kholid, M. M. (2017). The influence of social capital and entrepreneurship orientation on business strategy and performance of micro, small and medium enterprises: A study in Timor Tengah Utara. *Russian Journal of Agricultural and Socio-Economic Sciences*, 72(12), 93–99.
100. Young, D. R. (2009). Alternative perspectives on social enterprise. *Nonprofits and Business*.
https://books.google.com/books?hl=en&lr=&id=XVXCBB4MTRUC&oi=fnd&pg=PA21&dq=social+enterprise&ots=2uTqNpIf9U&sig=FgZfs_pAifzyeyny2bk4zVl51V4
101. Yusuff, Y. Z., Bakar, A. A., & Ahmad, S. (2018). Relationship Between Financial Capital, Social Capital and Women Entrepreneurs' Business Performance. *Advanced Science Letters*, 24(1), 202–204.
102. Zid, M., Alkhudri, A. T., Casmana, A. R., Marini, A., & Wahyudi, A. (2020). Ex migrant workers of international women and social entrepreneurship: study at kenanga village in Indramayu Regency in West Java Province in Indonesia. *International Journal of Advanced Science and Technology*, 29(06), 1855–1861.

Exoticism and Global Myth: A Semiotic Analysis of International Hotel Franchises in Bali

1. Ronald Maraden Parlindungan Silalahi (ronald.maraden@upj.ac.id /
2. Clara Evi C. Citraningtyas (clara.citraningtyas@upj.ac.id)
3. Raden Roro Gisela Anindita Maheswari (anin1706@gmail.com)

Abstract

Over the past several decades, Bali has emerged as a popular tourist destination, celebrated for its vibrant culture, stunning natural scenery, and distinctive spiritual traditions. Its exotic reputation has been carefully crafted to attract visitors from around the world. This cultural appeal has attracted numerous foreign investors to Bali's hospitality sector. International hotel chains blend their global branding strategies with local Balinese cultural elements, such as traditional symbols, architecture, and rituals, creating a unique identity that resonates with both residents and international tourists. This study adopts a qualitative method to examine the marketing communication and cultural practices of various international hotel franchises in Bali, using Oswald's Marketing Semiotics framework. The research aims to identify how exoticism and global myths are constructed and communicated through advertisements of five international hotel brands operating in Bali. These brands come from different countries, providing insight into the global development and commercialization of myths. The study specifically focuses on video commercials, which are used as key tools for delivering cultural narratives to a worldwide audience. While research on Bali's tourism and cultural commodification is increasing, few studies apply semiotic analysis to investigate how international hotel chains use marketing to convey global myths. This gap highlights the limited understanding of how global branding intersects with local culture. This research addresses that gap by contributing both theoretical and practical knowledge, expanding semiotic approaches in global marketing, and offering hotel marketers valuable guidance on crafting culturally relevant messages that appeal locally and globally. The findings reveal that international hotel chains do more than just use Balinese culture as decoration; they incorporate it into their strategic storytelling and mythmaking to strengthen their global brand image. This study thus provides important insights for the tourism industry.

Keywords:

Marketing Semiotics, Mythology, Exoticism, Hotel, Global Franchise, Bali

1. INTRODUCTION

Bali has earned global recognition as a premier tourist destination, celebrated for its rich cultural heritage, stunning landscapes, and spiritual traditions. This reputation is not accidental but results from a deliberate image construction through tourism campaigns, media portrayals, and cultural commodification. These efforts have crafted Bali's identity as an exotic, mystical island, attracting global tourists seeking both leisure and cultural depth (Pedersen, 2022; Verheijen & Darma, 2019). This strong international appeal has led to substantial foreign investment, especially in the hospitality sector. Global hotel chains view Bali as a lucrative market, drawn by its steady flow of international visitors. While such investments stimulate the economy, they also spark debates about cultural authenticity and the marginalization of local communities (Sunaryo & Ardhana, 2023). Hotels, as key players in Bali's tourism economy, have become spaces where global and local identities intersect.

To meet visitors' expectations for authenticity, many international hotels incorporate Balinese architectural styles, cultural motifs, and traditional values into their designs and branding (Utama & Krismawintari, 2025). This practice reflects “glocalization,” where global businesses adapt to local cultures to enhance appeal. Thus, Bali becomes not just a physical destination but a symbolic landscape where meanings are continuously negotiated, shaped, and marketed (Widhiastuty & Murdana, 2024).

The hospitality industry in Bali exemplifies how global branding strategies integrate local cultural elements to construct a hybrid identity. Hotels embed Balinese aesthetics—such as temples, statues, and ritual symbols—into their architecture and services, not merely as decoration but as a marketing tool to evoke cultural authenticity (Indrianto, 2016). This fusion creates distinctive brand identities that stand out in a competitive global market (Addis et al., 2024).

For tourists, this hybrid experience enriches the perceived authenticity of their stay. For locals, it offers symbolic recognition, though often within a commercial framework. While this blending enhances the tourist experience, it also raises concerns about the commodification of culture and the potential dilution of traditional practices for economic gain. The balance between cultural preservation and commercial adaptation becomes a central tension within Bali’s tourism landscape (Picard, 1995).

This study investigates how exoticism and global myth are constructed in the promotional narratives of international hotel brands in Bali. These representations play a crucial role in shaping the island’s global image as a serene, culturally rich, and spiritually resonant destination. The research focuses on promotional video advertisements from five major international hotel chains operating in Bali.

Using Laura Oswald’s framework of *Marketing Semiotics* (2012), the study analyzes how visual and linguistic signs in hotel advertisements create symbolic meaning. Marketing semiotics enables the decoding of how brands employ cultural references—such as traditional Balinese rituals, architecture, and landscapes—to position Bali as a dreamlike, luxury escape within the global imagination. Through this analytical lens, the study reveals how these advertisements blend local cultural symbols with universal themes of tranquility, refinement, and escape. The result is a mythologized representation of Bali that appeals to cosmopolitan audiences while reinforcing certain cultural stereotypes. The hotels do not merely promote services; they sell an imagined version of Bali—one that is spiritual, peaceful, and timeless.

By examining how global hotel brands use cultural signifiers to construct place identity, the research contributes to a broader understanding of tourism marketing’s role in shaping cultural narratives. It sheds light on the power dynamics between global commerce and local tradition, and how branding practices can influence both tourist expectations and the lived realities of host communities.

2. LITERATURE REVIEW

Marketing semiotics is an analytical approach that examines signs and symbols within the context of marketing to understand how meaning is constructed and communicated to consumers (Oswald, 2015 & Adami, 2016). Drawing from the principles of semiotics, this strategy aims to uncover the underlying layers of meaning embedded in brand messages, thereby enabling companies to craft more effective and meaningful

communication (Dydyński & Mäekivi, 2018). The primary function of marketing semiotics is to identify and interpret how visual, verbal, and contextual elements shape consumer perceptions of

particular products or brands. In practice, marketing semiotics also explores mythology—collective narratives or cultural beliefs that exist in the minds of consumers—to create strong and relevant symbolic associations. This is closely linked to the development of brand identity, which refers to the internally constructed essence of a brand, and brand image, which denotes how the brand is perceived and interpreted by the public (Kapferer, 1994; Keller, 1993; Aaker, 1991; Kotler & Keller, 2012). Through this approach, companies can develop brand narratives that are not only visually appealing but also culturally and emotionally resonant.

3. METHODOLOGY

This research adopts a qualitative methodology within an interpretive framework to investigate how elements of exoticism and global myths are portrayed in international hotel advertisements in Bali. The study focuses on video advertisements from 12 global hotel brands; however, only 8 of these incorporate Balinese cultural elements and are thus included in the analysis. Applying Laura Oswald's Marketing Semiotics theory (Oswald, 2015), the research explores how cultural symbols are employed to construct brand narratives emphasizing authenticity, luxury, and spiritual retreat. The analysis proceeds through three stages: identifying prominent signifiers, interpreting their symbolic significance, and examining how these contribute to broader cultural mythologies. Consistent with the interpretive paradigm, the study highlights the socially constructed nature of meaning within marketing discourse. It demonstrates how global branding strategies use local cultural elements to shape, commodify, and disseminate representations of Bali's identity—appealing to global audiences while simultaneously reinforcing exotic cultural stereotypes.

4. FINDINGS AND DISCUSSION

Many international hotel advertisements in Bali highlight cultural elements, but not all video campaigns focus on Balinese traditions. Some global hotel brands prioritize promoting modern facilities and exclusive services over local cultural references. These ads showcase unique offerings such as luxury wedding venues, personalized guest experiences, and signature programs that set them apart from competitors. The emphasis lies in delivering comfort, innovation, and convenience to a global audience. This strategy aligns with a broader branding approach that positions the hotel as a premium destination through service excellence rather than cultural immersion. Consequently, Bali's cultural identity becomes secondary, with the spotlight shifting toward the hotel's distinctive advantages designed to attract international travelers seeking refined and memorable experiences.

Exotic Bali as a Sacred Paradise

International hotel in Bali employs a semiotic strategy that heavily draws on the discourse of exoticism to construct Bali as a mythical and transformative destination. In the promotional advertisement, Intercontinental hotel in Bali narrates:

They say beauty is in the eye of the beholder. Imagine seeing things through the eyes of royalty—stepping into a divine paradise where magical nature and sophisticated design

become one. Experience world-class service with attention to the details that matter most. Wake up to the sound of waves that will leave you feeling transformed, and be immersed in a realm of happiness as you enter a new dimension of tranquility. Rejuvenate your senses as you relax and unwind in enchanting solitude. Discover sacred moments of connection with your loved ones, with

yourself, with your soul. Treat yourself to authentic flavours and rediscover the romance of travel as you journey through breathtaking beauty. Here lies a trail of beautiful moments you'll want to revisit time and time again.

Intercontinental Bali Resort (https://youtu.be/UBowYJFHhJs?si=8FmFq_oHPx6-o47y)

Phrases such as “divine paradise,” “realm of happiness,” “enchanted solitude,” and “a new dimension of tranquility” function as signifiers that evoke not only physical beauty but also a spiritual otherness often associated with the exotic. These signifiers frame Bali as a place outside of the ordinary, where magical nature and sophisticated design seamlessly merge. From a Marketing semiotic perspective, the narrator’s choice of words reflects a deliberate construction of meaning. These meanings represent the symbolic messages that the hotel brand intends to convey to potential consumers. Through these linguistic signs, the brand communicates not only information but also shapes perceptions and imaginations about the experience it offers.

Table 1. Connotation

Phrase	Connotation
Royalty & divine paradise	Royalty & divine paradise evoke Bali as a luxurious, sacred escape. Imagery of temples and rituals presents the island as a spiritual haven for elite travelers seeking transcendence.
Magical nature & sophisticated design	Magical nature & sophisticated design frame Bali as a balance between untouched nature and curated elegance. Lush landscapes paired with modern villas symbolize exotic authenticity made accessible.
World-class service & details that matter	World-class service & details that matter repackage Balinese hospitality into luxury experiences—personalized services and refined attention to detail meet global standards while drawing on local warmth.
Sound of waves, solitude, tranquility	Sound of waves, solitude, tranquility construct Bali as a space of healing. Natural elements and spiritual practices are positioned as tools for inner peace and renewal.
Authentic flavours & romance of travel	Authentic flavours & romance of travel promote Bali as a place of cultural immersion. Local cuisine and rituals become emotional touchpoints in a curated journey of self-discovery.

In the advertisement, the consumer is portrayed by a foreign (Caucasian) model, representing a prospective guest who does not originate from the Balinese cultural context. This clearly indicates that the hotel’s promotional message targets international audiences. Within the framework of Oswald’s marketing semiotics (Oswald, 2015), such visual and linguistic signs contribute to the construction of a cultural myth in which Bali is framed not merely as a tourist destination, but as a sacred space that promises emotional and sensory

rejuvenation. Elements such as imagery of royalty, authentic local flavours, and sacred connections reinforce a narrative that commodifies Balinese culture as luxurious and mystical—hallmarks of exoticism. By utilizing such strategies, the brand constructs an idealized and romanticized version of the East, appealing to global elites seeking an escape from modern life. Thus, exoticism becomes central to the brand’s identity, shaping Bali as both an aesthetic and emotional experience designed for consumption.

