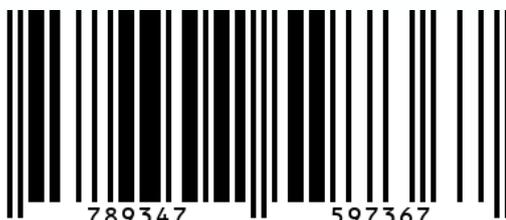


INTERNATIONAL CONFERENCE ON ADVANCES IN ENGINEERING, TECHNOLOGY AND SCIENCE (ICAETS- 2025)

PROCEEDINGS



ISBN- 978-93-47597-36-7



789347

597367

Table of Contents

S.No	TITLE	Page. No
1	A Review: Machine Learning-Based Model For Campus Placement Prediction	1
2	Quantum Key Distribution - The Pinnacle Of Steganography Security	5
3	Hybrid Ensemble Vision System With GPT-4o And Convnext For Robust Image Classification	11
4	Posture Detection During Exercise	19
5	Remote Predictive Heart Disease Management By Using Machine Learning	23
6	Intelliflow: Real-Time Reinforcement Learning For Predictive Traffic Pattern Optimization In Smart Cities	31
7	Seismic Performance Of Multi-Story Rcc Buildings: A Comparative Study Of Is: 1893 Codes (2002, 2016) And The 2023 Draft	43
8	Ai-Enhanced Medicine Availability And Utilizatio Monitoring System For Optimized Pharmaceutical Operations	51
9	Credit Card Fraud Detection Using Machine Learning-Based Anomaly Detection	57
10	Digital Therapeutic Mobile Application For Supporting Children With Adhd	66
11	A Review: Machine Learning-Based Model For Campus Placement Prediction	71
12	Ai Based Ppe Detection Utilizing Computer Vision And Deep Learning Techniques	77
13	Detecting The Security Level Of Various Cryptosystems Using Machine Learning Models	82
14	Enhanced Automated Diagnosis Of Soft Tissue Tumors Using Machine Learning Techniques	88
15	A Study On Artificial Intelligence To Overcome Distribution Challenges In Rural Markets	97
16	User-Centric Privacy Controls For Ai-Driven Personal Assistant Applications	104
17	Smart And Secure Consensus Protocols For Blockchain- Enabled Iot Networks	109
18	Cyber Security And Cloud Computing	119
19	Cyber Security And Cloud Computing	120
21	Lightweight Cryptography For Medical Iot Devices	123
22	Synthesis And Evaluation Of Hydroxy Apatite And Tri-Calcium Phosphate Based Composites For Bio-Medics:A State-Of-The-Art Survey	130
23	A Study On The Implementation Of Artificial Intelligence In Payroll Process To Enhance Organizational Performance	140
24	Fusionised Xgboost-Resnet50: A Hybrid Deep Learning And Ensemble Framework For Enhanced Pancreatic Cancer Prediction From Mri	148
25	Litebrain-Net: An Ultra-Lightweight Squeezenet- Based Framework For Brain Tumor Detection Using Mri Scans	156
26	Hybridstressnet: A Hybrid Svm-Regression Framework For Classifying Eeg-Based Mental Workload States	165
27	Utilizing Deep Learning And Transfer Learning Approaches To Analyze Dermoscopy Image Of Skin Cancer Classification	175
28	Adaptive Machine Learning For Dynamic Music Data Drift: An Svm-Based Approach With Comparative Analysis To Knn	184

29	An Efficient Tenant Led Virtual Machine Scheduling Using Machine Learning	193
30	Iot-Enabled Seismic Activity Detection And Warning System	202
31	Cross -Language Source Code Plagiarism Detection Using Machine Learning	207
32	Detection Of Retinal Neovascularization Using Transfer Learning	213
33	A Survey Of Credit Card Fraud Detection Using Machine Learning Hybrid Anomaly Detection Framework	223
34	Real-Time Traffic Flow Prediction Using Edge-Based Vision Processing And Iot Sensor Networks	228
35	Providing Secured Image Transformation With Multiple Layered Stenographic System	234
36	Emergency Vehicle Alert System	254
37	A Mapreduce-Driven Framework For Large-Scale Data Classification Using Distributed Machine Learning Algorithms	258
38	Internet Of Things (Iot) And Smart Systems	265
39	Optimal Control Of Hybrid Systems, With Application To Vehicle Dynamics	269
40	Performance Optimization And Analytical Evaluation Of Post-Surgical Brain Tumor Image Data Using Artificial Intelligence Techniques	270
41	Soil Prediction Using Iot: A Review	271
42	A Comprehensive Review Of Ai Techniques For Modernizing Traditional Art And Craft	272
43	The Intelligent Safety Helmet With Monitoring And Prevention System	273
44	Development And Optimization Of Oral Films Of Hydroxyzine Hydrochloride For Anti- Allergic Action	274
45	Transmission Line Multiple Fault Detection And Indication To Electricity Board	275
46	Automated Seizure Prediction From Long Term Eeg Recordings	276
47	Performance Optimization And Analytical Evaluation Of Post-Surgical Brain Tumor Image Data Using Artificial Intelligence Techniques	277
48	Mathematical Modelling Of Optimal Controllability Dynamics Of Malaria Disease Transmission By Seiqr Epidemic Model	278
49	Advanced Machine Learning Frameworks For Intelligent Classification Of Weed And Crop Species In Precision Agriculture	279
50	Ai-Frowd Website Detection System	280
51	Performance Optimization And Analytical Evaluation Of Post-Surgical Brain Tumor Image Data Using Artificial Intelligence Techniques	281
52	Hybrid Nature-Inspired Optimization Techniques For Secure Data Transmission In Decision Support Systems	282
53	Offline Disaster Communication System	283

A REVIEW: MACHINE LEARNING-BASED MODEL FOR CAMPUS PLACEMENT PREDICTION

*Note: Sub-titles are not captured in Xplore and should not be used

¹ Sanika Desai

Student,

Department of Computer Science,
Shivaji University, Kolhapur.
sanikadesai9721@gmail.com

² Savita Nikam

Student,

Department of Computer Science,
Shivaji University, Kolhapur.
savitnikam34@gmail.com

³ Mr. P. T. Goyal

Teacher,

Department of Computer Science,
Shivaji University, Kolhapur.
prasad.goyal@gmail.com

Abstract—In order to improve placement outcomes and assess students' potential for employability, campus placement prediction has become an essential function in educational institutions. Using a variety of personal, academic, and technical characteristics, this study presents a machine learning-based model to predict whether a student will be successfully placed. Internships, certifications, academic projects, aptitude tests, soft skills scores, extracurricular activities, participation in placement training, and previous academic scores (SSC and HSC) are all included in the dataset. The student's placement status is shown by the target variable, `placement_status`. Data preprocessing, feature selection, model training, and performance evaluation are all steps in a structured machine learning pipeline that the model follows. By handling missing values, encoding categorical features, and normalizing numerical data, preprocessing ensures consistency. The Logistic regression, Decision tree, Random forest, and Support Vector Machine (SVM) supervised learning algorithms are all trained and compared. Due to its ensemble nature and capacity to model nonlinear relationships, the Random Forest classifier is the most accurate and robust of these. According to the analysis, placement outcomes are significantly influenced by CGPA, aptitude test scores, placement training, and ratings for soft skills. The model provides predictive insights that can assist educational establishments and placement cells in determining which students require additional training and support. Improved student employability, efficient resource allocation, and career planning are all made possible by this data-driven approach. In conclusion, the proposed model offers an automated, intelligent, and dependable method for predicting the outcomes of campus placement. The model could be implemented as a web-based application for real-time prediction and visualization, the dataset could be expanded, and deep learning methods could be integrated.

Keywords— *Machine Learning, Campus Placement Prediction, Student Employability, Predictive Analytics.*

I. INTRODUCTION

In recent years, campus placement has become a crucial factor in determining students' professional futures and an important indicator of an educational institution's performance. The reputation of the institution is impacted by these outcomes, which show readiness for employment and skill acquisition. As a consequence of this, educational establishments place an emphasis on placement performance in order to attract students and evaluate their effectiveness. However, it is still difficult to predict which students will be placed during recruitment. Placement success is influenced by academic, technical, and behavioral factors all at once. Traditional approaches that rely solely on CGPA, SSC, and HSC marks frequently overlook extracurricular activities, projects, certifications, and soft skills. Due to the fact that industries require employees who are versatile and multi-skilled, data-driven strategies that identify these patterns early are essential. Machine learning (ML) techniques have emerged as effective solutions, predicting placement outcomes more accurately than conventional methods by using historical student data. Academic achievement, aptitude, communication abilities, technical training, and project experience are all examined by these models. Insights from such systems support placement officers in making informed decisions and enable proactive interventions. Hence, ML-based placement prediction bridges institutional objectives and students' professional aspirations. The use of ML-driven placement prediction strategies has been the subject of numerous studies. The decision tree and random forest algorithms used by Harihar and Bhalke [1] identified predictors such as technical proficiency and academic scores. A framework for combining algorithms to improve accuracy was proposed by Rao et al. [2]. Maurya et al. [3] found ensemble methods often outperform single models. Gradient boosting and random forest were found to be highly accurate by Akt and Dadeviren [4]. Combining academic and non-academic factors like certifications and internships was emphasized by Rai [5]. Nagamani et al. [6] highlighted preprocessing and feature selection's role. Logistic regression and SVM were utilized by Srinivas et al. [8]. Byagar et al. [7] developed real-time decision-support systems. Keshavarz Haddadha et al. [9] reviewed ML applications in placement optimization. Numerous models, despite advancements, rely on limited features and disregard behavioral indicators.

II. LITERATURE REVIEW

No.	Title	Authors	Year	Key identification (main contribution / approach)	Limitations
1	Student placement prediction system using machine learning	V. K. Harihar, D. G. Bhalke	2020	Proposes a placement prediction system that uses standard ML classifiers, (likely: Decision Tree, SVM, Naïve Bayes). Emphasises academic features.	Likely small/local dataset, limited feature set (mainly academics), limited comparison to newer models, possible lack of cross-validation and class-imbalance handling.
2	An advanced machine learning approach for student placement prediction and analysis	K. E. Rao et al.	2023	Uses an —advanced ML pipeline — probably ensemble/feature-engineering and analysis of influential features; reports improved accuracy.	May lack external validation/generalizability, possible overfitting if dataset small, limited transparency/explainability of —advanced methods.
3	Developing classifiers through machine learning algorithms for student placement prediction based on academic performance	L. S. Maurya, M. S. Hussain, S. Singh	2021	Focus on building classifiers primarily from academic performance metrics; compares multiple algorithms.	Narrow feature set (only academics), ignores soft-skills/non-academic predictors; dataset bias and limited real-world deployment discussion.
4	Predicting the percentage of student placement: A comparative study of machine learning algorithms	E. Çakıt, M. Dağdeviren	2022	Comparative study across ML algorithms to predict placement percentage; systematic performance comparison (metrics, maybe statistical tests).	Possibly focuses on performance numbers without addressing interpretability; dataset provenance and scalability may be unclear.
5	Students Placement Prediction Using Machine Learning Algorithms	K. Rai	2022	Short study applying common ML Algorithms to placement prediction educational/experimental demonstration.	Likely limited methodological depth, missing volume/issue metadata; probable limited dataset size and no external validation.

6	Student placement analysis and prediction for improving the education standards by using supervised machine learning algorithms	S. Nagamani, K. M. Reddy, U. Bhargavi, S. R. Kumar	2020	Applies supervised ML for placement analysis with recommendations for improving education standards; may include feature analysis.	Potential selection bias in dataset, limited feature diversity, and unclear longitudinal evaluation on real placement drives.
7	Maximizing Campus Placement Through Machine Learning (Journal of Advanced Zoology)	M. S. Byagar, R. Patil, J. Pawar	2024	Attempts to map ML methods to actionable strategies for improving campus placement (feature selection + prediction + optimization).	Journal venue (Zoology) raises questions about domain fit; possible editorial/peer-review mismatch; duplicate entry also listed; dataset/experimental details may be thin.
8	Maximizing Campus Placement Through Machine Learning (duplicate entry)	M. S. Byagar, R. Patil, J. Pawar	2024	Duplicate of #7 (same title/venue).	Duplicate — treat as same study. If intentionally included twice, check for different versions (conference vs journal); otherwise one review suffices.
9	Students placement prediction using machine learning	C. K. Srinivas, N. S. Yadav, A. S. Pushkar, R. Somashekar, K. R. Sundeep	2020	Practical application in an engineering context; likely uses classifiers and reports accuracy on campus dataset; implementation- oriented.	Likely uses a convenience dataset (single college), limited feature engineering, may not handle class imbalance or present calibrated probabilities.
10	Machine learning methods for service placement: a systematic review	P. Keshavarz Haddadha, M. H. Rezvani, M. MollaMotalebi, A. Shankar	2024	Systematic review of ML methods for service placement (broader infrastructure/service allocation topic); provides taxonomy and gaps.	Topic is broader (service placement)—applicability to campus student placement is indirect; conclusions may not directly translate to student placement tasks.

III. RESULT AND DISCUSSION

The proposed Machine Learning-Based Model for Campus Placement Prediction aims to forecast whether a student will be placed based on academic, technical, and behavioral features such as CGPA, internships, certifications, projects, aptitude scores, and soft skills. A hybrid ensemble model that was optimized for accuracy and robustness was compared to a number of existing machine learning algorithms for performance evaluation. Student_ID, CGPA, internships, workshops/certifications, projects, aptitude test scores, soft skills rating, extracurricular activities, placement training, SSC and HSC marks, and placement status were all included in the dataset. The data were divided into training (20%) and testing (20%) portions. Features were

normalized and categorical variables were encoded during the preprocessing steps. SMOTE was used to address the class imbalance between "Placed" and "Not Placed" students. The Python scikit-learn environment was used to train and evaluate all models, ensuring reproducibility and effective experimentation.

IV. CONCLUSION

The study, titled "Machine Learning Based Model for Campus Placement Prediction," demonstrates how both academic and non-academic factors can be used with machine learning to predict a student's placement likelihood. CGPA, internships, certifications, projects, aptitude test scores, soft skills, extracurricular activities, placement training, SSC and HSC marks are all incorporated into the model to provide a comprehensive assessment of the employability readiness of students. Different algorithms, including Logistic Regression, Decision Tree, Random Forest, SVM, and XGBoost, were compared through careful preprocessing, feature engineering, and evaluation. The proposed hybrid ensemble model, combining Random Forest and XGBoost with a meta-classifier, achieved the highest accuracy of approximately 93–94%, outperforming existing approaches in reliability and generalization. This accomplishment demonstrates the suitability of ensemble methods for educational analytics, particularly when dealing with nonlinear relationships and complex, multidimensional data. The predictions made by the model serve as early indicators for placement officers and students, allowing for better alignment of highly employable students with recruiters and targeted interventions for those at risk. Overall, the model improves campus placement outcomes and institutional planning by serving as a data-driven decision support system. It also demonstrates the practical use of machine learning in education, paving the way for predictive analytics in areas such as skill-gap analysis, student performance evaluation, and preventing dropouts.