Bali is portrayed as a dream destination for international tourists, primarily due to its hospitality and rich cultural heritage. This representation offers a distinctive experience compared to other common tourist destinations, positioning Bali as an exotic space ideal for emotional and spiritual escape. This is reflected in the narrative from Conrad Hotel's promotional video, which states:

... Check in to a slice of paradise at Conrad Bali, where every moment feels like a dream and unparalleled hospitality greets you from the moment you arrive. [...]Discover the art of Balinese calligraphy through a hands-on experience with a master of Aksara script writing
Conrad Bali (https://youtu.be/-ZaUikpI23I?si=WTVS_IOOwTTIUz7I)

The phrase “Check in to a slice of paradise at Conrad Bali, where every moment feels like a dream and unparalleled hospitality greets you from the moment you arrive” constructs a powerful semiotic narrative that blends luxury, emotional appeal, and exoticism. The expression “slice of paradise” positions Bali as a mythic, idyllic destination, evoking connotations of escape, purity, and natural beauty that resonate with Western fantasies of the tropical Other. By stating that “every moment feels like a dream,” the text elevates the guest experience to a surreal and euphoric level, suggesting that time spent at the resort transcends ordinary reality. Meanwhile, “unparalleled hospitality” signals global standards of perfection and personal care, aligning the brand with elite service expectations. Together, these signifiers encode a cultural myth in which Conrad Bali becomes not just a hotel, but a gateway to emotional rejuvenation and aesthetic pleasure—offering an exotic yet controlled experience curated for international travelers.

In international hotel promotional videos set in Bali, exoticism is constructed through rich non-verbal cultural elements. Features such as traditional Balinese architecture, ceremonial attire, Kecak dance, gamelan music, and Hindu temples and purification sites serve as semiotic signs that portray Bali as sacred, artistic, and spiritually magical. Within Oswald's (2015) framework of marketing semiotics, these signs do more than represent local culture—they construct a mythical destination narrative, positioning Bali as a transcendent space for emotional escape and personal renewal.

Table 2. Non-Verbal Elements

Non-Verbal Element	Interpretation
Balinese Architecture	Local cultural heritage, authentic and spiritual aesthetics
Traditional Balinese Attire	Unique ethnic identity and visual beauty
Kecak Dance	Spirituality, mythology, and Balinese artistic expression
Gamelan Traditional Music	Sacred and meditative atmosphere typical of rituals
Traditional Spa and Massage	Physical and spiritual healing, inner retreat
Hindu Temple / Tanah Lot	Sacredness, religious grandeur, and spiritual exoticism
Balinese Statue	Representation of deities and sacred artistic aesthetics
Purification Site (Petirtaan)	Ritual cleansing, connection to nature and spiritual renewal
Balinese Chime	Sacred, soothing sound, symbolizing serenity
Local Souvenirs	Authentic cultural products, tangible memory
Western and European Cuisine	Modern comfort, global appeal for elite tourists
Traditional Farming Scene	Local lifestyle in harmony with nature

The visual juxtaposition of traditional spas, Balinese chimes, sacred statues, and the fusion of

Western and local cuisines reinforces the narrative of “where luxury meets heritage.” Exoticism thus becomes central to the brand identity—framing Bali as both an aesthetic and spiritual experience curated for global elites. By combining symbolic, iconic, and indexical signs, the advertisement shapes a visual experience that aligns the hotel’s brand with authenticity, exclusivity, and emotional luxury.

Mythology as a Marketing Strategy

This study show how mythology functions as a strategic tool in marketing by examining the shifting meanings of signs within promotional materials. Referencing Barthes' concept (Barthes, 1957) of mythology as a “second-order semiological system,” the research highlights how signs acquire connotative meanings that go beyond their initial, literal definitions. A hotel, for instance, is traditionally understood as a place for temporary lodging. However, in the context of marketing—especially within high-end tourism—this meaning shifts. The hotel is reimagined not merely as accommodation but as a curated experience rich with cultural and emotional value. This process is not arbitrary; it is shaped by the social and cultural contexts in which the symbol appears. Specifically, this research emphasizes the use of Balinese exoticism as a mythologized layer in the branding of international hotels in Bali. Through both verbal descriptions and visual imagery, the advertising constructs a new narrative around the hotel, positioning it within an imagined realm of luxury, spirituality, and exclusivity. Bali is not simply presented as a travel destination, but as a sacred paradise—a place imbued with mysticism and beauty that appeals to Western fantasies of the “exotic other.” This constructed myth serves marketing objectives by transforming the hotel into a symbol of escape and transcendence.

The exotic imagery of Bali used in these advertisements—temples, traditional dances, sacred rituals, and tropical landscapes—does more than decorate the promotional material. It functions to evoke a sense of awe and desire, suggesting that the experience offered by these hotels is unattainable anywhere else in the world. This exclusivity is central to the appeal; the myth of Bali becomes a differentiator in a highly competitive hospitality market. As such, the shifting connotative meaning of “hotel” reflects broader marketing efforts to embed emotional and cultural depth into the brand identity.

Ultimately, the study reveals that meaning is not static but evolves alongside consumer perception and cultural trends. By embedding mythology into branding strategies, marketers can cultivate a deeper emotional connection with consumers, encouraging them to associate the product—not just with service—but with aspiration, narrative, and myth. This strategic semiotic manipulation becomes key in elevating the brand from a commodity to a culturally resonant experience.

Mythology and Cultural Fusion in Hospitality Marketing

This study also explores how the mythologization of Bali is not merely about emphasizing local traditions but involves a deliberate blending of Eastern and Western cultural elements. The exoticism of Bali, as presented in international hotel advertisements, is often infused with Western cultural aesthetics to appeal to global audiences—particularly those from the West. This cultural fusion is not incidental but serves as a strategic device in constructing a layered and hybrid brand narrative.

One prominent example is the representation of Western cuisine served in a Balinese context. In many promotional materials, Western-style meals are presented using local settings and traditional Balinese aesthetics. Servers may wear traditional attire, and the food—although Western in origin is contextualized within a Balinese dining experience. This creates a symbolic harmony between

the familiar (Western food) and the exotic (Eastern setting), making the experience more accessible yet still culturally rich for foreign guests.

A particularly illustrative case involves the depiction of Western wedding rituals within a Balinese ceremonial setting. Some luxury hotel advertisements showcase Western brides and grooms partaking in wedding ceremonies that incorporate Balinese rituals, architecture, and symbolic items. Here, Western symbols of romance and celebration are intertwined with Balinese spiritual and cultural motifs, creating a hybrid narrative that simultaneously signifies modern luxury and traditional depth.

This intentional mixing of cultural codes illustrates how Western hospitality practices are not merely inserted into the Balinese context but are aesthetically merged with local traditions to form a distinct mythological space. The hotel is no longer seen merely as a temporary lodging; it becomes a curated cultural experience—an imaginative and symbolic space where East meets West in harmony.

Such fusion challenges the static, denotative meaning of a hotel and repositions it as a site of cultural performance. This myth-making strategy not only enhances the perceived value of the hotel brand but also appeals to aspirational desires among global consumers who seek both authenticity and luxury in their travel experiences. In this way, marketing semiotics helps illuminate how brand narratives evolve by embedding symbolic richness through cultural fusion and mythological storytelling.

5. CONCLUSION

This study demonstrates how international hotel advertisements in Bali strategically employ marketing semiotics to construct and commodify the myth of exotic Bali. By analyzing both verbal and non-verbal elements, the research reveals that exoticism is not merely an aesthetic choice, but a central narrative strategy that positions Bali as a sacred, spiritual, and emotionally transformative destination. Through evocative language, symbolic imagery, and cultural performances, brands reframe hotels not simply as accommodations but as curated experiences imbued with luxury, authenticity, and spiritual depth. Drawing on Barthes' theory of mythology, the study shows how connotative meanings transform conventional symbols—such as hotels, temples, or cuisine—into emotionally charged signs that serve marketing objectives.

Furthermore, the integration of Western elements within Balinese cultural contexts—such as the presentation of European cuisine in local settings or the fusion of Western wedding rituals with Balinese ceremonial aesthetics—reflects a deliberate cultural hybridization. This blending creates a unique narrative space where East meets West, enhancing the global appeal of the brand while maintaining a sense of authenticity and local charm. Ultimately, these semiotic strategies illustrate how myth and cultural fusion are instrumental in crafting aspirational brand identities. They not only reflect evolving consumer desires for meaningful, immersive travel experiences but also reveal how hospitality marketing constructs new cultural imaginaries through symbolic manipulation. As such, the study contributes to broader discussions on branding, cultural representation, and the role of semiotics in shaping global tourism narratives.

REFERENCES

1. Aaker, D. A. (1991). *Managing brand equity: Capitalizing on the value of a brand name*. Free Press.
2. Adami, E. (2016). Introducing multimodality. In O. García, N. Flores, & M. Spotti (Eds.), *The Oxford handbook of language and society*. Oxford University Press.
3. Addis, G., Nasino, S., Massi, M., & Camilleri, M. A. (2024). The artification of hospitality: Art hotels as the new frontier of sustainable cultural tourism. *In Tourism planning and destination marketing* (2nd ed., pp. 175–195). Emerald Publishing Limited. <https://doi.org/10.1108/978-1-80455-888-120241009>
4. Barthes, R. (1957). *Mythologies*. Paris: Seuil.
5. Dydyński, J., & Mäekivi, N. (2018). Multisensory perception of cuteness in mascots and zoo animals. *International Journal of Marketing Semiotics*, 6(1). https://ijms.files.wordpress.com/2018/09/ijms-vol-vi-2018_dydyński.pdf
6. Indrianto, A. T. (2016). The commodification of culture in Bali in the frame of cultural tourism. *ASEAN Journal on Hospitality and Tourism*, 4(2), 151–165. <https://journals.itb.ac.id/index.php/ajht/article/view/3031>
7. Kapferer, J. N. (1994). *Strategic brand management: New approaches to creating and evaluating brand equity*. Free Press.
8. Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing*, 57(1), 1–22. <https://doi.org/10.1177/002224299305700101>
9. Kotler, P., & Keller, K. L. (2012). *Marketing management* (14th ed.). Prentice Hall.
10. Oswald, L. R. (2012). *Marketing semiotics: Signs, strategies, and brand value*. Oxford University Press.
11. Pedersen, D. J. (2022). *Pandemic in paradise: A critical political economy analysis of Bali's tourism sector* [Master's thesis, University of Wollongong]. <https://ro.uow.edu.au/ndownloader/files/50384184/1>
12. Picard, M. (1995). Cultural heritage and tourist capital: Cultural tourism in Bali. In M. Lanfant, J. B. Allcock, & E. M. Bruner (Eds.), *International tourism: Identity and change* (pp. 44–66). SAGE Publications Ltd. <https://doi.org/10.4135/9781446250402.n3>
13. Sunaryo, F. D. S., & Ardhana, I. K. (2023). Becoming a global village: The history of globalisation in Bali. *Paramita: Historical Studies Journal*, 35(1), 59–66. <https://journal.unnes.ac.id/journals/paramita/article/download/23715/3682/72091>
14. Utama, I. G. B. R., & Krismawintari, N. P. D. (2025). Revitalizing Bali: Exploring the growth of health and wellness tourism. *International Journal of Current Science Research and Review*, 8(3), 1405–1417. <https://journal.unnes.ac.id/journals/paramita/article/download/23715/3682/72091>
15. Verheijen, B., & Darma Putra, I. N. (2019). Balinese cultural identity and global tourism: The Garuda Wisnu Kencana Cultural Park. *Asian Ethnicity*, 21(3), 425–442. <https://doi.org/10.1080/14631369.2019.1641691>
16. Widhiastuty, N. L. P. S., & Murdana, I. K. (2024). The impacts of Bali cultural tourism to the local Balinese community. *LACULTOUR: Journal of Language and Cultural Tourism*, 3(1), 45–51. <https://ejournal.ppb.ac.id/index.php/lacultour/article/download/1360/915/5381>

Between Duty and Dreams: Indonesian Millennials and Gen Z Perspectives on the Sandwich Generation in Home Sweet Loan

Kheren Immanuel¹, Clara Evi C. Citraningtyas² Universitas Pembangunan Jaya, Indonesia

ABSTRACT

Home Sweet Loan is an Indonesian family drama film that portrays the struggles of the sandwich generation, individuals caught between supporting their parents and siblings while pursuing personal aspirations, such as owning a dream home. This study aims to examine how members of the sandwich generation interpret their representation in Home Sweet Loan, using Stuart Hall's reception theory as an analytical framework. Employing a descriptive qualitative method within the constructivist paradigm, this research utilizes reception analysis to explore audience interpretations. The analysis draws on Hall's encoding/decoding model and incorporates supporting concepts including the sandwich generation, family drama, new media, film studies, and film as a form of social representation. The central research question is: How does the sandwich generation position the meaning of their representation in the film Home Sweet Loan based on Stuart Hall's reception theory? Data were collected through interviews, observations, and documentation, and subsequently analyzed using open, axial, and selective coding.

Keywords: Sandwich Generation, Audience Interpretation, Film, Stuart Hall's Reception Theory.

Research on Consumer Purchasing Behavior: A Study of the apparel Industry in Ho Chi Minh City

**Dr. Tran Thi Ninh, Dr. Nguyen Thi Tuyet Nga, MBA. Do Thi Anh Phuong Hong Bang
International University**

Corresponding Author: Dr. Tran Thi Ninh, Email: nhinhtt@hiu.vn

ABSTRACT

This research explores consumer buying behavior pattern in Ho Chi Minh City's apparel industry through a survey involving 719 consumers. The objective is to identify key relationships influencing consumer purchasing behavior within this market. The findings offer valuable insights for managers of Vietnamese enterprises in the apparel industry, enabling a deeper understanding of consumer behavior and the relationships that shape it. Based on these insights, businesses can develop and implement effective strategies to boost consumer purchasing behavior in the apparel industry.

Keywords: Consumer purchasing behavior, coupons, sales promotion, discount, brand image, the apparel industry, Vietnam.

1. INTRODUCTION

Vietnam is currently one of the top five apparel-exporting countries in the world, with export turnover ranking second among industries in Vietnam, while also creating numerous employment opportunities [1]. In the context of rapidly developing global business, enterprises face fierce competition to attract an increasing number of customers. This competition is also clearly reflected in meeting market demands. A study analyzing the relationship between consumer purchasing behavior and the apparel industry shows that marketing strategies in the sector need to focus on the desires and capabilities of the younger generation, as this group tends to embrace modern fashion styles [2]. An ethical approach plays a key role in defining the relationship between consumer behavior and promotional activities in the apparel industry, as this sector must comply with cultural and traditional norms along with specific cultural factors in the mechanisms for promoting products in the global market [3]. The topic of consumer purchasing behavior has been the focus of many previous studies aiming to identify influencing factors and determine the impact relationships within the apparel industry. The current research continues to expand and clarify these relationships to gain a deeper understanding of customer shopping behavior in the apparel sector.

2. THEORETICAL FRAMEWORK AND RESEARCH MODEL

2.1. Theoretical Background

The Theory of Planned Behavior, proposed by [4], is a theoretical framework that analyzes consumer behavior from both psychological and social perspectives. According to this theory, three main factors—Attitude, Subjective Norms, and Perceived Behavioral Control—play a key role in explaining an individual's intention as well as subsequent actions. Based on this model, Ajzen [4] also noted that factors such as coupons, discounts, and promotional programs influence brand image and customers' purchasing decisions.

2.2. Research Concepts

Consumer purchasing behavior (CPB) is defined as the entire set of decisions related to the selection, use, and disposal of goods and services, as well as the timing and ideas of individuals or

groups making decisions [5].

Brand image (BI) is understood as brand perception, expressed through the associations that consumers hold in their memory [6].

Sales promotion (SP) is the application of various temporary incentive strategies aimed at encouraging customers to quickly purchase certain types of goods or services. Common forms of promotion include coupons, direct discounts, contests, or sweepstakes [7]. According to Kotler et al [8] sales promotion is considered a set of short-term incentive tools designed to encourage customers or distributors to buy faster and in larger quantities of a specific brand, product, or service.

A coupon is understood as a sales tool that allows customers to purchase products at a discounted price when they possess the coupon or buy goods and services at a lower price. It is one of the common techniques used by businesses to stimulate consumer demand by issuing tickets or receipts that encourage customers to buy products or services in the future at a preferential price [9].

According to Fill [10], a "discount" is a tactic of offering goods or services at a lower price to consumers, commonly applied in supermarkets. Khan et al [11] explain that discounts are promotional techniques aimed at boosting sales by reducing the prices of products and services.

2.3. Research Hypotheses

Agmeka et al [12] pointed out that sales promotion plays a key role in building brand image. In fact, promotional activities need to be closely linked and effectively support brand identity. At the same time, sales promotion should be designed so that consumers can easily receive and clearly visualize the company.

H1: SP has a positive influence on BI.

Consumers often prefer purchasing branded products in the market [13]. Tekin et al [14] analyzed the influence of brand image on luxury buying behavior among Turkish consumers. Chovanova et al [15] explored how branding affects consumer behavior. Durrani et al [16] studied the impact of brand image on teenage consumers' purchasing decisions. Riaz [17] examined the relationship between brand image and purchasing behavior in the Pakistani apparel industry, finding a significant positive correlation.

H2: BI has a positive influence on CPB.

The study by Ahmad et al [18] indicated a significant relationship between attitudes toward promotional tools such as price discounts, coupons, and buy-one-get-one-free programs and purchasing behavior. These findings were also confirmed by Shahzad et al [19], who discovered a strong positive correlation between discounts and coupons with consumer behavior. Therefore, it can be affirmed that promotional forms such as coupons and discounts play a positive role in encouraging customers' purchase decisions.

H3: Coupons have a positive influence on CPB. H4: Discounts have a positive influence on CPB.

The mediating role of brand image in the relationship between sales promotion and consumer purchasing behavior.

Based on Ajzen's Theory of Planned Behavior [4], customers' perceptions of sales promotion activities play an important role in enhancing brand awareness as well as stimulating consumer purchasing behavior. Helmi et al [20] also pointed out that sales promotion significantly affects brand image and customers' purchase decisions, while brand image positively and clearly influences their product choices. Based on these arguments, the research team proposes the following hypothesis:

H5: BI mediates the relationship between SP and CPB.

3. RESEARCH METHODOLOGY

The initial questionnaire was developed in English based on previous studies and underwent a peer review process by sixteen experienced experts in the fields of sales and marketing to refine its content. Subsequently, a large-scale online survey was conducted

targeting customers of apparel industry enterprises in Ho Chi Minh City, Vietnam. Prior to the survey, the author sent invitation letters explaining the research objectives and assuring participants of confidentiality. The survey was expected to take approximately 15 minutes to complete. A total of 750 questionnaires were distributed, with 735 returned; after screening, 16 questionnaires were excluded due to excessive missing responses, leaving 719 valid questionnaires for statistical analysis, corresponding to a response rate of approximately 95.87%. According to Hair et al [21], this sample size exceeds the minimum requirement for conducting Partial Least Squares Structural Equation Modeling (PLS-SEM) (Table 1).

The study used a five-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree), to evaluate all items. The coupon measurement scale, based on Pacheco and Rahman [22], consists of three observed variables. Meanwhile, the sales promotion scale from Sharma and Avasthi [23] includes six measurement items. Similarly, the brand image scale by Weiss et al [24] is constructed with six observed variables. Consumer purchasing behavior was measured by Valaei and Nikhashemi [2] through six factors, while the discount scale from Dibb et al [25] uses four observed variables.

4. RESULTS AND DISCUSSION

The results show that all Cronbach's Alpha coefficients exceed the recommended threshold of 0.7, and the composite reliability values also surpass the same benchmark [21]. The AVE value is considered meaningful when it is greater than 0.5 [26]. Thus, each construct in the study demonstrates high reliability and strong validity (Table 2).

The results show that all VIF values are below the threshold of 5, indicating that there is no multicollinearity among the variables [21]. At the same time, the Average Variance Extracted (AVE) of all constructs exceeds the recommended benchmark of 0.5, demonstrating satisfactory convergent validity [26] (Table 3).

Based on the cross-loading analysis results in Tables 3 and 4, the factor loadings of the measurement variables are all higher than the cross-loadings of other research constructs, indicating that the convergent validity meets the required standards [21].

The research results show a strong positive relationship between SP and BI ($\beta = 0.606$; $p < 0.05$), thus hypothesis H1 is accepted. Previous studies have also confirmed the important role of SP activities in building BI [12].

The research results show that BI has the strongest positive impact on CPB ($\beta = 0.613$; $p < 0.05$), thus hypothesis H2 is accepted. Previous studies have also confirmed the existence of a significant relationship between BI and CPB [14, 15].

The research results show that both coupons and discounts have a positive impact on CPB, with coefficients of $\beta = 0.157$ and $\beta = 0.170$ respectively ($p < 0.05$); therefore, hypotheses H3 and H4 are accepted. Previous studies have also acknowledged the positive and strong relationship between discounts, coupons, and CPB [19].

The research results indicate that BI plays a mediating role in the relationship between SP and CPB ($\beta = 0.372$; $p < 0.05$), thus hypothesis H5 is accepted. Moreover, SP is found to have a significant impact on both BI and CPB, while brand image also positively and significantly

influences purchase decisions [20].

The mediating role of BI in the relationship between SP and CPB is assessed through the Variance Accounted For (VAF) index. According to Hair et al [27], a VAF level below 20% indicates no mediation, 20–80% indicates partial mediation, and above 80% indicates full mediation. The analysis results show a VAF of 100%, confirming that BI fully mediates the relationship between SP and CPB.

This research helps businesses in the apparel industry better understand the factors driving BI and CPB, thereby enabling them to develop appropriate marketing strategies to enhance brand effectiveness and stimulate consumer purchasing behavior. At the same time, the study also clarifies the roles of factors such as SP, coupons, and discounts in the relationship between BI and CPB.

5. CONCLUSION AND MANAGEMENT IMPLICATIONS

The research results reveal four significant direct relationships from H1 to H4: (H1) SP positively affects BI; (H2) BI positively influences CPB; (H3) Coupons have a positive impact on CPB; (H4) Discounts positively affect CPB; and (H5) Brand image mediates the relationship between sales promotion and customer purchasing behavior. Based on these findings, leaders in the apparel industry should consider several strategic implications: first, develop reasonable policies for SP, coupons, and discounts tailored to each stage to optimize effectiveness. Next, they need to implement policies that enhance BI and stimulate CPB, specifically as follows:

(1) Developing reasonable SP policies tailored to specific stages aims to optimize effectiveness by clearly defining objectives, segmenting target customers, and selecting appropriate tools and channels. The goals include achieving outcomes such as increasing revenue efficiently, introducing new products, generating potential customers, accelerating inventory turnover, nurturing customer relationships, while simultaneously enhancing brand image and stimulating CPB.

(2) Developing appropriate coupon policies is an effective marketing strategy aimed at stimulating customers to purchase products and building a positive BI. Coupons not only serve as a token of appreciation from the seller but also help businesses segment their customer base, thereby enhancing CPB.

(3) Developing reasonable discount policies aims to boost sales, efficiently clear inventory, promote products to the market, and enhance brand recognition, with the ultimate goal of improving CPB.

(4) Developing policies to enhance BI aims to create differentiation from competitors, increase awareness, strengthen customer trust, promote brand advocacy, and build a professional BI. A strong BI increases positive perceptions and impressions in customers' minds, helping businesses consolidate their market position and more effectively stimulate CPB.

(5) Additionally, developing policies to improve CPB involves applying effective marketing strategies such as sales promotion activities, coupons, discounts, and brand image building. The goals of these policies are to stimulate shopping demand, increase sales, enhance brand awareness, and encourage CPB.

Table 1. Characteristics of the Research Sample (719)

	Indicator	Frequency	%
Gender	Male	312	43.39
	Female	407	56.61
Age	>= 25	139	19.33
	26 - 30	298	41.45
	31 - 35	169	23.50
	36+	113	15.72
Job position	Staff	299	41.59
	Manage	420	58.41
Income	> 15 million VND	151	21
	15 - million VND	312	43.39
	> 30 million VND	256	35.61
Degree	University	597	83.03
	Postgraduate	122	16.97

Source: Analysis results from the author's survey data, 2025

Table 2. Reliability and validity

Variance & Construct	Outer Loadings	α	CR	AVE	VIF
X1 – Coupons		0.833	0.900	0.749	
X11	0.863				1.891
X12	0.873				1.944
X13	0.862				1.935
X2 – Sales promotion (SP)		0.892	0.918	0.651	
X21	0.834				2.480
X22	0.718				1.684
X23	0.800				2.182
X24	0.856				2.779
X25	0.800				2.203
X26	0.825				2.388
X3 – Discounts		0.842	0.894	0.679	
X31	0.824				1.921
X32	0.855				2.092
X33	0.778				1.625
X34	0.837				1.948
Z – Brand image (BI)		0.915	0.934	0.703	
Z1	0.820				4.502
Z2	0.828				3.601
Z3	0.819				2.182
Z4	0.818				2.368
Z5	0.757				2.562
Z6	0.827				2.178
Y – Consumer purchasing behavior (CPB)		0.896	0.921	0.659	
Y1	0.894				2.147
Y2	0.849				2.241
Y3	0.807				2.295
Y4	0.821				2.213
Y5	0.843				1.783
Y6	0.812				2.261

Source: Analysis results from the author's survey data, 2025

Table 3. Fornell and Larcker criteria

	X1	X2	X3	Y	Z
X1	0.866				
X2	0.498	0.807			
X3	0.499	0.563	0.824		
Y	0.520	0.550	0.582	0.812	
Z	0.480	0.552	0.529	0.662	0.838

Source: Analysis results from the author's survey data, 2025

Table 4. HTMT

	X1	X2	X3	Y	Z
X1					
X2	0.576				
X3	0.596	0.648			
Y	0.602	0.613	0.670		
Z	0.550	0.607	0.604	0.730	

Source: Analysis results from the author's survey data, 2025

Table 5. Hypothesis testing.