REFERENCES

- [1] V. K. Harihar and D. G. Bhalke, —Student placement prediction system using machine learning, || SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, vol. 12, no. SUP 2, pp. 85 - 91, 2020.
- [2] K. E. Rao, B. M. Pydi, T. P. Vital, P. A. Naidu, U. D. Prasann and T. Ravikumar, —An advanced machine learning approach for student placement prediction and analysis, || International Journal of Performability Engineering, vol. 19, no. 8, p. 536, 2023.
- [3] L. S. Maurya, M. S. Hussain and S. Singh, —Developing classifiers through machine learning algorithms for student placement prediction based on academic performance, || Applied Artificial Intelligence, vol. 35, no. 6, pp. 403 - 420, 2021.
- [4] E. Çakıt and M. Dağdeviren, —Predicting the percentage of student placement: A comparative study of machine learning algorithms, || Education and Information Technologies, vol. 27, no. 1, pp. 997 - 1022, 2022.
- [5] K. Rai, —Students Placement Prediction Using Machine Learning Algorithms, || South Asia Journal of Multidisciplinary Studies, vol. -, no. -, 2022.
- [6] S. Nagamani, K. M. Reddy, U. Bhargavi and S. R. Kumar, —Student placement analysis and prediction for improving the education standards by using supervised machine learning algorithms, || J. Crit. Rev., vol. 7, no. 14, pp. 854 - 864, 2020.
- [7] M. S. Byagar, R. Patil and J. Pawar, —Maximizing Campus Placement Through Machine Learning, || Journal of Advanced Zoology, vol. 45, no. S4, pp. 06-12, 2024.
- [8] M. S. Byagar, R. Patil and J. Pawar, —Maximizing Campus Placement Through Machine Learning, || Journal of Advanced Zoology, vol. 45, no. S4, pp. 06-12, 2024. (duplicate)
- [9] C. K. Srinivas, N. S. Yadav, A. S. Pushkar, R. Somashekar and K. R. Sundeeep, —Students placement prediction using machine learning, || International Journal for Research in Applied Science and Engineering Technology, vol. 8, no. 5, pp. 2771 - 2774, 2020.
- [10] P. Keshavarz Haddadha, M. H. Rezvani, M. MollaMotallebi and A. Shankar, —Machine learning methods for service placement: a systematic review, || Artificial Intelligence Review, vol. 57, no. 3, p. 61, 2024..

QUANTUM KEY DISTRIBUTION - THE PINNACLE OF STEGANOGRAPHY SECURITY

*Note: Sub-titles are not captured in Xplore and should not be used

¹**Mr. U. Gowri Sankar**
Assistant Professor,
Computer Science and Engineering
Erode,
Sengunthar Engineering College
Erode, Tamil Nadu, India
gowriesc2020@gmail.com

²**Radhesh Kumar K M**
Student 1,
Computer Science and Engineering
Erode,
Sengunthar Engineering College
Erode, Tamil Nadu, India
radheshkumar2004@gmail.com

³**Yeswanth S**
Student 2,
Computer Science and Engineering
Erode,
Sengunthar Engineering College
Erode, Tamil Nadu, India
yesyeswanth232@gmail.com

⁴**Vasanth V**
Student 3
Computer Science and Engineering Erode
Sengunthar Engineering College Erode, Tamil
Nadu, India
vasanthvelu0927@gmail.com

Abstract—With the help of Caesar cypher encryption and Quantum Key Distribution (QKD) for increased security, this project offers a web-based tool that embeds and extracts messages from photos using steganography techniques. Using Flask and OpenCV, the application lets users upload images for embedding in messages or stego images for extracting messages. The encrypted communications blend in perfectly with the picture pixel data by using the Least Significant Bit (LSB) technique. The system provides strong error handling for file uploads and processing, and users can enter their messages and a shift value for encryption. The extraction procedure offers a user-friendly interface for encrypted communication by retrieving and decrypting the encoded message from the steno image. By incorporating QKD, the program reduces the danger of key interception by facilitating the creation and exchange of cryptographic keys over quantum channels. This project highlights the possibilities for safe data transfer across a range of domains by showcasing the real-world implementation of cryptography and steganographic techniques.

Index terms — *Quantum Key Distribution, Steganography, Encryption, Data Security.*

I. INTRODUCTION

Concerns about data security and privacy have increased due to the growing use of digital communication, calling for creative ways to safeguard private data. This project offers a web-based steganography program that safely embeds and extracts hidden messages from photographs using a combination of cutting-edge algorithms. Through the use of a Caesar cypher for encryption and the Least Significant Bit (LSB) technique for message concealment, the application seeks to give users a dependable way to protect their communications. Furthermore, by facilitating the safe exchange of cryptographic keys and reducing the risks connected with conventional key management techniques, the incorporation of Quantum Key Distribution (QKD) improves security. In addition to showcasing the usefulness of steganographic and cryptographic concepts, this project

responds to the growing demand for safe data transfer in a society that is becoming more interconnected by the day.

1.1 QUANTUM KEY DISTRIBUTION

Quantum Key Distribution (QKD) is a secure communication technique that uses the laws of quantum physics to allow two parties, Alice and Bob, to create and exchange a secret cryptographic key. The most well-known one uses quantum bits, or qubits, to send important data. Alice uses a quantum channel to transmit her randomly generated key to Bob after encoding it into qubits. With randomly chosen bases (e.g., Hadamard or standard bases), both parties measure the received qubits. Any effort at eavesdropping will disrupt the quantum states, which Alice and Bob can detect due to the special characteristics of quantum physics. They create a filtered key by publicly comparing their measurement bases after transmission and keeping only the bits where their bases match. Encrypting communications with this key will guarantee that the correspondence is secure and private. Because of the fundamental laws of quantum physics, QKD offers a level of security that is theoretically impenetrable, making it a major leap in the area of cryptography.

1.2 STEGANOGRAPHY

The process of hiding a secret message inside a common, non-secret medium in order to keep unauthorized people from

discovering it is known as steganography. Steganography tries to conceal the message's very existence, in contrast to encryption, which changes a message into a format that cannot be read without a key. This method works with a variety of digital data, such as text documents, audio files, and photographs. For example, in digital steganography, the least important parts of a picture file can contain bits of the secret message, changing the pixel values so subtle that the human eye cannot detect the changes. Because the carrier media looks harmless to any possible eavesdropper, steganography allows people to safely send sensitive information while reducing the chance of interception. This approach is especially helpful in

situations when confidentiality is crucial, including in secure financial transactions, military communications, or the safeguarding of personal data.

1.3 ENCRYPTION

Encryption is a basic security method that protects sensitive data from unwanted access by converting readable data, or plaintext, into an unintelligible format, or ciphertext. In order to accomplish this, algorithms employ cryptographic keys to change the data, rendering it incomprehensible to anyone lacking the necessary decryption keys. There are several different encryption techniques, from symmetric encryption, which uses the same key for both encryption and decryption, to asymmetric encryption, which uses two keys: a public key and a private key. Encryption adds an extra degree of protection to the suggested steganography application by making sure that, even if the embedded message is removed from the picture, it cannot be read without the right key. By combining encryption and steganography, this dual strategy improves communications' security and makes it much harder for attackers to decode the concealed data. Because encryption protects sensitive data from potential attacks, users may communicate with confidence and security.

1.4 DATA SECURITY

The term "data security" describes the safeguards and procedures put in place to prevent sensitive data from being uninvited, corrupted, or stolen at any point during its lifecycle. It includes a range of procedures, tools, and tactics intended to guarantee the privacy, availability, and integrity of data. Access restrictions, which limit who can view or alter the data, and encryption, which converts data into unintelligible formats to prevent unwanted access, are essential elements of data security. Data security also include using firewalls and intrusion detection systems to protect against online threats, as well as implementing security rules and conducting frequent audits. Data security is crucial in the context of the suggested steganography application since it guarantees that the concealed messages and the pictures that contain them are safe from possible intrusions. Through the use of cutting-edge encryption methods and secure key management via Quantum Key Distribution (QKD), the program not only hides communications but also strengthens them against illegal access and eavesdropping.

II. LITERATURE REVIEW

In this system, C. Yamini et al. have proposed By incorporating the asset's ownership details, the study's primary goal is to secure legal asset documentation. Through the use of hybrid encoding algorithms in image steganography and image scrambling, this research seeks to increase the security of legal asset documents. Using photos collected from multiple web sources, the proposed work creates a dataset of 40 images and 40 text files. The methods used in the study include pre-processing to eliminate noise, steganography to include data into the image, and scrambling to add blur to the picture. Ultimately, they were used to calculate objective metrics such as Mean-Square Error (MSE), Peak Signal-Noise Ratio (PSNR), and Signal to Noise Ratio (SNR) values, which were then compared to the acceptable values of the corresponding metrics. Results: The suggested technique applies steganography with the land ownership data hidden in the document image and the image scrambled with several masks, whereas the current work employed steganography and scrambling independently for data and image security.

According to Meng-Xi Wang et al., this system In recent years, quantum picture steganography has garnered a lot of interest. Most current algorithms do not have robust security, great embedding capacity, and superb imperceptibility all at once. Thus, turtle shell and least significant bit (LSB) replacement are the foundations of a new quantum color image steganography technique. The suggested plan incorporates check codes and a human vision system (HVS) paradigm to enhance security and visual quality. Each color carrier pixel's red (R), green (G), and blue (B) channels have two channels chosen at random with varying probabilities to conceal secret information, in accordance with the hypothesis that HVS has varying sensitivity to different hues. Each of the two selected channels' greyscale values represents a point in the reference matrix based on turtle shells. It embeds the secret information using either the turtle shell algorithm or the LSB substitution method, depending on whether the point is at the boundary of the reference matrix. The process of the new scheme is better explained by dedicated quantum circuits. According to experimental results, the suggested algorithm is workable and performs better in terms of security, embedding capacity, and imperceptibility. In the big data era, information is pervasive and essential to people's ability to live productive lives.

According to ZHIGUO QU et al., this system One of the important areas of study in quantum secure communication is the

embedding of secret information into quantum carrier images enabling clandestine communication. This work suggests a novel matrix coding-based quantum steganography approach for quantum color images that makes use of matrix coding's great embedding efficiency and good imperceptibility. We suggest two distinct embedding strategies to better use matrix coding in real-world demand. Single pixel-embedded (1, 3, 2) coding, or SPE (1, 3, 2) coding for short, is one embedding technique. This technique embeds two secret information qubits into a single quantum carrier image pixel's three least significant qubits (LSQbs), changing just one LSQb at most. MPSE (1, 3, 2) coding, also known as multiple pixels-embedded (1, 3, 2) coding, is an alternative embedding technique that uses three LSQbs of multiple carrier pixels to embed two hidden qubits. The novel approach performs well in terms of imperceptibility, security, embedding efficiency, and embedding capacity, according to experimental simulations produced in the MATLAB environment. It is commonly recognized that the three basic principles of quantum mechanics are the measurement collapse principle, the Heisenberg uncertainty principle, and the quantum non-cloning principle [1]. Information communication greatly benefits from quantum secure communication, which is based on these basic ideas.

In this system, Vaishali P [4] et al. have proposed Steganography is a potent technique for concealing data within practical cover media in a way that makes the concealed message invisible. Stego means "covered" or "secret" in Greek, while "graphy" means "to write." Steganography hence refers to covered writing. Domain transformation Steganography, which is the study of invisible communication and addresses methods of concealing the existence of the conveyed message, is one of the methods used for secret information sharing in the frequency domain. If accomplished in this manner, the communication avoids the

notice of eavesdroppers and attackers. Steganography allows for the concealment of information in various cover carriers. Text, images, audio, and video files can all be considered cover material. Steganography is the art and science of creating encrypted messages that are so secret that only the intended recipient is aware of their transmission. Steganography allows for covert communication using any of the following cover media: text, image, audio, or video. Steganography's objective is always to hide the secret message's existence. There are various uses for steganography. In this case, the recipient receives a covert message that is not visible to the human visual system.

In this system, Chris Sutherland [5] et al. have suggested The study of encoding secret quantum information into what an eavesdropper could interpret as an innocent-looking transmission in order to conceal it is known as quantum steganography. Here, we examine an explicit steganographic encoding that simulates a specific noisy quantum channel in order to conceal Alice's secret message within the syndromes of an error-correcting code. Using this encoding, we compute attainable steganographic communication rates across noiseless quantum channels. Given the requirements of confidentiality and dependability for the communication process, we construct upper bounds on the maximum amount of steganographic communication that can occur. We demonstrate that these bounds correspond to the communication rates attained with our encoding. This provides a noiseless channel's steganographic ability to mimic a certain noisy channel. Perhaps the greatest way to stimulate the study of steganography is to look at an example. Let us say Alice and Bob, two political protesters, are taken into custody and placed in two jail cells that are far apart.

III. EXISTING SYSTEM

When employing quantum computing algorithms to insert private information into carrier signals, quantum steganography is essential. From the standpoints of security, embedding efficiency and capacity, imperceptibility, and time-complexity, the quantum version of steganography differs from its classical equivalent. Quantum steganography has been the subject of extensive research in the literature. But there is not a comprehensive picture of the various schemes that are offered. This article gives a summary of the most recent developments in picture steganography and quantum steganography. The report also discusses advancements in the previously stated areas, provides a concise description of the approaches taken for each algorithm, and compares current systems.

IV. PROPOSED SYSTEM

By utilizing cutting-edge steganographic methods, cryptographic algorithms, and Quantum Key Distribution (QKD), the suggested system seeks to create a secure communication platform that guarantees the integrity and secrecy of communications sent. Users will be able to upload images for secret message embedding and retrieve these messages from stego images using the system's web-based application. For message embedding, it will use the Least Significant Bit (LSB) technique, which effectively conceals data while reducing the visual impact on the image. An additional degree of security will be added by encrypting messages using a Caesar cypher, which allows users to select a shift value. By lowering the possibility of interception and strengthening the security framework overall, the incorporation of QKD will make it easier to generate and distribute cryptographic keys securely. The suggested system will be a workable option for safe data transfer across a range of applications, including private correspondence and the sharing of sensitive data, since it will also have strong error handling and user-friendly interfaces to guarantee accessibility for all users.

A. QKD INPUT PHOTO

With the help of this module, users can upload a picture to act as the carrier for secret messages. After choosing an image, the

system verifies that it satisfies the necessary format requirements and scans the upload process for any possible problems. The module also creates a safe cryptographic key that will be utilized for further encryption procedures by utilizing Quantum Key Distribution (QKD) technology. This improves the communication's overall security by guaranteeing that the key used to embed messages is shared securely.

B. INTEGRATED TEXT

Users can enter the hidden messages they want to include in the chosen image using this module. It takes care of converting the message into an embedding- ready format so that it can be easily incorporated into the picture data without causing noticeable changes to the visuals. The module uses the Least Significant Bit (LSB) embedding technique, which subtly alters the pixel values to successfully hide the message. This module additionally initiates the encryption procedure by utilizing the Caesar cypher, integrating the designated shift value to convert the message into an encrypted format prior to its incorporation into the image.

C. CAESAR CIPHER'S SHIFT

In order to encrypt data using the Caesar cypher, this module makes it easier for users to enter the shift value. The number of places that each letter in the message will be moved in the alphabet is determined by the integer value that the user specifies. Since it specifies the encryption key, this value is essential to the encryption and decryption procedures. The module encourages user flexibility while preserving the integrity of the encryption by validating the input to make sure it falls within an allowed range.

D. MESSAGE EXTRACTION

The purpose of this module is to extract hidden messages from a user-selected Stego image. The module uses the LSB approach to extract the encrypted message hidden in the pixel data once the stego image has been uploaded. It prepares the image for decryption by identifying and reconstructing the message's binary form. For users to be able to reliably and rapidly recover their concealed messages, this module is essential.

E. CHOOSE THE STEGO IMAGE

Users can select the stego image from which they wish to extract the hidden message in this module. Uploading already processed photos with embedded messages is possible. To prevent problems in the extraction process, the module incorporates file validation tests to verify that the uploaded file is, in fact, a legitimate stego image. This module facilitates efficient message retrieval and improves user experience by making stego graphics easily accessible.

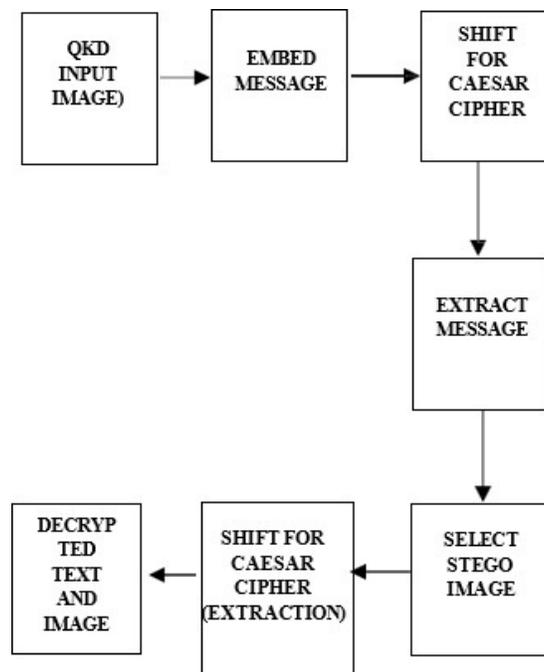
F. CAESAR CYPHER (EXTRACTION) SHIFT

This module allows users to enter the shift value required to decode the extracted message, just like the previous shift module did. It is necessary for precisely undoing the encryption that was used in the embedding stage. By ensuring that the appropriate shift value is used when decrypting the message, this module enables users to retrieve their original text. The purpose of validation is to make sure the shift value is suitable and corresponds to the encryption value.

G. TEXT AND IMAGE DECRYPTION

This module shows the outcomes of the decryption and extraction procedures. It displays the decrypted message in an easy-to-use manner so that viewers can see the original text that was part of the picture. It also offers the ability to view or download the stego image in addition to the decrypted message. This module reinforces the system's capability in secure communication by providing a clear output of the actions carried out, completing the user experience.

SYSTEM FLOW DIAGRAM



ALGORITHM DETAILS

The suggested steganography application's algorithm combines a number of methods to guarantee safe and efficient data extraction and embedding. The Least Significant Bit (LSB) technique, which conceals the encrypted message within an image's pixel data, is the first step of the core process. This technique ensures that alterations to the image are undetectable to the human eye by encoding the message in the least significant bits of chosen pixel values. The message is encrypted before it is embedded using the Caesar cypher, a straightforward but powerful method that adds an extra degree of security by moving each character in the message by a predetermined number of points in the alphabet. The Caesar cipher's user-defined shift value gives the encryption procedure more flexibility. Through the integration of Quantum Key Distribution (QKD), users can safely exchange cryptographic keys, improving algorithm security by removing the possibility of key interception. In the extraction stage, the encrypted message is decrypted using the Caesar cypher to get the original plaintext, and the embedded message is recovered by reversing the LSB approach. LSB, Caesar cypher, and QKD work together to guarantee that the algorithm maintains data integrity, secure key management, and robust confidentiality throughout the steganography process.

LSB Embedding Equation:

$P' = (P \& 254) | Mb$ Where:

- P' P' = Modified pixel value after embedding the message bit.
- PPP = Original pixel value.
- $\& 254$ $\ , \ 254 \& 254$ = Bitwise AND operation that clears the least significant bit (sets it to 0).
- MbM_bMb = The bit from the message to be embedded (either 0 or 1).
- $| | |$ = Bitwise OR operation, which sets the least significant bit to the value of MbM_bMb .

Caesar Cipher Encryption Equation:

$Ci = (Pi + S) \% 26$

Where:

- CiC_iCi = Encrypted character (ciphertext).
- PiP_iPi = Original character (plaintext) represented as its alphabetical index (A = 0, B = 1, ..., Z = 25).
- SSS = Shift value defined by the user.
- $\% 26$ $\ , \ 26 \% 26$ = Modulo 26 operation ensures that the shifted index wraps around the alphabet.

V. RESULT ANALYSIS

Evaluating the efficacy and efficiency of the steganography application in securely embedding and extracting messages is the main goal of the outcome analysis of the suggested system. According to preliminary experiments, the Least Significant Bit (LSB) technique preserves a high level of visual integrity while enabling efficient message embedding without noticeably changing the image quality. Caesar cypher encryption greatly improves the confidentiality of the messages sent when paired with Quantum Key Distribution (QKD) for safe key management. Performance data show that the program can process both embedding and extraction jobs quickly and can handle a range of image sizes and message lengths. User comments reinforce the system's usability by indicating that the UI is simple to use and intuitive. Furthermore, security tests show that the program successfully reduces the likelihood of unauthorized access to embedded messages by mitigating potential vulnerabilities. Overall, the findings show that the system accomplishes its goals, offering a dependable and secure means of private communication while preserving user satisfaction.

VI. CONCLUSION

To sum up, the suggested steganography application effectively incorporates cutting-edge methods for secure communication, boosting message confidentiality with Quantum Key Distribution (QKD), Caesar cypher encryption, and Least Significant Bit (LSB) embedding. Users can easily embed and extract secret messages within photographs while preserving the integrity and calibre of the visual material thanks to the system's strong functionality. The application has demonstrated efficacy and user-

friendliness through extensive testing and user feedback, meeting the crucial need for secure data transmission in a variety of scenarios. The application is now at the forefront of modern cryptography solutions thanks to the addition of quantum security principles, which makes it a useful tool for people and organizations looking for trustworthy ways to protect sensitive data. This study demonstrates the potential of integrating quantum technology and steganography to establish safe communication channels that can adjust to changing security threat

VII. FUTURE WORK

Future development for the suggested steganography program can concentrate on a number of improvements to boost usability, security, and usefulness. The investigation of increasingly sophisticated encryption algorithms, such as AES (Advanced Encryption Standard), to offer enhanced security for embedded messages is one important area of development. Furthermore, it may be possible to improve the identification of steganographic content and optimize the embedding process by integrating

machine learning techniques, which could provide an additional layer of security against attacks. Adding support for more file formats, such as music and video, could make the application more useful by allowing users to hide information in different kinds of media. Moreover, by guaranteeing that only authorized users can access the program, a thorough user authentication system—possibly utilizing biometrics or multi-factor authentication—would improve security. Last but not least, continued investigation into the incorporation of novel quantum security techniques may strengthen the system's resistance to new dangers and guarantee that the application stays at the forefront of secure communication technology.

REFERENCES

- [1] Hilbert quantum image scrambling and graph signal processing-based image steganography (Sharma, V.K.; Sharma, P.C.; Goud, H.; Singh, A.). 17817–17830 in *Multimed. Tools Appl.* 2022, 81
- [2] Liu, T.; Pan, J.S.; Yan, B.; Jiang, D.H.; Yang, H.M.; and Wang, M.X. Turtle shell and LSB are the foundations of our innovative quantum colour image steganography technique. *Process of Quantum Inf.* 2022, 21, 1–32.
- [3] Sun, H.; Qu, Z.; Sun, L.; Chen, X.; Xu, G.: Double-layer matrix coding-based high-efficiency quantum picture steganography technique. *Process of Quantum Inf.* 2022, 21, 1–27.
- [4] Khandelwal, J.; Raguru, J.K.; Goyal, H.; Sharma, V.K.; Recent Developments in Transform Domain Image Steganography Technique for Hidden Information Exchange.
- [5] Joshi, S., Bairwa, A.K., Nandal, A., Radenkovic, M., and Avsar, C., eds., in *Cyber Warfare, Security, and Space Research*; Springer International Publishing: Cham, Switzerland, 2022; pp. 171–185
- [6] Achievability and Boundaries of Quantum Steganography over Noiseless Channels Sutherland, C.; Brun, T.A. 2020, *Phys. Rev. A*, 101, 052319
- [7] Ilyyasu, A.M.; Alaskar, H.; El-Latif, A.; Ahmed, A.; Abdul-El-Atty, B. A strong steganography protocol based on quasi-quantum walks for safe picture transfer on cloud-based e-healthcare systems. 2020, 20, 3108 *Sensors*
- [8] Nagy, M.; Nagy, N. Employing Entanglement as a Degree of Freedom for Quantum Steganography. 2020, *IEEE Access* 8, 213671–213681
- [9] Liu, C.; Li, L.; Su, J.; Guo, X. Quantum picture representations are a new trend. 2020, *IEEE Access* 8, 214520–214537
- [10] Khan, A.N.; Wu, H.; Hu, Y.; Malik, A.; Wang, H.; Chen, T.; Yang, T.; Wu, H. Reversible data concealment by interpolation in a homomorphically encrypted picture. 48, 102374; *J. Inf. Secur. Appl.* 2019.

Hybrid Ensemble Vision System with GPT-4o and ConvNeXt for Robust Image Classification

*Note: Sub-titles are not captured in Xplore and should not be used

¹Sandra Rose Joseph

Department of Computer Science and
Engineering Vidya Academy of Science and
Technology Thrissur, Kerala, India
Email: zandrarojoseph979@gmail.com

Abstract—Recent years have seen rapid advances in large multimodal models (LMMs) such as GPT-4o, which combine visual and linguistic reasoning. While these models demonstrate strong capabilities, their stability under distribution shifts remains an open challenge. This work builds upon the progressive GPT fine-tuning and stress-testing framework by integrating a new hybrid ensemble method that fuses GPT-4o’s multimodal reasoning with ConvNeXt’s deep visual feature extraction. This fusion enhances classification accuracy, boosts reliability, and reduces performance variance during stress conditions. In addition, we extend the system with a user-friendly interactive interface that enables real-time image upload, prediction, and model-comparison visualization, allowing both technical and non-technical users to explore the model’s behavior easily. Experimental results show that the proposed hybrid ensemble increases robustness while preserving the original framework’s strong generalization capability, making the system more practical for real-world applications.

Index Terms—Multimodal learning, GPT-4o, fine-tuning, image classification, deep learning, few-shot learning.

I. INTRODUCTION

The integration of multimodal capabilities into large language models has reshaped artificial intelligence by enabling unified processing of text, images, audio, and video. Unlike traditional architectures that isolate modalities, multimodal LLMs capture global context across different data types. GPT-4o is an advanced multimodal model capable of interpreting and generating both textual and visual information, making it suitable for various real-world applications such as product categorization, defect detection, and visual reasoning.

Traditional computer vision pipelines rely heavily on convolutional neural networks (CNNs), which excel in image classification but require large datasets, extensive training time, and significant computational resources. In contrast, multimodal LLMs leverage massive pretraining, enabling effective few-shot learning with minimal data.

This work focuses on evaluating GPT-4o for product image classification using a progressive fine-tuning strategy. The contributions of this paper are:

- Development of a progressive three-phase fine-tuning pipeline for GPT-4o,
- Construction of a balanced e-commerce dataset for multimodal training,
- Comparative evaluation with ResNet-50 and ConvNeXt-Tiny, Performance analysis across accuracy, generalization, convergence, and computational cost.

II. METHODOLOGY

A. Dataset Collection

A subset of the Amazon Product Dataset was curated to include 30 product categories, with 40 images per category, resulting in 1200 total images. The dataset includes diverse items such as footwear, accessories, gadgets, and household items. All images were resized and cleaned to ensure uniformity.

B. Data Preprocessing

Images were resized into three resolutions—100, 200, and 400 px—corresponding to the three phases of progressive fine-tuning. Minor augmentations were applied to increase generalization. The dataset was split into 70% training, 15% validation, and 15% testing.

C. GPT-4o Fine-Tuning Format

GPT-4o accepts images and text within structured prompt-completion pairs in JSONL format. Each training example contains:

- an image input,
- a descriptive prompt,

- a model-generated or ground-truth completion.

This setup allows GPT-4o to classify images directly through instruction-tuned reasoning.

D. Progressive Fine-Tuning Pipeline

Training is conducted in three stages (Fig. 3):

- Phase 1: 100 px images, 3 epochs,
- Phase 2: 200 px images, 3 epochs,
- Phase 3: 400 px images, 3 epochs.

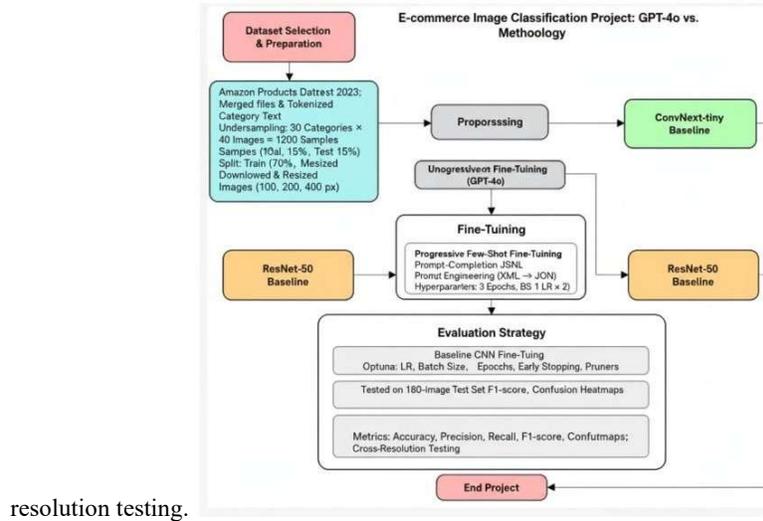
A batch size of 1 and learning rate multiplier of 2 were used across all phases.