Hypothesis	Relationship	Path coefficient	p values	Result
H1	X2 -> Z	0.552	0.000	yes
H2	Z -> Y	0.438	0.000	yes
H3	X1 -> Y	0.180	0.000	yes
H4	X3 -> Y	0.261	0.000	yes
H5	X2 -> Z -> Y	0.242	0.000	yes

Source: Analysis results from the author's survey data, 2025

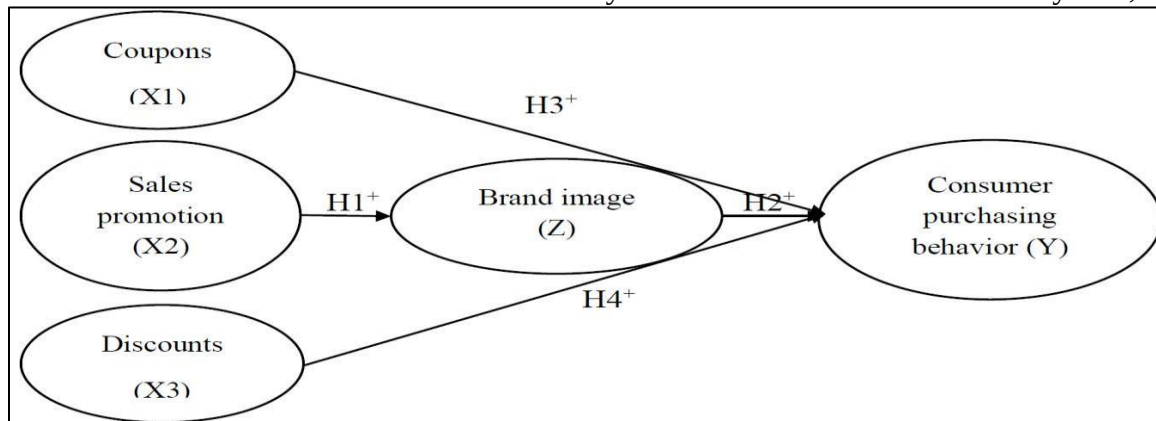


Figure 1. Research Model Source: Author's Proposal

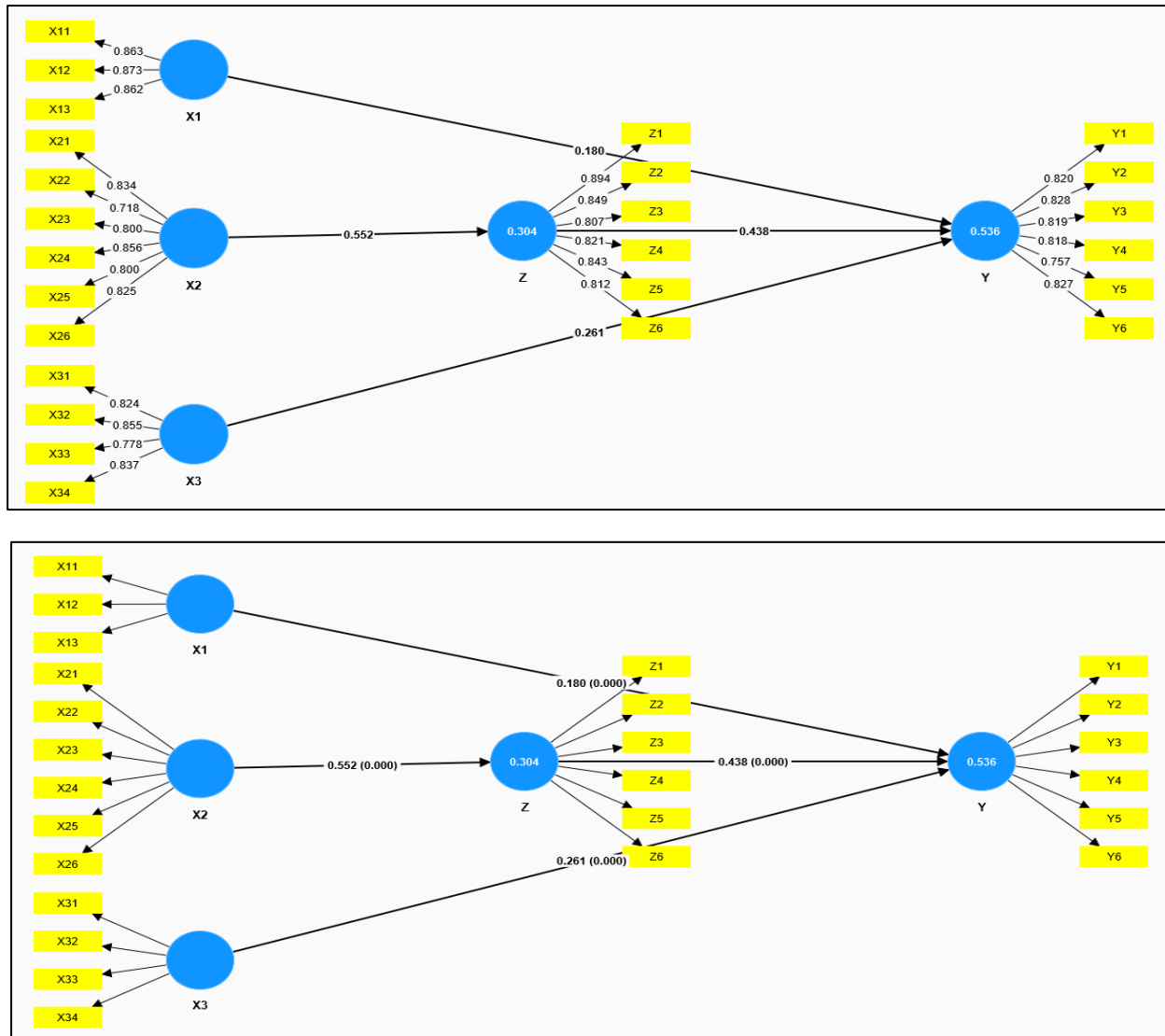


Figure 2. SEM model

Source: Analysis results from the author's survey data, 2025

Acknowledgements:

“This work is funded by Hong Bang International University under grant code GVTTC18.08”.

*Note: Consumer purchasing behavior (CPB), sales promotion (SP), brand image (BI).

REFERENCES

1. [1]. Huynh, N. T. (2024). Status and challenges of textile and garment enterprises in Vietnam and a framework toward industry 3.5. *International Journal of Logistics Research and Applications*, 27(2), 346-357.
2. <https://doi.org/10.1080/13675567.2022.2147490>.
3. [2]. Valaei, N., & Nikhashemi, S. R. (2017). Generation Y consumers' buying behaviour in fashion apparel industry: A moderation analysis. *Journal of Fashion Marketing and Management*, 21(4), 523–543. <https://doi.org/10.1108/JFMM-01-2017-0002>.
4. [3]. Cronin, A. M. (2004). Regimes of mediation: Advertising practitioners as cultural intermediaries? *Consumption Markets & Culture*, 7(4), 349–369. <https://doi.org/10.1080/1025386042000316315>.
5. [4]. Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T).
6. [5]. Gautam, N. and Jain, K. (2008), *Consumer Behaviour*, Gyan Books.
7. [6]. Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of marketing*, 57(1), 1-22.
8. <https://doi.org/10.1177/002224299305700101>.
9. [7]. Kotler, P. & Keller, L.K. (2016). *Marketing Management*. Pearson Education.
10. Fifteenth Edition.
11. [8]. Kotler, Keller, K. L., Brady, M., Victor Goodman, M. R., & Hansen, T. (2019). *Marketing management 4th European edition*. In *Soldering & Surface Mount Technology* (Vol. 13, Issue 3).
12. [9]. Khan, Is., & Warraich, U. A. (2021). Sales promotional tools – A catalyst to speed up sales. *Journal of Marketing Strategies*, 3(2), 38–53. <https://doi.org/10.52633/jms.v3i2.68>.
13. [10]. Fill, C. (2002). *Marketing communications: Contexts, strategies and applications*, 3rd ed.
14. [11]. Khan, M. A., Tanveer, A., & SohaibZubair, S. (2019). Impact of sales promotion on consumer buying behavior: A case of modern trade, Pakistan. *Governance and Management Review (GMR)*, 4(1), 39.
15. [12]. Agmeka, F., Wathoni, R. N., & Santoso, A. S. (2019). The Influence of Discount Framing towards Brand Reputation and Brand Image on Purchase Intention and Actual Behaviour in e-commerce. *Procedia Computer Science*, 161, 851–858. <https://doi.org/10.1016/j.procs.2019.11.192>.
16. [13]. Fennis, B.M. and Pruyn, T.H. (2006), “You are what you wear: brand personality influences on consumer impression formation”, *Journal of Business Research*, Vol. 60 No. 6, pp. 634-639.
17. [14]. Tekin, G., Yiltay, S. and Esra, A.Y.A.Z. (2016), “The effect of Brand image on consumer behaviour: case study of louiss Vuitton-Moet hennessy”, *International Journal of Academic Values Studies*, No. 2, pp. 1-24.
18. [15]. Chovanova, H.H., Korshunov, A.I. and Babcanova, D. (2015), “Impact of brand on consumer behavior”, *Procedia Economics and Finance*, Vol. 34, pp. 615-621.
19. [16]. Durrani, B.A., Godil, D.I., Baig, M.U. and Sajid, S. (2015), “Impact of brand image on buying behavior among teenagers”, *European Scientific Journal*, Vol. 11 No. 5,
20. pp. 155-168.
21. [17]. Riaz, H.A. (2015), “Impact of Brand image on consumer buying behavior in clothing sector: a comparative study between males and females of Central Punjab (lahore) and

- Southern Punjab (multan)”, Kuwait Chapter of Arabian Journal of Business and Management Review, Vol. 4 No. 9, p. 24.
22. [18]. Ahmad, S. A., Mehmood, W., Ahmed, S. A., Mustafa, M., Khan, M. F. T., & Yasmeen, M. (2015). Impact of Sales Promotion on consumer buying behavior in Pakistan. *International Interdisciplinary Journal of Scholarly Research*, 1(3), 13- 22.
 23. [19]. Shahzad, K., Bhatti, A., Islam, T., & Javaid, A. (2020). Impact of sales promotion on consumer buying behavior: A case of garments industry of Pakistan. *International Journal of Advanced Research in Engineering and Technology*, 11(10).
 24. [20]. Helmi, S., Ariana, S., & Supardin, L. (2022). The Role of Brand Image as a Mediation of the Effect of Advertising and Sales Promotion on Customer Purchase Decision. *Journal of Economics and Sustainable Development*, 13(8), 90–99. <https://doi.org/10.7176/JESD/13-8-09>.
 25. [21]. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Prentice Hall.
 26. [22]. Pacheco, B. G., & Rahman, A. (2015). Effects of sales promotion type and promotion depth on consumer perceptions: the moderating role of retailer reputation. *The International Review of Retail, Distribution and Consumer Research*, 25(1), 72-86.
 27. [23]. Sharma, A., & Avasthi, H. (2019). Retail sales promotion in Indian Apparel Industry. *International Journal of Research in Management*, 6(2).
 28. [24]. Weiss, A. M., Anderson, E., & MacInnis, D. J. (1999). Reputation management as a motivation for sales structure decisions. *Journal of marketing*, 63(4), 74-89. <https://doi.org/10.1177/002224299906300407>.
 29. [25]. Dibb, S., Simkin, L., Pride, W. M., & Ferrell, O. C. (2005). *Marketing: Concepts and strategies*. (p. 850). Houghton Mifflin.
 30. [26]. Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>
 31. [27]. Hair, J. F., Ringle, C. M., & Sarstedt, M. (2014). Corrigendum to “editorial partial least squares structural equation modeling: rigorous applications, better results and higher acceptance” [LRP 46/1-2 (2013) 1–12]. *Long Range Planning*, 47(6), 392.s

Aggressive Driving Traits and Tailored Technical Solutions for Two-Wheeler Mobility

Kshipra Moghe* , Saniya Atalatti*

**Assistant Professor, Department of Applied Sciences, COEP Technological University,
Pune, India**

Computer Engineering, Pune Institute of Computer Technology, Pune, India

kam.appsci@coeptech.ac.in, saniyaatalatti05@gmail.com

Abstract—Driving is a potentially stressful activity, and various situations encountered on roadways can elicit angry responses with minimal provocation. This study examines the relationship between personality traits and aggressive driving behavior among individuals who operate two-wheeled vehicles. Utilizing a survey of 315 participants ranging in age from 18 to 60 years, data were collected regarding anger triggers, aggression levels, and personality traits as defined by the Big Five Personality framework. Results indicate significant correlations between traits such as neuroticism, extroversion, and conscientiousness and aggressive driving tendencies. Contextual factors, including traffic conditions and personal experiences, were also found to influence driving aggression. Statistical analysis, including Chisquare tests and correlation matrices, highlight variations in aggression triggers and responses. The findings underscore the importance of tailoring road safety interventions to individual personality traits to mitigate aggressive driving behaviors and promote safer commuting practices among two-wheeler users. Index Terms—Aggressive Driving, Personality Traits, Road Safety, Two-Wheeler Drivers, Big Five Model

1. INTRODUCTION

Aggressive driving behavior is deliberate conduct that may increase collision risk, motivated by impatience, annoyance, hostility, or attempts to save time. Aggressive driving behavior is characterized as intentional conduct that potentially elevates collision risk, driven by factors such as impatience, irritation, hostility, or time-saving attempts. Various incidents can spark anger and aggressive behaviors on roadways. These include but are not limited to: hindered progress, observing reckless driving, and receiving hostile gestures. However, these angerprovoking situations do not elicit uniform responses from all drivers.