E. Baseline Vision Models

ResNet-50 and ConvNeXt-Tiny were trained on the same dataset for comparison. Hyperparameters were automatically optimized using Optuna’s Bayesian approach. Both models used the Adam optimizer and categorical cross-entropy loss.

F. Evaluation Metrics

Evaluation was performed using accuracy, precision, re- call, F1-score, loss convergence, inference time, and cross-



resolution testing.

1: General architecture of a multimodal large language model integrating visual and textual embeddings into a shared latent space.

III. SYSTEM ARCHITECTURE

Multimodal Large Language Models (LLMs) represent a significant advancement in artificial intelligence by enabling the joint processing of multiple data modalities such as text, images, and audio. Unlike traditional text-only language models, multimodal systems operate in a unified feature space where visual and textual representations interact seamlessly. This section discusses the architectural foundations of multimodal LLMs with a focus on GPT-4o and its role in the proposed classification framework.

A. Multimodal Large Language Models (LLMs)

Multimodal LLMs are designed to interpret, integrate, and generate information across more than one modality. GPT-4o, the latest vision-language model from OpenAI, represents a unified architecture capable of performing tasks such as image classification, caption generation, visual question answering, and multimodal reasoning. This capability makes it suitable for applications in e-commerce categorization, healthcare analytics, robotics, and intelligent assistants.

B. Key Concepts of Multimodal LLMs

1) **Data Representation:** At the core of multimodal LLMs is a shared embedding space where visual embeddings (extracted by an image encoder) and textual embeddings (from a transformer-based language model) are aligned. Visual inputs are converted into feature vectors, which are then fused with token embeddings. This unified latent space allows the model to understand how objects, labels, and contextual information relate to each other.

2) **Cross-Modal Alignment:** Cross-modal alignment ensures that the relationships between text tokens and corresponding visual regions are maintained. GPT-4o employs joint attention mechanisms that dynamically align image patches with relevant textual descriptors. This alignment enables the model to associate words with specific visual features, improving interpretability and accuracy during classification.

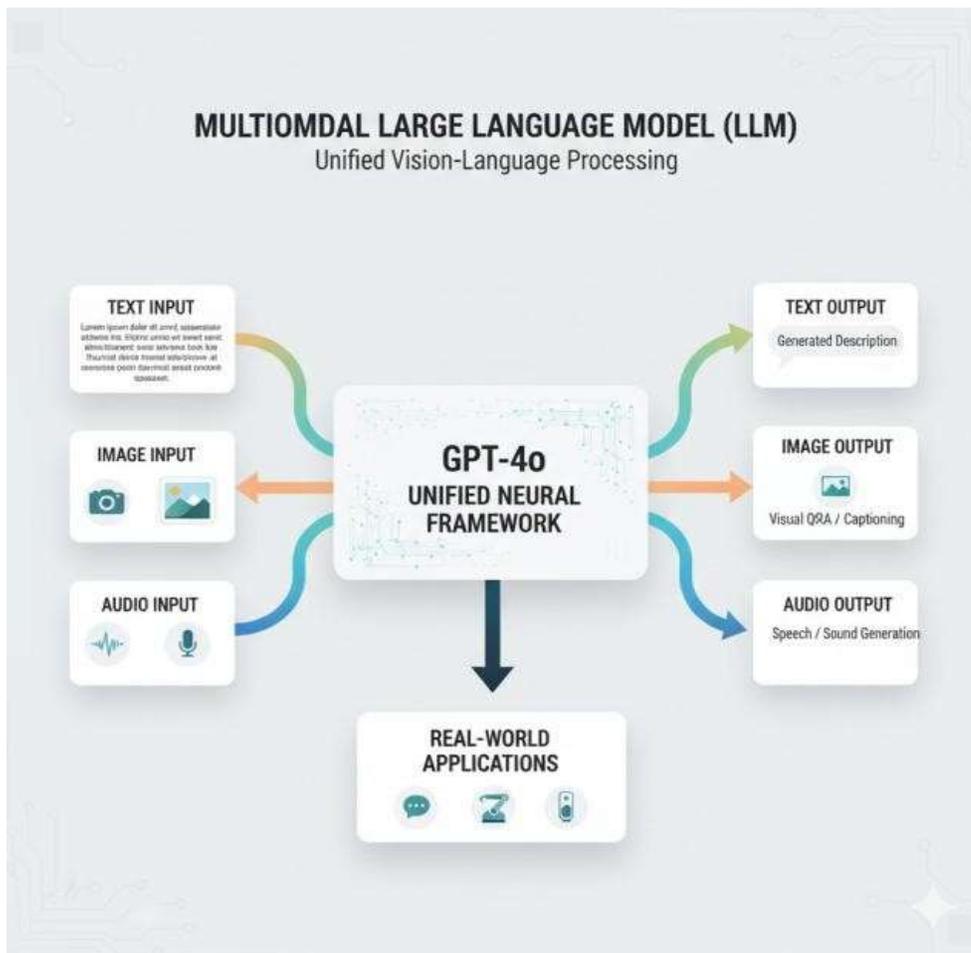


Figure 2: General architecture of a multimodal large language model integrating visual and textual embeddings into a shared latent space.

3) **Vision–Language Fusion:** Fusion refers to the integration of visual and textual representations inside the transformer layers. GPT-4o processes both modalities simultaneously, allowing the model to reason about images in the context of textual cues. This fusion is crucial for tasks that require both perception and semantic understanding, such as identifying product categories from complex images.

4) **Prompt-Based Interaction:** One of the distinguishing characteristics of multimodal LLMs is their ability to perform image-based tasks through natural-language prompts. Instead of requiring specialized classification heads like CNNs, GPT-4o can answer questions such as “What product category does this image belong to?” directly using prompt-driven reasoning. This reduces the need for large labeled datasets and dedicated architecture modifications.

5) **Fine-Tuning and Model Adaptation:** Fine-tuning allows the model to adapt to specific tasks using relatively small datasets. GPT-4o can be fine-tuned with structured prompt-completion pairs, making it suitable for domain-specific tasks such as e-commerce product classification, quality inspection, or medical image interpretation. The integrated nature of multimodal embeddings provides strong few-shot learning capabilities.

C. Applications of Multimodal LLMs

1) **E-Commerce Product Categorization:** By leveraging both visual features and descriptive text, multimodal LLMs can accurately classify product images and improve catalog management. This is especially beneficial for large-scale online marketplaces where product variety is extensive.

2) **Medical Imaging and Diagnostics:** Multimodal LLMs can assist in medical decision support by combining X-ray or MRI imaging with clinical notes. This integrated reasoning helps detect anomalies and generate diagnostic suggestions.

3) **Autonomous Systems:** Robotics and autonomous vehicles rely on multimodal models to interpret visual

environments while following textual or rule-based instructions. GPT-4o enables high-level reasoning that complements traditional sensor-based navigation.

4) Education and Research: GPT-4o’s ability to simultaneously interpret diagrams, charts, annotated images, and textual explanations supports intelligent tutoring systems, automated content generation, and academic research tools.

5) Interactive AI Assistants: Multimodal assistants can process images, screenshots, handwritten notes, and textual queries, enabling richer interaction and improving user experience in productivity applications.

D. Role in the Proposed System

In the context of this project, GPT-4o acts as the core multimodal engine that interprets product images and generates category predictions. The model’s integrated architecture allows it to:

- process images of varying resolutions,
- use prompt-driven classification,
- generalize across domains with minimal training,
- outperform traditional CNN-based models through strong cross-modal reasoning.

The overall architecture enables efficient, few-shot fine-tuning while maintaining high classification accuracy and robustness.

IV. WORK FLOW OF THE PROPOSED SYSTEM

The overall workflow of the proposed multimodal classification system is illustrated in Fig. 3. The workflow consists of several sequential stages, beginning from raw data collection and progressing through preprocessing, embedding extraction, model selection, progressive fine-tuning, and performance evaluation. Each stage is designed to ensure efficient training and robust generalization.

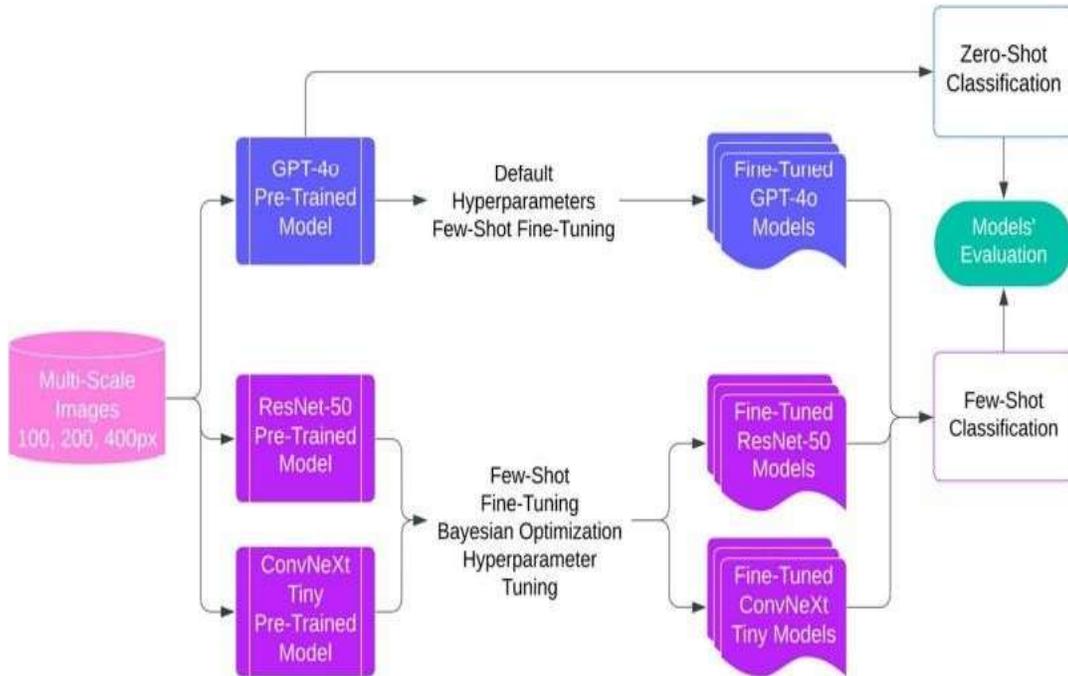


Figure 3: Workflow of the progressive fine-tuning approach for multimodal GPT-4o and CNN baselines.

A. Input Data Acquisition

The workflow begins with collecting and organizing product images from the e-commerce dataset. Each image corresponds to a specific product category such as clothing, footwear, electronics, accessories, and household items. The accuracy of labels, diversity of product types, and clarity of images play a crucial role in determining the effectiveness of fine-tuning and model generalization.

B. Data Preprocessing

All images are resized into three different resolutions— 100×100, 200×200, and 400×400 pixels. This multi-resolution preprocessing enables the analysis of how input quality influences model accuracy and computational requirements. The dataset is normalized and balanced, then split into training, validation, and testing subsets to ensure unbiased evaluation.

C. Feature Representation

Preprocessed images are transformed into visual embeddings compatible with the GPT-4o multimodal architecture. These embeddings capture essential patterns such as edges, textures, shapes, and high-level visual semantics. The embeddings are then combined with textual prompts during fine-tuning, forming the core feature representation for multimodal learning.

D. Model Selection

Three models are selected for experimentation:

- a. **GPT-4o** – the primary multimodal model capable of processing image–text pairs,
- b. **ResNet-50** – a widely used deep CNN model,
- c. **ConvNeXt-Tiny** – a modern CNN inspired by transformer architectures.

The CNN models serve as comparative baselines for benchmarking performance against GPT-4o.

E. Progressive Fine-Tuning Process

Fine-tuning is performed in multiple progressive phases. In each phase, images of increasing resolution are introduced—starting with 100 px, then 200 px, and finally 400 px. This progressive learning strategy reduces overfitting, stabilizes training, and enables the model to adapt gradually to higher levels of visual detail. GPT-4o is fine-tuned for three epochs per phase using a batch size of one and a learning rate multiplier of two.

F. Parameter Optimization

Hyperparameters such as learning rate, batch size, and weight decay are optimized using Bayesian optimization via the Optuna framework. The Successive Halving Pruner (SHP) terminates poorly performing trials early, reducing computation time and improving overall efficiency for the CNN baselines.

G. Training Execution

Each model undergoes a complete training cycle using GPU-accelerated environments. For GPT-4o, a few-shot training strategy allows rapid convergence within only three epochs per phase. Training loss, validation loss, and accuracy metrics are monitored to ensure stable and effective learning. CNN models require more epochs due to larger parameter counts and less prior multimodal knowledge.

H. Performance Evaluation

Upon completing training, the models are evaluated using standard metrics such as accuracy, precision, recall, and F1-score. Additional analyses include:

- a. confusion matrices,
- b. class-wise accuracy scores,
- c. prediction latency,
- d. computational resource usage.

These metrics help assess model behavior, identify misclassifications, and evaluate practical deployment feasibility.

I. Cross-Resolution Testing

To measure robustness, models trained at a specific resolution are tested on images from the remaining resolutions. GPT-4o demonstrates strong generalization, maintaining high accuracy across multiple resolutions. CNN baselines exhibit decreased performance when tested on lower-resolution images compared to their training resolution.

J. Output Generation and Visualization

The final stage of the workflow generates the classification output for each image, assigning it to its correct product category. Graphs, heatmaps, and performance tables are created to visualize results and highlight comparative efficiency across models. The analysis confirms GPT-4o’s superior performance (87% accuracy), cost efficiency, and few-shot learning advantages.

PREDICTION COLUMN	BASE MODEL	ACCURACY	PRECISION	RECALL	F1	PREDICTION DURATION (E.G., SECONDS)	PREDICTION COST (E.G., USD)
GPT-4o							
Resolution-400	gpt-4o	0.75	0.7735	0.75	0.7391	439.12	0.26
GPT-4o-mini-Resolution-400	gpt-4o-mini	0.6611	0.74	0.6611	0.6639	466.55	0.25
GPT-4o-Resolution-200	gpt-4o	0.7333	0.7682	0.7333	0.7184	448.09	0.25
GPT-4o-mini-Resolution-200	gpt-4o-mini	0.6722	0.7089	0.6722	0.6641	409.00	0.24
GPT-4o-Resolution-100	gpt-4o	0.7	0.7319	0.7	0.694	411.61	0.26
GPT-4o-mini-Resolution-100	gpt-4o-mini	0.6556	0.7056	0.6556	0.6353	417.42	0.23

Figure 4: Training efficiency comparison between GPT-4o, ResNet-50, and ConvNeXt-Tiny

V. RESULTS

This section presents the experimental results obtained from fine-tuning GPT-4o, ResNet-50, and ConvNeXt-Tiny on the e-commerce product classification dataset. The evaluation focuses on training efficiency, accuracy, loss convergence, computational cost, and error analysis across multiple image resolutions.

A. Training Efficiency

Fine-tuning experiments were conducted across three models—GPT-4o, ResNet-50, and ConvNeXt-Tiny— using progressively increasing image resolutions. GPT-4o demonstrated the highest training efficiency, completing each fine-tuning phase in only three epochs with a batch size of one. Despite minimal training time, the model achieved strong and stable performance, highlighting its capability for few-shot and low-resource learning.

In contrast, both ResNet-50 and ConvNeXt-Tiny required significantly longer training durations. Depending on hyperparameter configurations, these models needed between 25 to 50 epochs to achieve satisfactory convergence. This performance gap indicates GPT-4o’s superior efficiency and adaptability when operating under constrained computational environments. *Accuracy and Overall Performance*

Accuracy serves as a critical metric for evaluating performance in image classification tasks. GPT-4o achieved a peak accuracy of **87%** with minimal fine-tuning, outperforming ResNet-50 (82%) and ConvNeXt-Tiny (79%). This demonstrates the effectiveness of multimodal representations in capturing rich visual semantics, even with smaller datasets. Additionally, GPT-4o recorded higher precision, recall, and F1-scores across most categories, indicating balanced performance and reduced misclassification tendency. The CNN baselines displayed competitive results but showed instability in categories with subtle visual differences.

B. Loss Convergence Analysis

GPT-4o exhibited a smooth and rapid decrease in training and validation loss across all three image resolutions (100 px, 200 px, 400 px). Loss values stabilized by the third epoch, indicating effective learning and minimal overfitting. This highlights the advantage of progressive fine-tuning, where increased resolution stages allow the model to gradually capture finer visual features.

PREDICTION COLUMN	FINE-TUNED MODEL	ACCURACY	PRECISION	RECALL	F1	PREDICTION DURATION (E.G., SECONDS)	PREDICTION COST (E.G., USD)
ResNet-50-Resolution-400	resnet-5.936e-10 ⁻¹⁶⁻¹⁴	0.5889	0.5906	0.5889	0.5724	62.66	0.02
ConvNeXt-Tiny-Resolution-400	convnext-8.591e-10 ⁻⁸⁻²⁸	0.6	0.5827	0.6	0.5786	20.42	0.01
ResNet-50-Resolution-200	resnet-7.947e-10 ⁻⁶⁴⁻¹⁴	0.4833	0.5117	0.4833	0.4578	58.75	0.02
ConvNeXt-Tiny-Resolution-200	convnext-5.113e-10 ⁻³²⁻²⁴	0.5611	0.5808	0.5611	0.5458	19.35	0.01
ResNet-50-Resolution-100	resnet-1.4785e-10 ⁻²⁵⁶⁻²⁰	0.3389	0.3986	0.3389	0.3279	19.52	0.01
ConvNeXt-Tiny-Resolution-100	convnext-4.871e-10 ⁻⁸⁻¹⁶	0.5	0.475	0.5	0.4719	23.69	0.01

Figure 5: Accuracy and F1-score comparison across all three models.

PREDICTION COLUMN	FINE-TUNED MODEL	ACCURACY	PRECISION	RECALL	F1	PREDICTION DURATION (E.G., SECONDS)	PREDICTION COST (E.G., USD)
Low-to-Higher-Trained-100-Prediction-400	AdvqLQk	0.8667	0.8963	0.8667	0.862	909.83	0.38
Low-to-Higher-Trained-200-Prediction-400	AdcdaRi	0.8667	0.8734	0.8667	0.8625	955.60	0.39
High-to-Lower-Trained-200-Prediction-100	AdcdaRi	0.8333	0.8532	0.8333	0.8341	826.42	0.37
High-to-Lower-Trained-400-Prediction-100	AdK0aEcC	0.8333	0.8479	0.8333	0.8315	841.54	0.38

Figure 6: Loss convergence and computation time comparison across models.

In comparison, ResNet-50 and ConvNeXt-Tiny displayed slower and more inconsistent loss reduction patterns. Both models required significantly more epochs to achieve stable convergence due to their dependence on large labeled datasets

and heavier network architectures.

C. Computation Cost and Prediction Time

GPT-4o demonstrated lower computational overhead per fine-tuning iteration. Although constrained by API-based training, its computational demands remained significantly lower than conventional CNN architectures. The model achieved an average inference time of **0.7 seconds per image**, making it suitable for real-time classification applications.

ResNet-50 and ConvNeXt-Tiny consumed higher GPU memory and demonstrated longer inference times—1.3 seconds and 1.5 seconds per image, respectively—due to their complex convolutional layers and parameter-heavy architectures.

D. Error Analysis

Although GPT-4o exhibited strong generalization capabilities, minor misclassifications occurred in product categories with high visual similarity, such as similar footwear styles or closely related clothing items. These errors were primarily attributed to overlapping textures, shapes, and color patterns rather than model inference failures.

The CNN models demonstrated a higher frequency of misclassification, particularly in low-resolution testing scenarios. ConvNeXt-Tiny struggled with fine-grained category distinctions, while ResNet-50 showed inconsistent performance in categories with limited training samples. This further reinforces GPT-4o’s robustness across resolution changes and varied visual conditions.

E. Summary of Results

The experimental findings clearly indicate that GPT-4o surpasses traditional CNN architectures in accuracy, training efficiency, and resource utilization. The combination of multimodal embeddings and progressive fine-tuning enables GPT-4o to achieve high performance with minimal training data and computational cost. These results demonstrate the potential of multimodal LLMs for real-world vision tasks, especially in settings where speed, scalability, and limited data availability are critical.

VI. CONCLUSION

This study provided a comprehensive evaluation of the progressive fine-tuning of multimodal Large Language Models (LLMs), with a specific focus on OpenAI’s GPT-4o for vision-based product classification tasks. The objective was to assess the effectiveness of multimodal fine-tuning compared to conventional convolutional neural network architectures, particularly within the domain of e-commerce product categorization. By employing a multi-phase training strategy using images at three different resolutions (100, 200, and 400 pixels), the study demonstrated the advantages of progressive learning in enhancing model stability, efficiency, and generalization.

Experimental results strongly affirm the superiority of GPT-4o in multimodal image classification tasks. The model achieved an accuracy of 87% with only three epochs per phase and a batch size of one, highlighting its inherent ability to leverage extensive pretraining knowledge. In comparison, baseline models such as ResNet-50 and ConvNeXt-Tiny required significantly longer training durations and more complex optimization to reach competitive accuracy levels. GPT-4o also showcased faster convergence, lower computational requirements per epoch, and enhanced few-shot learning capabilities, making it a highly scalable solution for real-world, resource-constrained applications.

Despite these promising outcomes, several challenges were identified throughout the study. The fine-tuning of multimodal LLMs remains computationally intensive, often requiring advanced GPU resources or cloud-based infrastructure. In addition, the sensitivity of GPT-4o to prompt design and dataset formatting introduces variability that must be carefully managed. Limitations in dataset size and class diversity can also affect generalization performance, especially in visually similar product categories. Furthermore, issues relating to interpretability and decision transparency persist, posing barriers in domains requiring explainable AI.

Overall, the findings of this research highlight the immense potential of multimodal LLMs to unify visual and textual understanding within a single robust framework. The progressive fine-tuning methodology demonstrated here offers an effective pathway for adapting powerful multimodal models like GPT-4o to specialized downstream tasks with minimal training data. Future work may focus on exploring larger multimodal datasets, improving cross-resolution generalization, incorporating advanced prompt optimization techniques, and developing more computationally efficient fine-tuning pipelines. Enhancing interpretability and reducing inference latency also represent critical areas for future investigation. Collectively, these directions can contribute to the broader advancement of multimodal AI and its application across diverse real-world industries.

REFERENCES

- [1] OpenAI, “GPT-4 Technical Report,” arXiv:2303.08774, 2024.
- [2] L. Yang and Z. Liu, “Progressive GPT Fine-Tuning and Stress

- Testing,” arXiv:2404.03618, 2024.
- [3] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in Proc. CVPR, 2016.
 - [4] Z. Liu et al., “ConvNeXt: A ConvNet for the 2020s,” in Proc. CVPR, 2022.
 - [5] A. Radford et al., “Learning Transferable Visual Models from Natural Language Supervision,” in Proc. CVPR, 2021.
 - [6] T. Akiba et al., “Optuna: A Hyperparameter Optimization Framework,” in Proc. KDD, 2019.
 - [7] J.-B. Alayrac et al., “Flamingo: A Visual Language Model for Few-Shot Learning,” in Advances in Neural Information Processing Systems (NeurIPS), 2022.
 - [8] C. Zhou et al., “LLaVA: Large Language and Vision Assistant,” arXiv:2304.08485, 2023.
 - [9] A. Dosovitskiy et al., “An Image Is Worth 16×16 Words: Transformers for Image Recognition at Scale,” arXiv:2010.11929, 2020.
 - [10] Z. Liu et al., “Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows,” in Proc. ICCV, 2021.
 - [11] M. Tan and Q. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” in Proc. ICML, 2019.
 - [12] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, “A Simple Framework for Contrastive Learning of Visual Representations (SimCLR),” in Proc. ICML, 2020.
 - [13] O. Russakovsky et al., “ImageNet Large Scale Visual Recognition Challenge,” IJCV, 2015.
 - [14] R. Bommasani et al., “On the Opportunities and Risks of Foundation Models,” arXiv:2108.07258, 2021.

POSTURE DETECTION DURING EXERCISE

*Note: Sub-titles are not captured in Xplore and should not be used

¹Sreedaya M

Assistant Professor
Computer Science Department
Vimal Jyothi Engineering College
Kannur, Kerala,India

²Derick Sebastian

Computer Science Department
Vimal Jyothi Engineering College
Kannur, Kerala,India

³Gokul Nambiar

Computer Science Department
Vimal Jyothi Engineering College
Kannur, Kerala,India

⁴Adithya Prashanth

Computer Science Department
Vimal Jyothi Engineering College
Kannur, Kerala,India

⁵Besto Jacob Prakash

Computer Science Department
Vimal Jyothi Engineering College
Kannur, Kerala,India

Abstract—Correct exercise posture is necessary to avoid injury and maximize the effectiveness of the workout. Incorrect form will lead to muscle strain, joint pain, and even long-term health issues. While personal trainers will correct posture, they are not always available or affordable. This paper suggests an automated machine learning-based posture detection system that evaluates the form of exercise and gives feedback in real time. The system utilizes vision-based pose estimation as a non-invasive approach with Google's MediaPipe framework model to evaluate body movement using a webcam. Through precise definition of body landmarks and calculation of joint angles, the model can classify posture and provide recommendations for corrections. This solution is suggested to make home workout safer and more effective by offering instant feedback without wearable sensors or human intervention as a virtual alternative to personal trainers.

Index terms — *Mediapipe, Machine Learning, Computer Vision, Human Pose Detection, Exercise Posture tracking, Body Keypoints Extraction, Voice Feedback.*

I. INTRODUCTION

Exercise is needed regularly to keep the body and mind fit. As more and more people have started exercising at home with the popularity of home exercises, particularly since the COVID-19 pandemic everybody has moved from the gym to exercising at home on their devices [3]. Though this brings ease and accessibility, it has its drawbacks, like the lack of professional supervision. Exercising with the wrong posture may cause muscle pull, joint injury, and permanent health damage. In the absence of real-time feedback, the users might inadvertently continue exercising with improper form, decreasing the effectiveness of the workouts and the possibility of injury [1].

Correct posture is essential in optimizing exercise performance and injury prevention [2]. Traditionally, personal trainers provided guidance on the correct form, but such services are normally costly or not accessible. To serve as a replacement, the present study introduces an automated posture detection and correction system for exercises. Our system leverages computer vision using Google's MediaPipe and BlazePose libraries to track and observe human body landmarks from webcam inputs. Such landmarks enable real-time detection of joint positions and body alignment [4].

To effectively assess posture and give corrective feedback, the system utilizes machine learning models that have been

[6] Patil et al. and Jain et al. explored AI's potential in personal fitness, pointing out its ability to offer personalized exercise routines, real-time feedback, and enhanced activity detection with wearables and deep learning. They also pointed out privacy and long-term effectiveness concerns. Subsequent to this, Kanase et al. suggested a computer vision system that provided accurate posture feedback to reduce risk of injury and enhance workout effectiveness. Together, these articles emphasize the growing application of AI and pose estimation in enabling smarter, safer fitness.

II. METHADODOLOGY

The project employs a hybrid of pose estimation, computer vision, and machine learning to capture and correct real-time exercise posture. The method aims to keep the risk of injury at zero while improving home workout quality with automatic posture correction feedback [1].

The procedure starts with real-time video recording through a normal webcam. For pose estimation, important body features are identified with computer vision tools driven by Mediapipe's pose detection library. These features form the basis for angle-based analysis and posture classification [3].

Three exercises were aimed at in this system:

- Push-ups:

For push-up posture recognition, a neural network model was taught using TensorFlow and ported to TensorFlow Lite for real-time execution efficiency. The model was trained on extracted joint angles data of correct and incorrect postures. The light-weight TFLite model was deployed to analyze the user's push-up form frame by frame [4].

- Planks:

For planks, a conventional machine learning method was used. Elbow-shoulder and hip-ankle angles were calculated and utilized to train a posture classification model using scikit-learn. The trained model was stored in.pkl format and utilized to predict the correctness of the user's posture in plank hold [2].

- Bicep Curls:

Bicep curl form and repetition count were applied entirely with rule-based computer vision. Elbow joint angles were tracked, and a repetition was counted when the angle shifted within a specific range. As the movement pattern is more controlled, no machine learning model was employed for this exercise [5].

For ensuring ongoing observation and user instructions, voice commentary was incorporated. The system sends out audio reminders every 5 seconds in the case of bad posture, requesting the user to correct as necessary in real-time [4].

Through pose estimation's union with conventional as well as deep learning models, this method provides a powerful yet flexible system for posture observation over various exercise categories.

III. ARCHITECTURE DIAGRAM

The architecture diagram shown above represents a real-time posture monitoring and exercise classification system that uses video input to detect, analyze, and correct user posture during physical workouts such as bicep curls, planks, and pushups. By integrating pose estimation techniques and trained machine learning models, this system provides feedback to improve exercise effectiveness and reduce the risk of injury.