Aggressive driving behavior represents a form of externalizing anger, with certain individual and environmental factors mediating the connection between anger triggers and their manifestation on roads. Situational factors linked to road rage incidents include increasing road congestion, high temperatures, poor road conditions, vehicle make, and the anonymity of other drivers.

A driver's response to these triggers is further influenced by individual factors such as personality. Given the extensive research potential in the area of aggressive driving among two-wheeler operators and the heightened risk of death and injury associated with it, the current study was designed to investigate the personality traits that trigger driving-related aggression and its expression among two-wheeler drivers.

The objectives of the paper include:

- Examine the association between personality characteristics and aggressive driving behavior among two-wheeler operators.
- Investigate how specific personality traits, including conscientiousness, neuroticism, and extroversion, correlate with aggressive driving tendencies.
- Review of Aggressive Driving Identification Methods and Personality Tailored Interventions.

2. LITERATURE REVIEW

Aggressive driving, driven by impatience, annoyance, and hostility, significantly increases collision risks [4], [7]. Road rage stems from anger triggers like congestion, reckless driving, and hostile gestures [5]. Factors such as anonymity, high temperatures, and poor road conditions exacerbate these behaviors [3], [9].

Personality traits influence drivers' responses to these triggers. extroversion fosters sociability and assertiveness, potentially escalating confrontations [2]. Neuroticism, linked with emotional instability, heightens reactivity to stress [9]. Conscientious individuals, marked by discipline and responsibility, are less likely to exhibit aggression [11].

Anger expression manifests as active (e.g., honking, verbal aggression) or passive (suppressing anger), reported more by males, impacting road safety and mental well-being [1],[6].

Although research on aggressive driving is extensive, studies focusing on two-wheeler drivers remain limited. This study addresses the gap by exploring the interplay between personality traits, situational factors, and aggressive behaviors in this vulnerable group.

3. DATA COLLECTION

The data were collected through a survey, targeting individuals aged 18 to 60. A total of 315 responses were obtained. The survey comprised four sections: Part 1 collected information on methods of expressing anger; Part 2 focused on triggers of anger while driving; Part 3 assessed aggression levels; and Part 4 gathered data on the Big Five personality traits.

Anger Level	High Aggression - 144 (45.86%)							Low Aggression - 170 (54.54%)						
Demographics														
Gender	Male			Female			Male			Female				
	94 (65.28%)			50 (34.72%)			103 (60.58%)			67 (39.42%)				
Age	18-30			30+			18-30			30+				
	112 (77.78%)		32 (22.22%)			103 (60.58%)		67 (39.42%)						
Employment	Student		Working Professional		Not employed			Student		Working Professional		Not employed		
	105 (72.91%)		39 (27.08%)		2 (0.01%)			94 (55.3%)		59 (34.6%)		17 (0.1%)		
Type of Vehicle	Electric		Geared		Non-Geared			Electric		Geared		Non-Geared		
	17 (11.8%)		53 (36.8%)		74 (51.4%)			12 (7.05%)		44 (25.9%)		114 (67.05%)		
Route	Highway	Suburban	Narrow Roads	Highway, Suburban	Suburban, Narrow Roads	Highway, Narrow Roads	Highway, Suburban, Narrow Roads	Highway	Suburban	Narrow Roads	Highway, Suburban	Suburban, Narrow Roads	Highway, Narrow Roads	Highway, Suburban, Narrow Roads
	18 (10.6%)	39 (22.9%)	25 (14.7%)	7 (4.12%)	27 (15.9%)	10 (5.9%)	18 (10.6%)	15 (8.82%)	36 (21.2%)	51 (30%)	8 (4.7%)	31 (18.23%)	12 (7.05%)	17 (10%)
Distance (km)	0-10	10-20	20-30		30-40	40-50	50+	0-10	10-20	20-30		30-40	40-50	50+
	60 (35.3%)	41 (24.1%)	27 (15.9%)		8 (4.7%)	6 (3.5%)	2 (1.2%)	84 (49.4%)	52 (30.6%)	26 (15.3%)		4 (2.3%)	0 (0%)	4 (2.3%)

4. Fig. 1. Demographics

RESEARCH METHODOLOGY

The purpose of this study is to examine the relationship between personality traits and aggressive driving behavior among two-wheeler operators, and to investigate how specific personality traits, namely conscientiousness, neuroticism, and extroversion, etc correlate with aggressive driving tendencies. To determine the correlation and significance between factors, we utilized the chi-square test and correlation matrix functions in Microsoft Excel. Excel was selected as a costeffective, accessible, and user-friendly tool for data analysis. The Buss-Perry Aggression Questionnaire (BPAQ) was employed to assess aggression levels. Parameters were weighted based on the Likert scale to facilitate a more comprehensive analysis of each variable. To evaluate the personality of each respondent, we employed The Big Five Personality Test, which is widely recognized as one of the most validated psychological personality trait assessments.

TABLE I
GENDER OF RESPONDENTS

Gender	Number of Respondents	Percentage(%)
Male	198	63
Female	117	37

TABLE II
AGE OF RESPONDENTS

Age	Number of Respondents	Percentage(%)
18-30	215	68
30-60	100	32

TABLE III
OCCUPATION OF RESPONDENTS

Occupation	Number of Respondents	Percentage(%)
Students	198	63
Students & Working	14	5
Employed	93	29
Unemployed	10	3

5. RESULTS AND DISCUSSIONS

The surveyed population of 315 individuals has been categorized into two distinct groups based on aggression levels: High aggression and Low aggression. A Likert scale was utilized, and individuals scoring below the mean were classified as low aggressive, while those scoring higher than the mean were classified as high aggressive. This segmentation provides a comprehensive analysis of how aggression levels correlate with various demographic factors, elucidating potential patterns or associations within the surveyed population.

The subsequent section of the survey examined the factors that elicit an aggressive response, specifically the aggression triggers. The 15 anger triggers were analyzed using a Chisquare test, with groups categorized based on aggression level, age, gender, type of vehicle driven, and employment status.

Chi-square tests were conducted to establish the relationship between the triggering factors and the aforementioned groups. Chi-square tests were also performed to establish the relationship between anger triggers and the Big Five personality traits. Of the Big Five personality traits, three (extroversion, Conscientiousness, and Neuroticism) were deemed most relevant to this study.

Subsequently, the relationship between anger expression (active and passive methods of aggression) and the Big Five was analyzed using the Chi-square test. To conclude, a correlation matrix was obtained between high and low aggression, anger expression, and the three traits selected from the Big Five. The matrix serves as a graphical representation of the relationship between seven key factors considered crucial in understanding the interplay between personality traits and aggressive driving behaviors among two-wheeler drivers. These factors include anger triggers, aggression levels, and Big Five personality traits.

The matrix employs a standard scale ranging from -1 to 1 to quantify the degree of correlation between these factors. Positive values on the scale indicate a positive correlation, signifying that as one factor increases, the other factor also tends to increase. Conversely, negative values signify a negative correlation, indicating that as one factor increases, the other tends to decrease. The significance level for the correlation matrix is 0.05.

A. Personality Analysis

Out of the five Big Five personality traits, extroversion, Conscientiousness, and Neuroticism, were deemed to be relevant to this experiment. The data presented in the provided table indicates that the Big Five personality traits do not exert a significant influence on the expression of anger. The correlation between anger expression and Big Five personality traits such as extroversion, neuroticism, and conscientiousness are termed parameters A, B, and C.

H_0 = Anger expression is independent of personality traits such as extroversion, neuroticism, and conscientiousness

A – Anger expression and extroversion as the calculated value of $p(0.73428) > 0.05$, the H_0 is accepted. This concludes that the expression of anger is independent of extroversion.

B – Anger Expression and Neuroticism as the calculated value of $p(0.629841) > 0.05$, the H_0 is accepted. This concludes that anger expression is independent of neuroticism.

C – Anger Expression and Conscientiousness as the calculated value of $p(0.86663) > 0.05$, the H_0 is accepted. This concludes that anger expression is independent of conscientiousness. This shows that H_0 is accepted for all these cases.

The study analyzed the relationship between anger expression and three personality traits from the Big Five model: extroversion, Neuroticism, and Conscientiousness. The p-values obtained for these traits were all exceeding the significance threshold of 0.05. Consequently, the H_0 was accepted in each case, indicating no statistically significant relationship between anger expression and these personality traits.

The factors that influence anger triggers have been grouped into 5 broader categories based on their similarities. The first category, Road and Traffic Conditions, includes factors such as traffic jams (caused by accidents, political rallies, or quarrels), roads in bad condition (potholes, open drainage), encroachment on roads (illegal shops, incorrect parking), heavy vehicles (trucks, JCBs), overtaking from the wrong side, and general indiscipline on the road (speeding, lane cutting). The second category, Noise and Disturbances, encompasses honking (loud or irritating horn noise), loud music (from cars or during festivities), and distractions caused by others using earphones or talking on the phone. The third category, Hawkers and Roadside Activities, includes hand carts, hawkers at signals (windshield cleaners, garland sellers, etc.), and drivers spitting on roads. The fourth category, Environmental and Personal Factors, covers the impact of weather conditions (hot temperatures, rain), pollution (dust, smoke), and personal factors such as mood, hunger, and stress levels. Lastly, the fifth category, Time Pressure and Urgency, focuses on situations involving deadlines or important meetings. These categories provide a simplified framework to analyze how different factors contribute to aggression.

1) Neuroticism Analysis: Upon examination of the data using the Chi-Square test, it becomes apparent that individuals surveyed, irrespective of their neuroticism scores, exhibited limited sensitivity to various environmental factors.

This indicates that neuroticism does not exhibit a disproportionately strong or weak influence on any specific anger trigger. Instead, the effects of neuroticism appear to be uniformly distributed across all these triggers. The consistently high p-values across all categories underscore the absence of significant variation, meaning there is no statistically meaningful evidence to suggest that individuals with higher neuroticism levels respond differently to specific types of anger triggers.

TABLE IV
ANGER TRIGGERS VS NEUROTICISM

Anger Triggers vs Neuroticism			
Alpha = 0.05			
Anger Triggers	χ^2	P-value	Inference
Road and Traffic Conditions	16.578	0.680	H_0 Accepted
Noise and Disturbances	15.126	0.769	H_0 Accepted
Hawkers and Roadside Activities	2.63	0.999	H_0 Accepted
Environmental and Personal Factors	4.51	0.99988	H_0 Accepted
Time Pressure and Urgency	5.48	0.99944	H_0 Accepted

2) Extroversion Analysis: Analysis of the data using the Chi-square test reveals that individuals surveyed, irrespective of their extroversion scores, demonstrated limited sensitivity to various environmental factors.

TABLE V
ANGER TRIGGERS VS EXTROVERSION

Anger Triggers vs Extroversion			
Alpha = 0.05			
Anger Triggers	χ^2	P-value	Inference
Road and Traffic Conditions	35.708	0.0166	H_0 Rejected
Noise and Disturbances	24.684	0.2138	H_0 Accepted
Hawkers and Roadside Activities	4.102	0.99994	H_0 Accepted
Environmental and Personal Factors	9.902	0.9699	H_0 Accepted
Time Pressure and Urgency	5.75	0.99920	H_0 Accepted

However, individuals with both high and low extroversion scores exhibited susceptibility to certain environmental factors. Specifically, factors including others using earphones/talking on the phone, encroachment on roads, traffic jams, honking, weather, and bad roads were found to affect individuals with varying levels of extroversion. Notably, individuals with higher extroversion scores demonstrated a heightened susceptibility to factors such as others using

earphones/talking on the phone, encroachment on roads, traffic jams, honking, weather, and bad roads compared to those with lower extroversion scores.