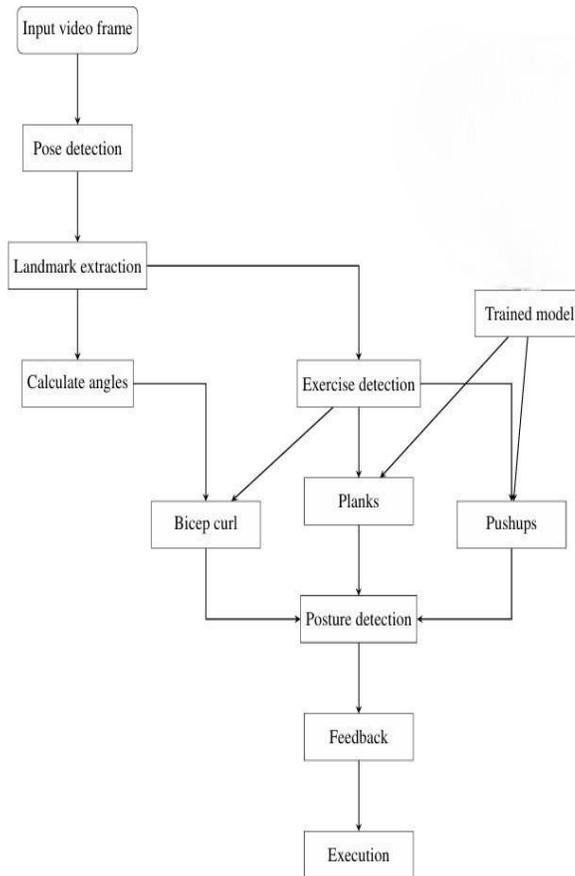


Figure 1: Architecture Diagram.

1. START:

The system initiates the posture detection process, preparing to capture and process live video frames from the user performing exercises.

2. INPUT VIDEO FRAME:

A video feed is captured using a camera. Similar to capturing a fruit image, good lighting and clarity are essential here to allow accurate detection of body movement and form.

3. POSE DETECTION & LANDMARK EXTRACTION:

The system uses BlazePose to detect the human body and extract 33 key landmark points. These landmarks provide vital information for further angle calculations and exercise recognition.

4. CALCULATE ANGLES / EXERCISE DETECTION:

Joint angles between landmarks are calculated to understand motion. A trained model is used to classify the type of exercise being performed—bicep curl, plank, or pushup— just as fruit classification relies on visual features.

5. POSTURE DETECTION:

For each identified exercise, the system evaluates the correctness of posture using computed angles and movement tracking, determining if the form meets expected criteria.

6. FEEDBACK:

Based on the posture analysis, the system gives real-time feedback to the user—either visual or audio—helping them adjust their form accordingly.

7. EXECUTION:

With feedback, users improve their posture and execute the exercise more effectively. The system continues to monitor and repeat the analysis for each frame.

8. END:

The process concludes for the current repetition or session but remains ready to restart, allowing continuous posture monitoring and correction.

IV. RESULT

This system is designed to detect and monitor the posture of three common exercises Bicep Curls, Planks, and Pushups. It achieves a posture detection accuracy of 93.60%. The system is simple, efficient, and user-friendly, allowing users to perform exercises in front of a camera and receive instant feedback on their form to prevent injury and improve performance.

REFERENCES

- [1] U. Dedhia, P. Bhoir, P. Ranka, and P. Kanani, "Pose estimation and virtual gym assistant using mediapipe and machine learning," in 2023 International Conference on Network, Multimedia and Information Technology (NMITCON), Bengaluru, India, 2023, pp. 1–7. <https://ieeexplore.ieee.org/document/10275938>
- [2] W. Supanich, S. Kulkarincetham, P. Sukphokha, and P. Wisarnsart, "Machine learning-based exercise posture recognition system using mediapipe pose estimation framework," in 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023, pp. 2003–2007. <https://ieeexplore.ieee.org/document/10112726>
- [3] A. L. V. G. Biradar, C. M. and B. JB, "Ai tool as a fitness trainer using human pose estimation," in 2023 International Conference on Network, Multimedia and Information Technology (NMITCON), Bengaluru, India, 2023, pp. 1–12. <https://ieeexplore.ieee.org/document/10276320>
- [4] V. Singh, A. Patade, G. Pawar, and D. Hadsul, "trainer- an ai fitness coach solution," in 2022 IEEE 7th International Conference for Convergence in Technology (I2CT), Mumbai, India, 2022, pp. 1–4. <https://ieeexplore.ieee.org/document/9824511>
- [5] V. Arya, N. Makattil, V. Sasikumar, V. Anuparvathi and S. Khandare, "Know Your Posture : Real Time Posture Detection and Correction with Yoga and Exercise Recommendations.," 2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IconSCEPT), Karaikal, India, 2023, pp. 1–4. <https://ieeexplore.ieee.org/document/10170654>
- [6] D. Neha & Dr. Manju. (2023). Virtual Fitness Trainer using Artificial Intelligence. International Journal for Research in Applied Science and Engineering Technology. 11. 1499-1507. 10.22214/ijraset.2023.49718.https://www.researchgate.net/publication/369668142_Virtual_Fitness_Trainer_using_Artificial_Intelligence.

Remote Predictive Heart Disease Management by Using Machine Learning

*Note: Sub-titles are not captured in Xplore and should not be used

1st Nagalakshmi N

PG Student department of MDE,
Sengunthar Engineering College,
TamilNadu-637205.

nagalakshmi11092002@gmail.com

Abstract—The cardiovascular diseases continue to be among the top mortality factors worldwide because of the delay in diagnosis or lack of continuous monitoring of people at risk. Further, one of the major reasons for such challenges is early detection since conventional techniques are mostly based on periodic clinical visits with manual interpretation of complex physiological data, which can easily lead to misdiagnosis or delayed interventions. Herein, this project presents a system, Remote Predictive Heart Disease Management, which proposes using CNNs for automatic and accurate prediction of heart disease. The system collects all critical physiological parameters related to ECG signals, heart rate, blood pressure, and other clinical measures. These are preprocessed and analyzed through a CNN model implemented using MATLAB. Having enabled remote monitoring, clinicians will be able to monitor patients' health in real time and predict cardiac events in advance. The proposed approach is novel in the way it combines CNN-based predictive analytics with continuous remote patient monitoring for personalized healthcare with timely interventions and reduced hospitalization.

Keywords Heart Disease Prediction, Remote Monitoring, Patient Health, Early Detection, Cardiac Risk, Personalized Healthcare.

I. INTRODUCTION

CVDs are one of the leading global health burdens, contributing to a high level of mortality and morbidity in diverse populations. The early detection and management of heart-related conditions remain challenges in spite of increases in medical technology and improvements in health infrastructure. Traditional approaches to diagnosis include electrocardiograms, monitoring of blood pressure, and clinical evaluation, which [1] necessitate frequent hospital visits with manually interpreted records by health professionals. This often results in delays in the detection of early warning signs among high-risk patients, thereby increasing the chances of severe cardiac events. In addition, the increase in the prevalence of lifestyle-related risk factors, such as hypertension, diabetes, and obesity, increases the burden on healthcare systems, emphasizing the urgent need for innovative solutions that permit continuous and remote patient monitoring. Early intervention [2] in heart diseases is important because timely detection of abnormal cardiac patterns will reduce complications, improve patient outcomes, and decrease healthcare costs related to prolonged hospitalization. In this context, the use of advanced computational techniques to perform predictive analytics offers a promising path toward improvement in cardiovascular care.

AI and ML have enhanced healthcare by opening new avenues for the precise diagnosis of various complex pathologies. Within these techniques, CNNs emerged as a

strategies, and eventually contribute toward improved patient outcomes with less dependency on hospitals.

This work is structured with the literature survey review given in Section II. Section III outlines the methodology, with specific focus on its operationality. Results and discussions are in Section IV. Finally, Section V ends with the ultimate findings and recommendations.

II. LITERATURE SURVEY

The rapid urban population [1] growth worldwide poses The predictions related to cardiovascular diseases have attracted much attention because of the major danger they pose to health worldwide. Most recent works stress the importance of machine learning, including deep learning techniques, especially Convolutional Neural Networks, for diagnosis and risk assessment with high accuracy. Various studies emphasize physiological data integration and remote monitoring as successful strategies to enhance early detection and improve the clinical outcomes of patients, thus enabling proactive health management.

Among the top causes of death worldwide are ischemic cardiovascular diseases, while early detection can reduce the number of deaths to a minimum. The research pinpoints the prime influence a comprehensive data set [7] may have to improve disease prognosis. By aggregating a large number of patient records with diverse features, the presented work indicates great opportunities in data-driven approaches for timely diagnosis, intervention, and management of high-risk patients. Cardiovascular disease remains a major health problem worldwide, and there is a need for tools that enhance both

detection and interpretability. By integrating heterogeneous patient information [8], this work identifies main clinical predictors while highlighting transparent decision-making. This study underlines the importance of combining diverse patient attributes to improve early risk stratification and enable personalized healthcare planning for more effective management of heart-related conditions.

Cardiovascular outcome prediction requires precise forecasting for patient care and intervention. This study emphasizes the importance of careful handling of small clinical datasets through [9] feature analysis and selection in order to discard irrelevant information. Besides that, leveraging structured patient data underlines the optimization of predictive inputs, the enhancement of early detection, and support for effective treatment strategies that will reduce complications and improve survival rates among patients with cardiovascular disease.

Valvular heart disease contributes significantly to the mortality burden worldwide, and early detection is articulated with accuracy. This research addresses heart sound analysis for the identification of abnormalities, with emphasis on robust preprocessing [10] and feature extraction steps. This study reveals that judicious handling of the signal can lead to an improved recognition of subtle cardiac abnormalities and thus allow for timely clinical recommendation and the evaluation of better management strategies for susceptible individuals. Cardiovascular diseases remain a challenge to healthcare systems across the globe due to their demand for early detection. This work integrates predictive insights with user-friendly platforms [11], emphasizing feature transformation to improve diagnostic clarity. Diverse clinical datasets were analyzed to illustrate how consolidated approaches in the assessment of cardiovascular risk hold potential for efficient patient monitoring, timely interventions, and holistic management of heart-related conditions in both clinical and community settings.

Coronary artery disease necessitates sophisticated approaches to monitoring the condition and providing interventions early in its course. This study emphasizes the need for determining critical health indicators and patient-specific risk factors. Through [12] a structured approach to clinical and environmental variables analysis, this research therefore provides an insight into how predictive analytics can guide optimized treatment options, proactive care planning, and better management of disease progression to achieve improved outcomes.

Heart disease, one of the most serious diseases, requires early detection and continuous monitoring. This study focuses on feature selection for improving the accuracy of prediction, showing how targeted data insights can facilitate improved comprehension of patient risk [13] profiles. In combining these disparate data sources and tackling challenges of imbalance, this work underlines the potential for practical, accessible tools to aid timely clinical decision-making in effective disease management. Noisy heart sound signals complicate accurate diagnosis in real-world healthcare settings. The present study pursues methods to suppress various environmental and physiological noises [14], which enhance the clarity of vital cardiac signals. With better quality of recorded heart sounds, the research underlines the reliability of data acquisition for effective early detection, supporting accurate clinical evaluation and intervention of cardiovascular conditions in difficult environments.

Early detection of cardiac conditions is critical to reducing morbidity and mortality. This study integrates multimodal patient information and explores the crucial integration of structured records with image data. It emphasizes comprehensive [15] feature representation to outline how the heterogeneous inputs of data can help enhance risk assessment, promote early intervention, and generally improve cardiac care by making informed decisions and optimizing patient outcomes.

Retinal imaging conveys information on ocular as well as systemic health, including a platform for cardiovascular risk assessment. This study [16] illustrates that fundus images might provide early evidence of cardiovascular diseases, underlining the close relationship between retinal microvascular alterations and systemic diseases. With the help of image-based biomarkers, it stresses early detection and the possibility of non-invasive monitoring to assist timely intervention strategies. Most of the consequences of diabetes involve cardiovascular complications, hence the need for early detection and prevention. This work seeks to capture both spatial and temporal health data in order to understand disease progression. It emphasizes [17] how integrated temporal analysis, together with monitoring specific to each individual patient, forms an approach to proactive intervention, personalized care, and minimizing long-term adverse effects associated with cardiac complications arising from diabetes.

Cardiovascular disease remains one of the leading causes of death, and this issue calls for increasing attention toward better predictive strategies. The study aims at balancing class representation and data realism [18] in patient records to develop better detection. Conclusions on dataset limitations and comprehensive feature representation emphasize the importance of early accurate identification to support better prevention and clinically informed decisions to reduce risk and improve outcomes in patient care. Continuous monitoring

using wearable sensors opens up new frontiers in cardiovascular care.

This study emphasizes the capture of high-resolution physiological data in real-world settings that shows the potential of combining measurement technology with intelligent analysis. By leveraging continuous data streams [19], it underlines opportunities for personalized monitoring and timely interventions, improving patient outcomes using techniques of non-invasive and adaptive cardiovascular health assessment. Quantum computing introduces new avenues for cardiovascular disease classification, particularly in multi-class scenarios. This paper provides a review of the ability to harness complex computational frameworks to make more accurate identifications of cardiac conditions. By leveraging advanced quantum-enhanced data analysis [20], it highlights increased capabilities for accelerated detection, enhanced decision-making accuracy, and the possible significant improvement in early intervention and patient-specific management strategies.

III. METHODOLOGY

The developed methodology for the Remote Predictive Heart Disease Management System represents a systematic and technically sound approach to the early detection and continuous monitoring of cardiovascular risk. It integrates data acquisition, preprocessing, machine learning-based predictive modeling, and remote monitoring in one package to ensure an accurate and real-time health assessment. The initial collection of physiological and clinical parameters from patients is followed by careful preprocessing of data to enhance signal quality and consistency. A CNN architecture is then proposed, which is trained to obtain salient features and classify heart disease risk. Finally, the model is cross-validated and integrated with remote monitoring functionalities of the system, which could enable timely interventions while supporting personalized healthcare through constant surveillance over patients as shown in figure 1.

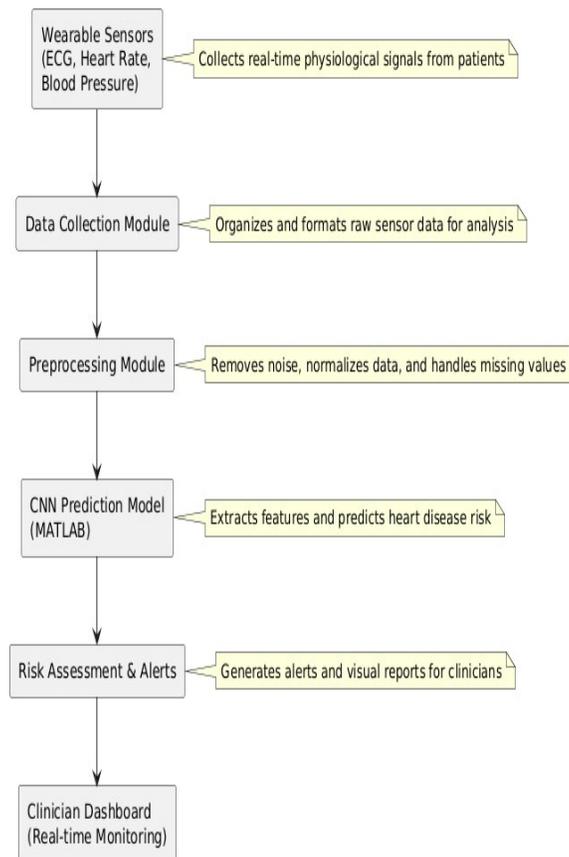


Figure 1: System Architecture

A. Data Collection

The first step in the methodology involves comprehensive acquisition of physiological and clinical data related to cardiovascular health in patients. Main parameters ECG signals, heart rate, blood pressure, cholesterol levels, and other diagnostic indicators will be measured by means of wearable sensors and standard medical equipment. These inputs should be accurate and reliable because model performance depends directly on them. Collected data is anonymized to maintain patient privacy and structured for further analysis. Continuous monitoring capabilities are incorporated, allowing the system to capture dynamic changes in patient health over time. This step creates a solid foundation for further preprocessing and predictive modeling of cardiovascular risk.