3) Conscientiousness Analysis: Analysis of the data using the Chi-Square test reveals that surveyed individuals, regardless of their conscientiousness scores, exhibited minimal sensitivity to various environmental factors.

TABLE VI
ANGER TRIGGERS VS CONSCIENTIOUSNESS

Anger Triggers vs Conscientiousness			
Alpha = 0.05			
Anger Triggers	χ^2	P-value	Inference
Road and Traffic Conditions	71.142	0.000001	H_0 Rejected
Noise and Disturbances	20.904	0.411	H_0 Accepted
Hawkers and Roadside Activities	5.632	0.999	H_0 Accepted
Environmental and Personal Factors	13.318	0.899	H_0 Accepted
Time Pressure and Urgency	4.794	0.9995	H_0 Accepted

For the anger trigger of road and traffic conditions, the null hypothesis is rejected, indicating that conscientiousness has a significantly uneven effect on this specific trigger. This suggests that individuals with varying levels of conscientiousness respond differently to road and traffic conditions, making it a key factor in understanding anger reactions for this category. In contrast, for other triggers such as noise and, hawkers and roadside activities, environmental and personal factors, and time pressure and urgency, the null hypothesis is accepted. This indicates that conscientiousness does not have a significantly uneven effect on these triggers, implying that the trait's influence is uniformly distributed across these categories. The rejection of the null hypothesis for road and traffic conditions highlights a notable exception, suggesting that conscientiousness plays a distinct role in how individuals perceive and react to anger in traffic-related scenarios.

B. High Aggression - Low Aggression analysis

For road and traffic conditions and environmental and C. Corelation Matrix TABLE VIII CORRELATION MATRIX personal factors, the null hypothesis is rejected. This indicates that aggression levels have a significantly uneven effect on these specific triggers, suggesting that individuals with varying aggression levels respond differently to these scenarios. These triggers

appear to evoke stronger or more distinct responses based on the individual's aggression level. In contrast, for noise and, hawkers and roadside activities, and time pressure and urgency, the null of the individual's aggression level. The findings highlight that certain triggers, such as road and traffic conditions and environmental and personal factors, are more sensitive to variations in aggression levels. In contrast, others like noise, hawkers, and time pressure show no significant variation, pointing to a more generalized influence of aggression in these contexts.

H_0 : the effect of aggression level on anger triggers is evenly distributed

TABLE VII
ANGER TRIGGERS VS AGGRESSION LEVEL ANALYSIS

Anger Triggers vs Aggression Level			
Alpha = 0.05			
Anger Triggers	χ^2	P-value	Inference
Road and Traffic Conditions	56.03	0.0000288	H_0 Rejected
Noise and Disturbances	19.12	0.514	H_0 Accepted
Hawkers and Roadside Activities	16.15	0.707	H_0 Accepted
Environmental and Personal Factors	46.05	0.000793	H_0 Rejected
Time Pressure and Urgency	17.60	0.614	H_0 Accepted

C. Corelation Matrix

TABLE VIII
CORRELATION MATRIX

	extroversion	Neuroticism	Conscientiousness	High	Low	Active	Passive
Extroversion	1	0.6988	0.7310	0.1283	0.1095	0.0157	-0.0157
Neuroticism	0.6988	1	0.6929	0.0148	0.0326	-0.0405	0.0405
Conscientiousness	0.7310	0.6929	1	0.2216	0.2007	0.0625	-0.0625
High	0.1283	0.0148	0.2216	1	-1	0.2901	-0.2901
Low	0.1095	0.0326	0.2007	-1	1	0.1622	-0.1622
Active	0.0157	-0.0405	0.0625	0.2901	0.1622	1	-1
Passive	-0.0157	0.0405	-0.0625	-0.2901	-0.1622	-1	1

active and passive anger expression, suggesting its role in managing aggression and anger-related behaviors. Furthermore, high aggression levels are closely linked to both active and passive anger expression, indicating a comprehensive relationship between aggression intensity and the ways anger is expressed. These findings underscore the intricate interplay between

aggression intensity and the ways anger is expressed. These findings underscore the intricate interplay between personality traits, aggression levels, and anger expression styles. Certain factors do not show significant correlations, highlighting distinct separations between specific traits and behaviors. Extroversion is not significantly correlated with either active or passive anger expression, suggesting that this personality trait does not strongly influence the style of anger expression. Similarly, neuroticism does not exhibit significant correlations with high or low aggression levels, nor with active or passive anger expression, indicating that emotional instability may not directly determine aggression intensity or anger expression styles. These findings underline the independence of these traits and behaviors, emphasizing the complexity of their interactions.

D. Review of Aggressive Driving Identification and Personality-Tailored Interventions

Understanding and mitigating road rage in two-wheeler drivers requires a dual approach: leveraging psychological insights and integrating intelligent technologies. The interplay between personality traits—such as neuroticism, extroversion, and conscientiousness—and aggressive driving, as revealed in this study, highlights the need for interventions that are both individualized and context-aware. Grounded in HumanComputer Interaction (HCI) principles, the following frameworks present implementable, scalable solutions that align with emerging methodologies such as real-time speech analysis, emotion classification via machine learning, and adaptive driver assistance systems [12]–[14]. These design models offer a pathway toward safer and more psychologically responsive commuting experiences.

1. Personality-Tailored Smart Helmet System

Informed by advances in real-time emotion recognition and contextual alert systems [14], the PersonalityTailored Smart Helmet integrates psychological profiling with situational analysis to deliver adaptive feedback to riders. Drawing from the concept of Advanced Driver Assistance Systems (ADAS), the helmet uses pre-ride personality data to shape in-ride auditory interventions tailored to rider profiles such as high neuroticism or low conscientiousness. By integrating traffic data, environmental inputs, and rider behavior (e.g., sudden braking, traffic congestion), the system proactively delivers stress-alleviating prompts. These prompts, ranging from deep breathing cues to structured behavioral reminders—are delivered via embedded speakers, making the helmet both a protective device and an emotional regulation interface.

User Flow:

- 1) The rider pairs the helmet with a mobile app and completes a Big Five Personality test.
- 2) The app generates a psychological profile and syncs it with the helmet system via Bluetooth.
- 3) During rides, the helmet utilizes traffic APIs and embedded motion sensors to detect situational stressors.
- 4) Based on personality and context, the helmet delivers tailored feedback (e.g., calming affirmations for neuroticism, assertive instructions for low conscientiousness).
- 5) Riders interact via voice or head gestures, preserving safety while controlling feedback.
- 6) Post-ride, the app offers a behavior summary and updates response patterns for future personalization.
- 7) The system evolves over time, aligning more closely with rider behavior and emotional patterns.

This system reflects principles from ADAS and psychological feedback theory, offering a low-friction, real-time intervention platform consistent with evidence on driver stress mitigation [14].

2. Personality-Tailored Smart Helmet System

Building on emotion classification models such as CNNs and Bi-LSTM networks [13], this mobile application translates behavioral data and emotional cues into an engaging, gamified self-regulation system. The app incorporates speech features such as Mel Frequency Cepstral Coefficients (MFCC), which have demonstrated high reliability in detecting emotional stress in noisy environments [12]. These features, coupled with personalized feedback loops, help riders track and improve emotional control during commuting. Challenges are configured based on individual traits: extroverted users receive socially competitive tasks, while high-neuroticism users engage with calming reinforcement modules. By combining gamification and psychological insight, the app encourages users to adopt safer behaviors and recognize their personal anger triggers in real-world contexts.

User Flow:

- 1) Onboarding begins with a personality test to determine the user's psychological profile.
- 2) The app configures gamified challenges and emotional tracking parameters based on trait scores.
- 3) Prior to each ride, users complete a quick emotion check-in to prime the system.
- 4) The app passively records speech signals (via MFCC extraction) and motion data to detect signs of stress or aggression [12].
- 5) During episodes of potential road rage, the app sends personalized, real-time nudges (e.g., relaxation prompts or motivational audio).
- 6) Users earn points for controlled behavior and unlock milestones or badges for streaks of safe driving.
- 7) Weekly summaries provide reflection tools, visual progress charts, and adaptive challenge recommendations.

This framework reflects the power of hybrid emotion recognition models [13] and behavior-change strategies grounded in HCI and affective computing, providing a lightweight, scalable solution to improve self-awareness and reduce road aggression.

6. CONCLUSION

This research establishes a significant link between personality traits—such as extroversion, neuroticism, and conscientiousness—and aggressive driving behaviors in two-wheeler riders, offering a psychological lens through which road rage can be better understood and addressed. By integrating behavioral insights with emerging technologies, the study proposes two practical, user-centered design frameworks: a **personality-tailored smart helmet system** and a **gamified emotional regulation mobile app**. These solutions apply Human-Computer Interaction (HCI) principles and leverage tools like **speech emotion recognition (using MFCCs)** and **machine learning models (e.g., CNNs, BiLSTM)** to provide real-time, context-sensitive feedback and motivation. By aligning interventions with individual personality profiles, these systems can deliver adaptive, nonintrusive cues that help mitigate aggression during high-stress traffic scenarios.

Beyond these frameworks, the findings support broader AI-driven strategies such as emotion-aware Advanced Driver Assistance Systems (ADAS), personalized training modules, and community-based engagement programs. These initiatives—combined with targeted public messaging, virtual situational training, and educational interventions—can foster safer, more emotionally intelligent road behavior at scale. As AI and affective computing continue to evolve, integrating psychological profiling into traffic systems holds great potential to reduce road-related aggression, prevent accidents, and build a more responsible and empathetic commuting culture.

REFERENCES

- [1] Bjõrklund, G. (2008). Gender differences in driver aggression: A review. *Transportation Research Part F: Traffic Psychology and Behaviour*, 11(2), 146-158.
- [2] Dukes, R. L., Clayton, S. D., Jenkins, K., Miller, R., & Rodgers, C. M. (2001). Effects of traffic congestion on drivers' emotions and aggressive behavior. *Environment and Behavior*, 33(1), 25-40.
- [3] Ellison, P., Govern, J., Petri, H. L., & Figler, M. H. (1995). Anonymity and aggressive driving behavior: A field study. *Journal of Social Behavior and Personality*, 10(1), 265- 272.
- [4] Galovski, T. E., & Blanchard, E. B. (2004). Road rage: A domain for psychological intervention. *Aggression and Violent Behavior*, 9(2), 105-127.
- [5] Hennessy, D. A., & Wiesensthal, D. L. (2002). The relationship between traffic congestion, driver stress, and aggression. *Transportation Research Part F: Traffic Psychology and Behaviour*, 5(2), 129-144.
- [6] James, L., & Nahl, D. (2000). *Road rage and aggressive driving: Steering clear of highway warfare*. Prometheus Books.
- [7] Lajunen, T., & Parker, D. (2001). Are aggressive people aggressive drivers? A study of the relationship between self-reported general aggressiveness, driver anger, and aggressive driving. *Accident Analysis & Prevention*, 33(2), 243-255.
- [8] Roberts, L. D., & Indermaur, D. (2005). The social construction of "road rage": The discursive construction of driving anger and aggression in talkback radio. *Psychology and Crime*, 11(4), 373- 388.
- [9] Shinar, D. (1998). Aggressive driving: The contribution of the drivers and the situation. *Transportation Research Part F: Traffic Psychology and Behaviour*, 1(2), 137-160.
- [10] Smart, R. G., Stoduto, G., Mann, R. E., & Adlaf, E. M. (2004). Road rage and collisions: A study of the relationship between drivers' anger and traffic accidents. *Accident Analysis & Prevention*, 36(1), 173-181.
- [11] Tasca, L. (2000). A review of the literature on aggressive driving behavior. *Canadian Transportation Research Forum*, 10, 1-25.
- [12] Norhaslinda Kamaruddin, Abdul Wahab Abdul Rahman, Khairul Ikhwan Mohamad Halim, Muhammad Hafiq, Iqmal Mohd Noh (2018). Driver Behaviour State Recognition based on Speech. *TELKOMNIKA Telecommunication Computing Electronics and Control*, 16(2):852-861.
- [13] Yang Liu, Wenjing Wang, Xinmin Xu. Road Rage Recognition System Based on Speech Features. 21st Inter- national Symposium on Distributed Computing and Appli- cations for Business Engineering and Science (DCABES), Chizhou, China, 2022, pp. 217- 220, 217-220 .
- [14] Fadi Al Machot, Ahmad Haj Mosa, Alireza Fasih, Christopher Schwarzlmüller, Mouhanndad Ali Ali, Kyan- doghere Kyamakya. (2012). A Novel Real-Time Emo- tion Detection System for

REVEALING ABORTION STIGMA: A QUALITATIVE STUDY IN EARLY ADULT WOMEN WHO HAVE AN ABORTION

Aaliyah Azzakiyah Tasning, Salsabila Hetri Soraya, Ratna Rusmiyati, Aries Yulianto
Department of Psychology, Universitas Pembangunan Jaya, Indonesia

Abstract

Women who experience an unwanted pregnancy may take action to keep the pregnancy or have an abortion. According to the WHO, there are approximately 73 million abortions annually worldwide and more than 45% of them are performed under unsafe conditions. Women who have abortions may experience abortion stigma, which is a negative attribute given to women who want to terminate the pregnancy by labeling them, both internally and externally, as a person who does not fit the ideal image of a woman. The objective of this study is to obtain a description of the abortion stigma experienced by early adult women who have abortions due to unwanted pregnancies. This study used qualitative research methods with a phenomenological approach. The sample selection used criterion sampling method, where two young adult women who had abortions were selected as informants. The results showed that both informants faced significant emotional and social distress, especially worries about judgment and self-judgment. Factors such as social environment and access to abortion information also influenced the informants' experiences. Informant 1 had a lower level of stigma due to social support that reduced the emotional impact she felt. This study reinforces the importance of advocacy to reduce the social stigma associated with abortion and expand access to information and psychological support for women in similar situations.

Keywords: abortion stigma, early adult women, unwanted pregnancy, abortion, phenomenology

1. Introduction

1.1 Unwanted pregnancy

Pregnancy is one of the important phases in a woman's life that marks the beginning of major changes, both physically, emotionally, socially, and in life priorities (Fatimah, 2023). In many cultures, pregnancy is considered a gift. However, not all women welcome pregnancy with feelings of happiness. Unwanted pregnancy becomes an unavoidable phenomenon, especially when pregnancy occurs outside of planning or in unsupportive conditions. Unwanted pregnancy occurs when one or both partners do not fully desire the pregnancy or only at that time (Sari, 2021). Data shows that in the 2015–2019 period, there were approximately 121 million cases of unwanted pregnancy globally, and more than 60% of them ended in abortion (Guttmacher Institute, 2019). In Indonesia, the BKKBN recorded that the rate of unwanted pregnancy reached 17.5% per year, occurring mainly in women aged 15–27 years (17.9%) and 40–49 years (21.4%) (BKKBN, 2019). Although this figure is still relatively low compared to the global scale, the majority of unwanted pregnancies in Indonesia still result in abortion (Fitriani, 2019).

1.2 Types of abortion and the global context

An abortion can occur naturally or be induced through human intervention. Medically, abortion is divided into three main types: (1) spontaneous abortion, which is a natural miscarriage due to factors such as infection or uterine abnormalities (Anindya, 2020), (2) medical abortion, which uses medication to induce the expulsion of the fetus. This method is more commonly used and considered safer than surgery (World Health Organization, 2020), and (3) surgical abortion, through physical procedures such as

vacuum aspiration or dilation and curettage (Putri, 2020). It is estimated that 73 million

abortions occur worldwide each year, and more than 45% of them are performed under unsafe conditions (World Health Organization, 2022). Approximately 97% of these unsafe abortions occur in Africa, Asia, and Latin America. Countries with strict abortion regulations generally have high abortion-related mortality rates, whereas countries that provide access to legal and safe abortion services have lower mortality rates (Bearak et al., 2020; Ushie et al., 2019).

1.3 Reasons for abortion and social stigma

The reasons women have abortions are very diverse, including health reasons, economic conditions, and social pressure due to unwanted pregnancies. In a health context, an abortion is performed to save the mother's life or to avoid severe risks to the fetus, making it a crucial medical procedure (Santoso, 2022). Economically, many women feel unable to meet the basic needs of the child to be born (Astuti, 2020). Meanwhile, morally, some women experience social and family pressure due to pregnancies considered non-normative (Gultom, 2021), especially if they occur outside of marriage.

Abortion is a complex and highly controversial global issue. In a conservative country like Indonesia, the act of abortion is not only considered a violation of the law (for example in Pasal 346 KUHP) but is also seen as conflicting with religious values and social norms (Hadijah, 2019). In Islam, the majority religion in Indonesia, abortion is viewed as a violation of the fetus's right to life (Mustofa, 2020). This view strengthens the stigma against women who have abortions. Consequently, many women do not receive empathy from their communities. Furthermore, women who have had an abortion tend to face challenges in their education and careers as a result of their actions (Khofifah & Atika, 2024).

1.4 Research focus: Stigma on early adult women

This study focuses on how the social stigma surrounding abortion is personally experienced by young adult women who have had an abortion outside of marriage. In a culture that upholds religious values and moral norms, the decision to have an abortion can create immense internal pressure and social stigma. Women are often faced with a dilemma between their personal needs and the pressure from a judgmental social environment. To gain a deeper understanding of how this stigma affects individuals, this research employs a qualitative phenomenological method, which allows the researcher to explore subjective experiences in depth. The purpose of this study is to uncover the meaning of the experience of stigma as felt by young adult women after undergoing an illegal abortion, as well as how they internalize the social and emotional pressures they face.

2. Literature review

2.1 Definition of abortion stigma

Abortion stigma was introduced by Kumar et al. (2009), who defined it as a negative attribute assigned to women who seek to terminate a pregnancy, labeling them, both internally and externally, as individuals who do not conform to the ideal image of a woman. Furthermore, abortion stigma is also defined as a shared understanding that

abortion is morally wrong and/or socially unacceptable. Abortion stigma manifests at various levels, including media, law and policy, institutions, communities, relationships, and the individual (Cockrill et al., 2013).

2.2 Aspects of abortion stigma

There are four aspects of abortion stigma: worries about judgement, isolation, self-judgement, and community condemnation (Cockrill et al., 2013). Worries about judgement is a state where an individual is worried about the negative views of others regarding their abortion and is also concerned about the deterioration of their relationships with others. Isolation is a state where an individual talks about their abortion to those closest to them to see the extent to which they feel supported by those people. Self-judgement is a state where an individual views themselves negatively for having had an abortion. Community condemnation refers to the perception of women who have had an abortion regarding the views of people in their community, who mostly hold negative views on abortion.

2.3 Factors of abortion stigma

There are five factors influencing why each individual has a different level of abortion stigma: age, ethnicity, level of education, as well as religion and degree of religiosity (Cockrill et al., 2013). Typically, younger individuals have more open views compared to older ones, who may be more influenced by traditional norms. Younger individuals tend to be more accepting of abortion than their older counterparts. Additionally, the cultural background and norms of an ethnic group can influence a person's views on abortion. Some ethnic groups may have traditions or values that are more accepting of abortion, while others may still be unable to accept it and tend to view it negatively. The level of education is also often related to a better understanding of reproductive health and individual rights. People with higher education tend to be more open-minded about abortion. And lastly, religious beliefs play a crucial role in a person's view of abortion. Many religious teachings consider abortion an unethical act, so people with a high degree of religiosity may be more inclined to have a negative view of abortion.

2.4 Impacts of abortion stigma

The stigma from the surrounding community regarding abortion can impact the psychological, physical, and reproductive health of women who have abortions. Negative community stigma related to abortion can cause women who have them to feel shame, guilt, anxiety, and even depression. This pressure from stigma can lead these women to feel isolated and afraid to speak out or seek help due to fear of being judged, thereby increasing the risk of mental health disorders. Furthermore, the negative stigma associated with abortion can cause women who have had one to choose not to get their physical and reproductive health checked post-abortion for fear of being judged, which can lead to various future health problems that could have been prevented (Jones & Jerman, 2017).

3. Problem definition/formulation

The research question in this study is: "How are the dynamics of abortion stigma in young adult women who choose to have an abortion due to premarital pregnancy?"

4. Research Informants

The researchers used a purposive sampling technique for informant selection, considering the suitability of the subjects' characteristics for the research objectives. Informants were purposely selected based on their relevant experience: having had an abortion outside of marriage and consequently facing social stigma. The number of informants in this study was two young adult

women who met the predetermined criteria.. The researchers ensured that the informants understood the purpose and scope of the research before the interviews were conducted, and obtained their voluntary consent to participate with an assurance of identity confidentiality.

5. Method of analysis

This study uses a qualitative research method with a phenomenological approach. A phenomenological approach is one that is used to describe the common meaning of a lived experience or a phenomenon that occurs for an individual or a group (Creswell & Poth, 2018). Data collection in this study was conducted over two meetings, each lasting 90 minutes for a total time of 180 minutes, and used three data triangulation methods: interviews, observation, and questionnaires.

6. Result and discussion

6.1 Physical and emotional effects after abortion

The interview results indicate that both subjects (SS and BI) experienced significant physical and emotional impacts after their abortions. Subject SS experienced severe abdominal pain accompanied by bleeding, which made her feel weak and caused her body weight to drop drastically. Subject BI also experienced bleeding following a surgical abortion procedure. Neither subject sought further medical care due to their fear of being judged by others for having had an abortion.

Regarding subject SS, she experienced emotional impacts in the form of increased sensitivity, feelings of guilt, and a tendency for self-blame. Nevertheless, SS is trying to accept her situation and come to terms with it:

“My feelings were all mixed up... um, sad, angry, frustrated, and in pain. That's what I felt because, well, no matter what, I had already grown to love my fetus. But I also have to get on with my life. So, like it or not, I just have to try and sincerely accept it.” (SS, 19 Nov 2024, 373)

In the case of subject BI, she emotionally experienced deep sadness, guilt, and trauma that surfaced whenever she encountered certain triggers:

“When there are triggers about abortion or pregnant minors on social media, it really haunts me.” (BI, 14 Des 2024, 52)

Unlike SS, who tried to be open with those closest to her, BI chose to hide her experience for fear of being judged, even abandoning her intention to seek professional help:

“After thinking it over, I decided not to go to a psychologist or for counseling after the abortion because I was afraid of being judged negatively. I just wasn't ready, you know, to face the stigma and negative opinions from other people.”
(BI, 14 Des 2024, 54)

On an emotional level, both subjects experienced deep sadness, feelings of guilt, and trauma after the abortion. However, their responses to these feelings differed. SS showed an effort to accept and come to terms with the situation, even though she was still grappling with guilt. She tried to look ahead and attempted to continue her life with more sincere acceptance. In contrast, BI tended to avoid conversations about the abortion, withdrew from her social circle, and

postponed seeking professional help for fear of facing stigma and negative judgment from others.

This difference indicates that social support and the ability to be open with one's social circle influence the emotional recovery process. SS, who bravely acknowledged her emotions and tried to find ways to accept her situation, began to enter the healing phase more quickly. Meanwhile, BI, who chose to keep the experience to herself, faced guilt and trauma for a longer period.

These findings reinforce the theory of abortion stigma (Cockrill et al., 2013), wherein self-judgement and worries about judgment play a significant role in exacerbating the post-abortion emotional impact. Individuals who fear judgment tend to isolate themselves, avoid support, and experience more severe psychological distress.

6.2 Living in the shadow of judgment

In SS's neighborhood, they never openly discuss their views on abortion. However, based on the beliefs held by SS, abortion is considered a sin and forbidden. Despite having made a mistake, SS's spirituality has grown stronger as they continue to seek redemption by strengthening their prayers and honoring their faith. Within SS's family environment, they also do not openly discuss their views on abortion, but SS's parents always advise their children to take care of themselves. SS occasionally hears negative comments about abortion, but these comments do not influence their decision, and SS believes that society should educate itself about abortion:

"Actually, according to my beliefs, I think abortion is a mistake, maybe even a sin, but I've kept the mistake I made a secret and I haven't made the same mistake again. I'm trying to be better, I'm trying to be more careful, I'm trying not to make the same mistake again. It's gotten stronger because after the abortion, honestly, whenever I want something, I have to pray. I just want to strengthen my prayers to God. I want to beautify and strengthen my faith from the mistake I made, which was very serious, and it's definitely considered a sin. But that's why now, while I can still honor my beliefs, I do it. No, there isn't. My family doesn't discuss abortion, but my parents often say things like, "Take care of yourself because we don't know what kind of people are out there, we don't know what kind of relationships people have, so maybe you should just take care of yourself." My parents mostly tell me to take care of myself and give me advice, but they don't talk about abortion or anything like that. Actually, if there are still many people who think it's okay to ignore it, there's no harm in finding out what abortion is, and there's no

harm in learning about abortion education. Perhaps if you still want to dismiss it, that's fine, but go back to the people around you and ask them what they think and how they feel about it. But it's called education, seeking knowledge, so why not? Right?" (SS, 19 Nov 2024, 455-473)

Meanwhile, in the BI subject's environment, people openly show that they view abortion as a negative and shameful act, and in their religion, abortion is considered a grave sin. This view causes the BI subject to feel uncomfortable and "slapped" when they hear people around them making negative comments about abortion without knowing that BI has done so:

"Well, the people around me think that abortion is a really negative thing to do, and they even see it as a big sin and something to be ashamed of. But that's just how it is, so I don't want the

people around me to know that I did it because it's a really negative thing, and in my religion, it's a really big sin. Hmm, from the perspective of the people around me, I feel like I've done something very wrong and unusual, so it affects me because my life becomes more anxious, and hearing others discuss it without knowing what I've done makes me feel like I'm being judged by what they think of me, you know?" (BI, 14 Des 2024, 146-148)

Subjects SS and BI stated that they believed abortion to be a grave sin. In the SS subject's neighborhood and family, they do not openly express their views on abortion, but the SS subject's mother and father often advise their children to take care of themselves. However, SS has occasionally heard about people's negative views on abortion and believes that the community should be educated to learn more about abortion. Meanwhile, in BI's neighborhood, people often openly express their views and consider abortion to be a negative act. However, BI's family never discusses their views on abortion. When BI hears people's negative views on abortion, she focuses only on her feelings and becomes increasingly worried.