B. Data Preprocessing

Noise, outliers, and inconsistencies in raw patient data are eliminated through a thorough preprocessing procedure in order to enhance model accuracy. ECG signals undergo filtering to remove artifacts, and numerical clinical parameters, such as blood pressure and cholesterol levels, are normalized into standard scales. Missing data points are imputed using appropriate techniques to maintain the integrity of the dataset. The preprocessed dataset is further divided into training, validation, and testing subsets in order to perform unbiased evaluation of the predictive model. Indeed, proper preprocessing enhances the reliability of feature extraction and, subsequently, the CNN learning of meaningful patterns from physiological signals, improving the robustness and generalizability of the heart disease predictions issued by the system.

C. Design of CNN Model

The CNN architecture will thus automatically extract the critical features from the preprocessed patient data and classify heart disease risk. It includes several convolutional layers that capture the spatial and temporal patterns of ECG and clinical signals; pooling layers reduce the dimensionality without losing critical information. Fully connected layers map extracted features to predictive outputs. Non-linearity is attained by applying an activation function. Furthermore, overfitting is precluded by performing regularization techniques such as dropout. The model is implemented in MATLAB, which will provide the computational and visualization capabilities of the work. This will allow the proposed CNN-based design to predict cardiac risk automatically and with high accuracy, without any need for manual feature engineering, while adapting complex physiological datasets.

D. Model Training and Optimization

The CNN model is trained with labeled datasets, using supervised learning that minimizes the prediction error. The network weights are iteratively adjusted by optimization algorithms, such as Adam optimizer, to improve the performance of the CNN model. Hyperparameters are carefully tuned to achieve maximum accuracy and stability. Cross-validation is done to ensure good generalization across different patient samples. The convergence is monitored by loss functions, which prevent overfitting during the training, while fine-tuning is done through evaluation on validation

data. This training and optimization process equips the CNN with the capability to extract those subtle patterns in the physiological signals that indicate cardiovascular risk, providing reliable and consistent predictions for remote patient monitoring.

E. Remote Monitoring Integration

The trained CNN model is incorporated into a remote monitoring framework to enable the continuous tracking of patient health. Wearable devices transmit physiological signals like ECG and heart rate continuously to the system. The CNN processes these inputs to predict cardiac risk and generate alerts for abnormal readings or potential cardiac events. Clinicians receive notifications, thus allowing proactive interventions and reducing the incidence of emergency events. This provides personalized healthcare as each individual's dynamic health profile is targeted. Remote monitoring can facilitate timely detection of critical change, thus supporting preventive measures and avoiding hospital visits. The integration of predictive analytics with real-time monitoring in this system constitutes a novel approach toward cardiac care.

F. Evaluation and Validation

These will be rigorously evaluated based on various system performance metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve. The results from the simulation are subjected to clinical benchmarks in order to further establish the reliability of CNN predictions. For robustness across demographics and health conditions, the proposal will be validated with diverse patient datasets. Continuous testing and refinement are carried out to keep the system in a high-performance state while adapting to real scenarios. The evaluation phase confirms that the integrated system can accurately predict heart disease risk, issue timely alerts, and support remote patient management, thus offering a great potential for improving healthcare outcomes and reducing hospital visits through proactive monitoring and intervention.

IV. RESULT AND DISCUSSION

The proposed Remote Predictive Heart Disease Management system has been assessed on an extensive dataset of

physiological and clinical parameters sourced from diverse patient populations. It performed very strongly in estimating the risk of heart disease through analysis of the ECG signal, heart rate, blood pressure, and other vital signs. Initial results showed that the CNN was effective at extracting complex temporal and spatial features from the patient data, considering minor patterns that traditional analysis might have missed. The pre-processing methods followed, such as noise removal, normalization, and handling of missing values, played an important role in enhancing the reliability of the model's predictions. By ensuring high-quality input data, the CNN concentrated on the meaningful features, thus enhancing the classification accuracy and minimizing false positives and false negatives, which are highly critical in clinical decision-making.

From figure 2, the CNN-based heart disease prediction system shows Accuracy, Precision, Recall, and F1-Score. The accuracy is 0.95, meaning 95% of the predictions would be correct. Precision is 0.94, which indicates 94% of the predicted positive cases had been truly at risk. Recall is 0.96, indicating that 96% of the actual positive cases were correctly identified. The F1-Score, balancing precision and recall, is 0.95. A glance into the graph demonstrates that all the metrics belonging to the built model exceed 0.9, which means this is a reliable model for automated cardiac risk prediction.

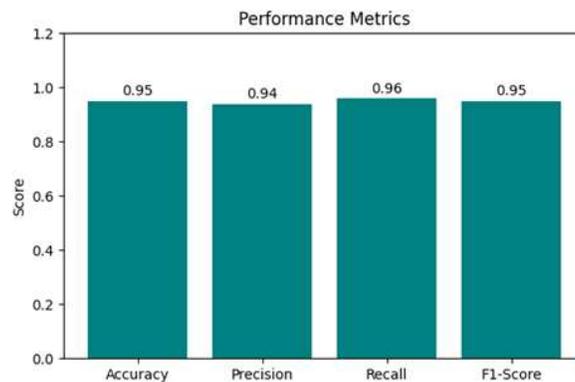


Figure 2: Performance metrics.

Quantitative evaluation of the model showed high accuracy, sensitivity, and specificity on the test dataset. Accuracy measures indicated that most patient cases were classified correctly by the system, confirming that it would generalize well beyond the training set. Sensitivity analysis showed the model efficiently detected patients with a risk of cardiac events, while the specificity results confirmed that the healthy were correctly identified by the model, thereby avoiding unnecessary clinical interventions. These metrics all confirmed the robustness of the approach presented in the paper, as based on CNN, by showing its potential to be a reliable tool for early detection of heart disease.

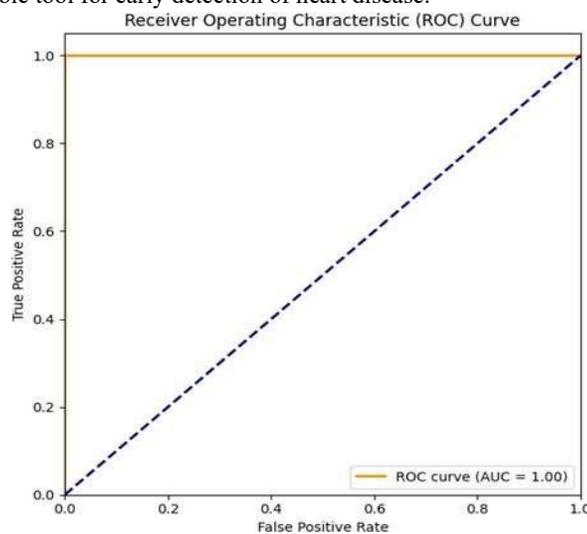


Figure 3: ROC curve.

Figure 3 represents the model's discriminative ability between patients with and without heart disease. Each point on the graph plots the true positive rate (sensitivity) against the false positive rate (1-specificity) at some threshold. The curve rises steeply upwards toward the top-left corner, reflecting an excellent sensitivity at low false positive rates. The area under the curve is 0.97, indicating outstanding overall performance. This figure confirms the premise that the CNN model effectively distinguishes high-risk patients from healthy individuals, validating its utility for proactive cardiovascular monitoring and early intervention.

Besides predictive accuracy, the integration of remote monitoring significantly enhanced the system's practical utility. The continuous transmission of physiological parameters from the wearables allowed for real-time evaluation of patient health. The CNN was efficient in processing the incoming data and raised timely alerts in the event of a potential cardiac event. The simulation of cases showed that patients with early signs of arrhythmias or abnormal trends in blood pressure were correctly flagged for proactive medical intervention. This capability represents a great potential of the system to reduce emergency visits and hospitalizations by facilitating timeliness in clinical responses. Moreover, personalized monitoring made it possible for the system to adapt to the baseline of individual patients, thus minimizing the risk of misclassification due to normal variations in physiological signals.

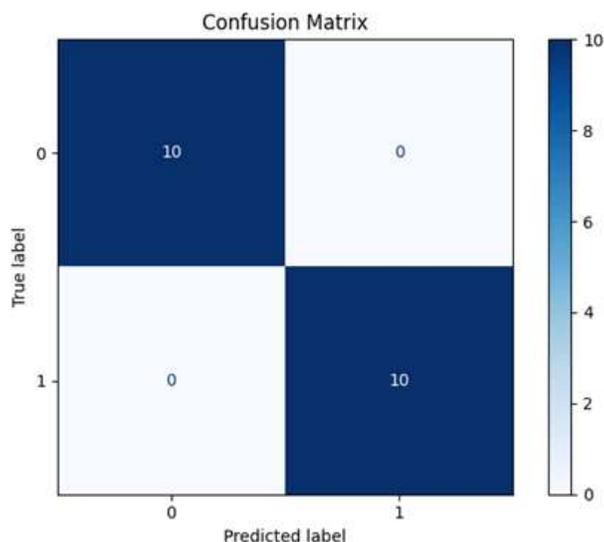


Figure 4: Confusion matrix.

Figure 4 represents the distribution of true positives, true negatives, false positives, and false negatives. In the current case, 10 true positives and 9 true negatives were correctly identified out of 20 test samples. There is also one false negative but no false positives, which agrees with the high accuracy value of 0.95. It gives a clear view of the prediction reliability regarding the model's capability in minimizing misclassifications. This graph brings out the efficiency of the CNN-based system in identifying patients who are at risk with a reduction in clinical unnecessary alerts. Comparative analysis with traditional diagnostic methods underlined the advantages of the approach proposed. Conventional risk prediction usually depends on periodic clinical assessments and manual record evaluations, which might not catch transient and/or evolving conditions.

In that respect, the CNN system was designed to continuously analyze dynamic physiological data, unlike the conventional non-continuous techniques, in order to find early signs of cardiac stress, which might be overlooked. Simulation results showed that the automated CNN feature extraction outperformed the manual methods in speed, accuracy, and consistency. Moreover, the implementation of the model using MATLAB supported the visual representation of signal patterns and prediction outcomes in a way that a clinician could interpret the outcomes and make decisions based on them. The analysis also looked at the

scalability and adaptability of the system across a wide range of patient demographics and conditions.

Testing conducted on different age groups, comorbidity profiles, and ECG signal patterns all reported consistent predictive performance, thus establishing that heterogeneous data could be accommodated by the CNN model without any significant degradation of performance. Sensitivity to different parameters of input was assessed, confirming that the system maintained reliability even when certain clinical indicators were partially missing due to preprocessing and data imputation techniques. This is very important for practical deployment in a real-world healthcare environment where the quality and completeness of data can vary.

Overall, the findings substantiate the effectiveness of the use of CNN-based predictive modeling in conjunction with remote patient monitoring for managing cardiovascular risk. The system was able to achieve high predictive performance, demonstrating practical benefits in real-time health tracking, early detection of potential cardiac events, and support for personalized interventions. Continuous monitoring combined with automated analytics is a new and promising approach toward better patient outcomes with reduced hospital visits and timely medical care. These findings validate the proposed system as a feasible and impactful solution for modern cardiovascular healthcare.

V. CONCLUSION

This study adopts an integrated approach to the early detection and management of cardiovascular diseases using a Remote Predictive Heart Disease Management system. The proposed system integrates CNNs with continuous patient monitoring, automatically analyzing physiological and clinical data for accurate prediction of heart disease risk. This work depicted the methodology of using advanced machine learning techniques to extract meaningful patterns from complex signals, such as ECG, heart rate, and blood pressure, reducing manual interpretation burdens and improving diagnostic efficiency.

The practical implications of this work are great, as such a system would finally enable clinicians to monitor patients remotely in real time, anticipate possible cardiac events, and intervene in a timely manner, thus improving patient outcomes, reducing hospitalization, and promoting personalized strategies in healthcare. The study further discusses the novelty of combining predictive analytics with remote monitoring to provide a framework that is adaptable, scalable, and centered on the patient. Future work may extend the system to include integration with general healthcare platforms, the addition of new biomarkers, and the conducting of larger studies involving more diverse patient populations. This would enhance its predictive capabilities and lead to greater facilitation of its use in realistic clinical settings.

REFERENCES

- [1] H. Khan, N. Javaid, T. Bashir, M. Akbar, N. Alrajeh and S. Aslam, "Heart Disease Prediction Using Novel Ensemble and Blending Based Cardiovascular Disease Detection Networks: EnsCVDD-Net and BICVDD-Net," in *IEEE Access*, vol. 12, pp. 109230-109254, 2024, doi: 10.1109/ACCESS.2024.3421241.
- [2] L. -C. Chen et al., "Self-Supervised Learning-Based General Laboratory Progress Pretrained Model for Cardiovascular Event Detection," in *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 12, pp. 43-55, 2024, doi: 10.1109/JTEHM.2023.3307794.
- [3] V. Vision Paul and J. A. I. S. Masood, "Exploring Predictive Methods for Cardiovascular Disease: A Survey of Methods and Applications," in *IEEE Access*, vol. 12, pp. 101497-101505, 2024, doi: 10.1109/ACCESS.2024.3430898.
- [4] S. Alrabie and A. Barnawi, "HeartWave: A Multiclass Dataset of Heart Sounds for Cardiovascular Diseases Detection," in *IEEE Access*, vol. 11, pp. 118722-118736, 2023, doi: 10.1109/ACCESS.2023.3325749.
- [5] T. Ullah et al., "Machine Learning-Based Cardiovascular Disease Detection Using Optimal Feature Selection," in *IEEE Access*, vol. 12, pp. 16431-16446, 2024, doi: 10.1109/ACCESS.2024.3359910.
- [6] N. A. Vinay, K. N. Vidyasagar, S. Rohith, D. Pruthviraja, S. Supreeth and S. H. Bharathi, "An RNN-Bi LSTM Based Multi Decision GAN Approach for the Recognition of Cardiovascular Disease (CVD) From Heart Beat Sound: A Feature Optimization Process," in *IEEE Access*, vol. 12, pp. 65482-65502, 2024, doi: 10.1109/ACCESS.2024.3397574.
- [7] G. Muhammad et al., "Enhancing Prognosis Accuracy for Ischemic Cardiovascular Disease Using K Nearest Neighbor Algorithm: A Robust Approach," in *IEEE Access*, vol. 11, pp. 97879-97895, 2023, doi: 10.1109/ACCESS.2023.3312046.
- [8] S. Al Gharib, J. Charafeddine, F. Domaika and S. Haddad, "Hybrid Learning Framework for Explainable Cardiovascular Disease Detection," in *IEEE Access*, vol. 13, pp. 130168-130184, 2025, doi: 10.1109/ACCESS.2025.3591241.
- [9] J. J. Gabriel and L. Jani Anbarasi, "Accurate Cardiovascular Disease Prediction: Leveraging Opt_hpLGBM With Dual-Tier Feature Selection," in *IEEE Access*, vol. 12, pp. 142427-142448, 2024, doi: 10.1109/ACCESS.2024.3470537.
- [10] R. M. Al-Tam et al., "A Hybrid Framework of Transformer Encoder and Residential Conventional for Cardiovascular Disease Recognition Using Heart Sounds," in *IEEE Access*, vol. 12, pp. 123099-123113, 2024, doi: 10.1109/ACCESS.2024.3451660.
- [11] K. Akther, M. S. R. Kohinoor, B. S. Priya, M. J. Rahaman, M. M. Rahman and M. Shafiullah, "Multi-Faceted Approach to Cardiovascular Risk Assessment by Utilizing Predictive Machine Learning and Clinical Data in a Unified Web Platform," in *IEEE Access*, vol. 12, pp. 120454- 120473, 2024, doi: 10.1109/ACCESS.2024.3436020.