This is in line with the aspect of abortion stigma identified by Cockrill et al. (2013), namely community condemnation, which refers to the perception of women who have had abortions regarding the views of people in their community, most of whom have negative views about abortion. Based on this explanation, it can be said that subject SS has a strong desire to educate people so that abortion can be more accepted by society. Meanwhile, BI is only focused on her feelings and concerns about the negative views of others that may be directed toward her.

7. Conclusion

The results of the study indicate that the dynamics of abortion stigma among young adult women who choose to have an abortion due to premarital pregnancy are significantly influenced by the internal and external stigma they face. Both study subjects experienced significant emotional and social pressure, particularly concerns about judgment and self-judgment. The first subject (SS) and the second subject (BI) had different experiences in dealing with stigma. The level of abortion stigma experienced by subject SS was lower than that of subject BI. This was due to the significant support received by SS from their partner and friends, which helped reduce emotional isolation and provided space to consider the decision more rationally. Conversely, subject BI, who

faced greater moral pressure from family and limited support, tended to experience higher levels of stigma. Her fear of family reactions led to emotional isolation and the decision to undergo an abortion at an illegal clinic out of necessity. Factors such as social environment and access to abortion information also influenced the experiences of both subjects. Subjects in the SS group utilized formal sources of information such as consultations with doctors and health apps, while subjects in the BI group relied more on informal channels, such as friends who had experienced similar situations. Strong social stigma, particularly in conservative societies, caused both groups to feel isolated and compelled to keep their experiences secret. This had an impact on mental health, including feelings of shame, guilt, and post-abortion trauma. Social support, whether from partners or close friends, has been shown to help reduce the emotional burden felt, although it does not completely eliminate the impact of stigma. This conclusion underscores the complexity of abortion stigma dynamics among young adult women in Indonesia, which involves the interplay of social, cultural, and individual factors. The lower level of stigma among SS subjects highlights the importance of social support in reducing the emotional impact felt. This study reinforces the importance of advocacy to reduce social stigma related to abortion and expand access to information and psychological support for women facing similar situations.

8. Reference

- 1) Adrian, K. (2025, January 2). Penyebab ibu hamil muda lebih emosional dan cara mengatasinya. Alodokter.
- 2) Auli, R. C. (2024, January 31). Bunyi pasal 346 KUHP tentang aborsi.
- 3) Biggs, M. A., Brown, K., & Foster, D. G. (2020). Perceived abortion stigma and psychological well-being over five years after receiving or being denied an abortion. *PLoS ONE*, 15(1). <https://doi.org/10.1371/journal.pone.0226417>
- 4) Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>.
- 5) Cockrill, K., Herold, S., Upadhyay, U., Baum, S., Blanchard, K., & Grossman, D. (2013). Addressing abortion stigma through service delivery.
- 6) Cockrill, K., Upadhyay, U. D., Turan, J., & Foster, D. G. (2013). The Stigma of Having an Abortion: Development of a Scale and Characteristics of Women Experiencing Abortion Stigma. *Perspectives on Sexual and Reproductive Health*, 45(2), 79–88. <https://doi.org/10.1363/4507913>
- 7) Cowan, S. K. (2017). Enacted abortion stigma in the United States. *Social Science and Medicine*, 177, 259–268. <https://doi.org/10.1016/j.socscimed.2017.01.011>
- 8) Creswell, J. W., & Poth, C. N. (2018). *Qualitative Inquiry & Research Design* (Fourth). SAGE Publications, Inc.
- 9) Dhiyaurrahman, M. Z., Santosa, I., & Primadata, A. P. (2024). Pergeseran konstruksi sosial aborsi dalam kajian sosiologi hukum di Purwokerto. *Jurnal Dinamika Sosial Budaya*, 26(1), 144–152. <https://journals.usm.ac.id/index.php/jdsb/page144>
- 10) Edi, F. R. S. (2016). Teori wawancara psikodiagnostik. *LeutikaPrio*.
- 11) FK-KMK UGM. (2023, March 7). Mengenal aborsi dan penanganannya.
- 12) Ganatra, B., Gerds, C., Rossier, C., Johnson, B. R., Tunçalp, Ö., Assifi, A., Sedgh, G., Singh, S., Bankole, A., Popinchalk, A., Bearak, J., Kang, Z., &
- 13) Alkema, L. (2017). Global, regional, and subregional classification of abortions by safety, 2010–14: estimates from a Bayesian hierarchical model. *The Lancet*, 390(10110), 2372–2381. [https://doi.org/10.1016/S0140-6736\(17\)31794-4](https://doi.org/10.1016/S0140-6736(17)31794-4)
- 14) Hamidah, N. N., & Amnar, Z. (2021). Hukum Abortus atau Aborsi. *The Renewal of Islamic Economic Law*, 3(1).
- 15) Hessini, L. (2007). Abortion and Islam: Policies and Practice in the Middle East and North Africa. *Reproductive Health Matters*, 15(29), 75–84. [https://doi.org/10.1016/S0968-8080\(06\)29279-6](https://doi.org/10.1016/S0968-8080(06)29279-6)
- 16) Jones, R. K., & Jerman, J. (2017). Abortion Incidence and Service Availability In the United States, 2014. *Perspectives on Sexual and Reproductive Health*, 49(1), 17–27. <https://doi.org/10.1363/psrh.12015>
- 17) Juliana, F. (2017). Faktor-faktor yang mempengaruhi perilaku aborsi pada perempuan pekerja hiburan malam.
- 18) Kementerian Sekretariat RI. (2022). Kehamilan tidak diinginkan jadi ganjalan penurunan stunting.
- 19) Khofifah I.P.W, N., & Atika, T. D. (2024). Literature review: dampak aborsi baik secara fisik maupun psikis pada remaja. *JURNAL BIMTAS: Jurnal Kebidanan Umtas*, 8(2).
- 20) Kumar, A., Hessini, L., & Mitchell, E. M. H. (2009). Conceptualising abortion stigma. *Culture, Health and Sexuality*, 11(6), 625–639. <https://doi.org/10.1080/13691050902842741>.
- 21) Malhotra, N., Nunan, D., & Birks, D. (2017). *Marketing Research: An Applied Approach* (5th ed.). Pearson.
- 22) Mashabi, S., & Kuwado, F. J. (2020, June 30). BKKBN: Kehamilan Tak Diinginkan di

- Indonesia Rata-rata 17,5 Persen. Artikel BKKBN: kehamilan tak diinginkan di Indonesia rata-rata 17,5 persen. Kompas.Com.
- 23) Norris, A., Bessett, D., Steinberg, J. R., Kavanaugh, M. L., De Zordo, S., & Becker, D. (2011). Abortion Stigma: A Reconceptualization of Constituents, Causes, and Consequences. *Women's Health Issues*, 21(3), S49–S54.
<https://doi.org/10.1016/j.whi.2011.02.010>.
- 24) Salamah, U. (2013). Depresi pasca abortus.
- 25) Santrock, J. W. (2019). *Life-span Development* (17th ed.). McGraw-Hill Education.
- 26) Sebayang, W. B., Rambe, N. L., Ritonga, F., & Halawa, F. H. (2023). Penyuluhan dampak aborsi bagi kesehatan remaja di SMP Imelda Medan. *Jurnal Ilmiah Pengabdian Kepada Masyarakat (Ji-SOMBA)*, 3(1), 23–27. <http://jurnal.uimedan.ac.id/index.php/Ji-SOMBA>
Journal homepage: <http://jurnal.uimedan.ac.id/index.php/Ji-SOMBA>
- 27) Sedgh, G., Ashford, L. S., & Hussain, R. (2016). Unmet need for contraception in developing countries: Examining women's reasons for not using a method. <http://www.guttmacher>.
- 28) Shellenberg, K. M., & Tsui, A. O. (2012). Correlates of perceived and internalized stigma among abortion patients in the USA: An exploration by race and Hispanic ethnicity. *International Journal of Gynecology & Obstetrics*, 118(S2).
[https://doi.org/10.1016/S0020-7292\(12\)60015-0](https://doi.org/10.1016/S0020-7292(12)60015-0)
- 29) Suryani, L. (2021). Faktor-Faktor Pendorong dan Praktik Aborsi di Indonesia. *JSGA: Journal Studi Gender Dan Anak*, 8(2).
- 30) UNFPA. (2022). Nearly half of all pregnancies are unintended—a global crisis.
- 31) WHO. (2024). Abortion.
- 32) Widowati. (2021). Tindakan aborsi dalam sudut pandang hukum dan kesehatan di Indonesia. *Jurnal Fakultas Hukum Universitas Tulungagung*.
- 33) Wijayati, M. (2015). Aborsi akibat kehamilan yang tak diinginkan (KTD): Kontestasi antara pro-live dan pro-choice. In *ANALISIS: Jurnal Studi Keislaman* (Vol. 15, Issue 1).
- 34) Windrawan, T. (2016). Serba-serbi ngidam. *Kata Hati*.
- 35) World Health Organization. (2022). Abortion care guideline.
- 36) YKP. (2020, May 8). Kehamilan yang tidak diinginkan pada remaja.

Blended Learning in the Eyes of Students: Between Offline, Online Synchronous, and Asynchronous

Clara Evi Candrayuli

Citraningtyas Universitas Pembangunan Jaya, Indonesia E-mail:

clara.citraningtyas@upj.ac.id

Abstract

This study aims to understand how students respond to blended learning methods, by trying to capture their views on their first experience enrolled in a blended class in the context of Indonesian higher education. Data were collected from 93 respondents, consisting of 62 female and 31 male students, through a survey on the effectiveness of various blended learning methods: face-to-face meetings, synchronous online (Zoom), asynchronous learning through the Learning Management System (LMS), and asynchronous through Coursera. Although 81 respondents (87.1%) stated that face-to-face meetings helped them understand the material the most, the highest attendance rate was recorded in asynchronous sessions through the LMS. This could be an important signal that asynchronous learning has great potential to be further developed as part of a future blended learning strategy. However, all forms of online learning, both synchronous and asynchronous, still present their own challenges, especially related to unstable internet connections. These findings provide a rich initial picture of students' preferences and experiences in blended learning, which can serve as a basis for future improvements in curriculum design and teaching methods.

Strategic Statelessness and Cultural Resistance: Zomia, the Baduy Dalam, and the Politics of Autonomy in Southeast Asia

Sebastian Satrio Darminto Ritsumeikan Asia Pacific University, Japan

se24s6tb@apu.ac.jp

Abstract:

This paper explores statelessness as a deliberate political and cultural strategy, drawing on James C. Scott's "The Art of Not Being Governed" and the case of the Baduy Dalam in Indonesia. Scott's concept of Zomia reframes upland Southeast Asian societies as communities that actively evade state incorporation through geographic, cultural, and religious means. The Baduy Dalam, though geographically embedded within the Indonesian state, maintain autonomy through strict adherence to adat (customary law), ecological stewardship, and spiritual obligations—rejecting modern technology, formal education, and state documentation while relying on the Baduy Luar as intermediaries.

The paper addresses critiques of romanticizing statelessness, acknowledging its vulnerabilities, such as restricted access to healthcare and gender inequities, while situating it as a calculated choice to avoid exploitation and cultural erasure. It argues that Indonesia's recognition of the Baduy Dalam reflects a strategy of selective disengagement rather than forced integration. Ultimately, the study challenges developmentalist assumptions that equate progress with state incorporation, framing statelessness as both a cultural safeguard and a political resource in contemporary struggles for Indigenous sovereignty and self-determination.

Keywords: Statelessness, Zomia, Baduy Dalam, Southeast Asia, Indigenous Autonomy.