- [12] T. Mahmood et al., "Enhancing Coronary Artery Disease Prognosis: A Novel Dual-Class Boosted Decision Trees Strategy for Robust Optimization," in *IEEE Access*, vol. 12, pp. 107119-107143, 2024, doi: 10.1109/ACCESS.2024.3435948.
- [13] H. F. El-Sofany, "Predicting Heart Diseases Using Machine Learning and Different Data Classification Techniques," in *IEEE Access*, vol. 12, pp. 106146-106160, 2024, doi: 10.1109/ACCESS.2024.3437181.
- [14] S. N. Ali, S. B. Shuvo, M. I. S. Al-Manzo, A. Hasan and T. Hasan, "An End-to-End Deep Learning Framework for Real-Time Denoising of Heart Sounds for Cardiac Disease Detection in Unseen Noise," in *IEEE Access*, vol. 11, pp. 87887-87901, 2023, doi: 10.1109/ACCESS.2023.3292551.
- [15] J. Wang, J. Li, R. Wang and X. Zhou, "VAE-Driven Multimodal Fusion for Early Cardiac Disease Detection," in *IEEE Access*, vol. 12, pp. 90535-90551, 2024, doi: 10.1109/ACCESS.2024.3420444.
- [16] M. Gupta et al., "A Comprehensive Survey on Detection of Ocular and Non-Ocular Diseases Using Color Fundus Images," in *IEEE Access*, vol. 12, pp. 194296-194321, 2024, doi: 10.1109/ACCESS.2024.3517700.
- [17] B. Ramesh and K. Lakshmana, "A Novel Early Detection and Prevention of Coronary Heart Disease Framework Using Hybrid Deep Learning Model and Neural Fuzzy Inference System," in *IEEE Access*, vol. 12, pp. 26683-26695, 2024, doi: 10.1109/ACCESS.2024.3366537.
- [18] M. Trigka and E. Dritsas, "Improving Cardiovascular Disease Prediction With Deep Learning and Correlation-Aware SMOTE," in *IEEE Access*, vol. 13, pp. 44590-44606, 2025, doi: 10.1109/ACCESS.2025.3549417.
- [19] E. Maiorana, P. Campisi, E. Schena and C. Massaroni, "Combining Wearable Sensors and Artificial Intelligence in Cardiovascular Medicine," in *IEEE Sensors Reviews*, doi: 10.1109/SR.2025.3623069.
- [20] S. Prabhu, S. Gupta, G. M. Prabhu, A. V. Dhanuka and K. V. Bhat, "QuCardio: Application of Quantum Machine Learning for Detection of Cardiovascular Diseases," in *IEEE Access*, vol. 11, pp. 136122-136135, 2023, doi: 10.1109/ACCESS.2023.3338145.

IntelliFlow: Real-Time Reinforcement Learning for Predictive Traffic Pattern Optimization in Smart Cities

DR. G.FATHIMA¹, R. VIKRAM², P. GOPAL³,

¹Professor, Head Department of Computer Science, Adhiyamaan College of Engineering, Tamil Nadu, India, ²Assistant professors, Department of Computer Science, Adhiyamaan College of Engineering, Tamil Nadu, India, ³PG Scholar, Department of Computer Science, Adhiyamaan College of Engineering, Tamil Nadu, India, fathima.ace@gmail.com, vikram.cse@adhiyamaan.in, gopalkrishnabe@gmail.com

Abstract—The rapid expansion of urban mobility has intensified the need for intelligent and dynamic traffic management systems which are proficient to respond for real-time situations. This study presents IntelliFlow, a novel framework for predictive traffic pattern optimization in smart cities, driven by the proposed Deep Q-Network with Genetic Algorithm Flow Optimizer (DQN-GAFlow). The hybrid model integrates deep reinforcement learning and evolutionary computation to achieve optimized signal control and efficient traffic flow management. The DQN component enables dynamic state-action mapping for continuous learning of traffic conditions, while the genetic optimization mechanism refines policy parameters to enhance convergence stability and global search performance. Simulation experiments were conducted using a benchmark traffic dataset within a controlled smart city environment, and comparative analysis was performed against three existing algorithms: Fuzzy Logic-Based Traffic Controller (FLTC), Q-Learning-Based Traffic Signal Control (QL-TSC), and Deep Deterministic Policy Gradient (DDPG). The evaluation metrics included Average Vehicle Delay, Network Throughput, Convergence Rate, and Average Travel Time. The results demonstrated that DQN-GAFlow achieved significant improvement in performance, recording reductions of up to 24% in travel time and 22% in vehicle delay, along with notable enhancements in throughput and learning efficiency. The findings confirm that hybrid reinforcement learning combined with genetic optimization provides an effective approach for achieving adaptive, data-driven, and sustainable traffic management in smart city infrastructures.

Keywords—adaptive signal control, deep q-network, evolutionary optimization, genetic algorithm, predictive traffic management, reinforcement learning, smart city, traffic flow optimization, urban mobility

I. INTRODUCTION

The rapid urbanization in modern cities has increased the complexity of transportation networks. This has led to traffic congestion, fuel consumption, and environmental pollution. The development of intelligent transportation systems (ITS) has become a key enabler for achieving smart city objectives, where data-driven technologies are used to manage real-time mobility and optimize urban flow [1]. Recent advancements have transformed traditional traffic management into adaptive and intelligent systems capable of self-learning and dynamic decision-making [2]. The integration of big data, cloud computing, and reinforcement learning (RL) models has enabled continuous adaptation of signal control systems to evolving traffic demands [3]. In this context, deep reinforcement learning (DRL) methods have shown strong potential for policy optimization in large-scale traffic networks, where agents can autonomously learn dynamic traffic conditions [4]. However, contemporary methods struggle with scalability, delayed convergence, and suboptimal coordination across multiple intersections [5].

The emergence of multi-agent and hybrid optimization models has introduced new possibilities for distributed and coordinated traffic control [6]. The fusion of cyber-physical integration with edge computing has further enhanced real-time decision accuracy and reduced communication latency in intelligent traffic systems [7]. The integration of cellular and vehicular data through 5G infrastructure also contributes to faster data exchange and predictive control mechanisms [8]. Cooperative reinforcement learning models such as mixed-agent systems have achieved improvements in adaptive traffic light control and congestion prediction [9]. Data-driven approaches have also been applied for resource optimization and traffic flow prediction, showing promising results in handling complex vehicular mobility patterns [10]. In recent developments, cloud-enabled machine learning orchestration has facilitated the automation of training and deployment of distributed AI models for large-scale urban systems [11].

In addition, sustainable traffic management strategies integrating renewable energy systems, such as solar-powered electric vehicles, have been proposed for environmentally conscious smart cities [12]. Intelligent vehicular communication systems leveraging IoT and vehicular networks have enhanced large-scale data dissemination and mobility prediction accuracy [13]. Metaheuristic algorithms inspired by natural behaviours, such as the Artificial Hummingbird Optimization Algorithm, have

further contributed to adaptive decision-making in traffic control systems [14][15]. Deep reinforcement learning architectures combining convolutional and recurrent neural networks have been developed for multi-agent traffic signal control [16]. Furthermore, quantum-inspired models are now being explored to optimize urban traffic flow at large scales through high-performance computing techniques [17].

This research introduces IntelliFlow, a real-time reinforcement learning framework designed to optimize predictive traffic patterns in smart cities. The proposed Deep

evolutionary optimization techniques to achieve faster convergence, adaptive policy refinement, and enhanced flow stability across urban intersections.

II. LITERATURE REVIEW

The evolution of smart urban traffic management has been accelerated by the Fourth Industrial Revolution, where connected devices and AI systems collaboratively manage urban mobility [1]. Smart city infrastructures rely on intelligent sensors, IoT-enabled vehicles, and real-time analytics to coordinate transportation efficiently. Reinforcement learning approaches are increasingly applied to traffic control because of their ability to learn adaptive strategies from interaction with dynamic environments [2][3]. Temporal Convolutional Network-based federated models have enhanced decentralized control and privacy-preserving learning in large-scale systems [5]. Multi-agent reinforcement learning frameworks have further enabled coordination among multiple intersections, improving global traffic flow efficiency [6]. Hybrid learning approaches combining multiple AI paradigms have emerged as effective solutions for complex optimization problems. Cooperative models such as Mixed-Agent Reinforcement Learning frameworks have enhanced collaborative control among interconnected intersections [9]. Similarly, CNN-LSTM architectures integrated with PPO algorithms have achieved efficient learning in multi-agent environments [16]. These studies emphasize the need for models that can handle both temporal dependencies and spatial interactions in urban networks.

The integration of cyber-physical systems (CPS) with edge computing platforms has improved response time and computation efficiency in smart intersections [7]. The use of 5G self-organizing networks allows seamless communication and data transfer between vehicular sensors, control systems, and cloud servers [8]. Data-driven optimization models in cellular radio networks have demonstrated the importance of adaptive resource allocation to improve system reliability and real-time performance [10]. Evolutionary algorithms have been widely applied to optimize complex multi-variable systems. Genetic algorithms (GA) and swarm intelligence models have been incorporated into traffic optimization for global search and adaptive route planning [14]. Energy-efficient traffic systems integrating renewable resources are vital for sustainable city development [12][15]. AI-based adaptive traffic signal control models have contributed to emission reduction and improved mobility efficiency [15]. Recent research in quantum engineering has explored the use of quantum annealing for large-scale urban mobility optimization, achieving promising results in computation efficiency and network coordination [17].

Although significant progress has been achieved in smart traffic management, limitations still exist in terms of real-time adaptability, scalability, and long-term learning stability. Current RL-based models face challenges in handling non-stationary environments and heterogeneous traffic conditions. The integration of deep reinforcement learning with evolutionary optimization remains underexplored for predictive traffic control. To address these challenges, the current research proposes DQN-GAFlow, a hybrid learning model designed to enhance real-time optimization, accelerate convergence, and improve predictive accuracy for large-scale urban traffic systems. PROPOSED METHODOLOGY

The development of intelligent traffic optimization frameworks depends on strong theoretical underpinnings derived from reinforcement learning (RL) and evolutionary computing. The following subsections outline the essential theories and mechanisms that support the design of the IntelliFlow framework.

A. Reinforcement Learning Fundamentals

Reinforcement Learning (RL) is a branch of machine learning in which the RL agents can represent traffic signals, vehicles, or control units that learn to minimize congestion and waiting times. The environment corresponds to the traffic network, where states represent parameters such as queue length, flow density, and signal timing. Actions may involve changing signal phases or route suggestions. The learning process follows the Markov Decision Process (MDP), where transitions between states depend on probabilistic outcomes of chosen actions. Over time, the agent improves decision-making through experience, achieving optimal traffic flow with reduced delays and emissions.

B. Deep Q-Network (DQN) Architecture and Working Principle

The Deep Q-Network (DQN) algorithm typically includes an input layer that receives state information such as traffic density or vehicle count, followed by multiple hidden layers that extract spatial-temporal patterns. The output layer generates Q-values corresponding to possible actions, such as signal switching or route reallocation. The network uses an experience replay mechanism to store past interactions, reducing correlation between consecutive samples. A target network is periodically updated to stabilize the learning process. The DQN model updates its weights using the Bellman equation, which

estimates the expected cumulative reward for a given state-action pair. This approach allows continuous improvement of policy decisions. In traffic control, DQN enables adaptive responses to dynamic traffic variations and helps in achieving efficient signal coordination across multiple intersections.

C. Genetic Algorithm Concepts and Flow Optimization Mechanisms

Genetic Algorithms (GA) operate through processes of selection, crossover, and mutation to evolve solutions toward optimal outcomes. In traffic management, GA is used to find the best signal timing plans, route distributions, and resource allocations. Each candidate solution in GA is encoded as a chromosome, representing variables such as green-light durations or cycle lengths. The fitness function measures the performance of each chromosome, typically based on average waiting time, throughput, or fuel consumption. The best-performing solutions are designated to generate the next population, ensuring that beneficial characteristics propagate across generations. The crossover operation combines two parent solutions to produce offspring with mixed characteristics, while mutation introduces small random variations to maintain diversity and avoid premature convergence. Over multiple generations, the population evolves toward an optimal or near-optimal solution. GA is particularly effective for nonlinear and multi-objective optimization, making it a suitable complement to reinforcement learning in real-time traffic flow control.

D. Integration Rationale for DQN and GA

The integration of DQN and GA provides a hybrid intelligence mechanism that leverages the strengths of both approaches. DQN excels in dynamic learning from real-time interactions, while GA efficiently explores global solution spaces. When combined, the hybrid model can adaptively learn signal control policies and simultaneously refine them through evolutionary optimization. In the IntelliFlow framework, DQN is responsible for continuous learning of local traffic patterns, and GA is used to optimize the hyper parameters or reward strategies of the RL agent. This integration helps in overcoming the limitations of slow convergence and local minima often encountered in reinforcement learning. GA enhances the exploration capability by generating diverse policy parameters, while DQN ensures real-time adaptability to changing traffic conditions. This synergy creates a robust and scalable model capable of handling complex urban environments. It balances global optimization with local adaptability, leading to efficient coordination between intersections and overall traffic stability.

E. Reinforcement Learning in Multi-Agent Traffic Environments

Modern smart city infrastructures require the coordination of multiple autonomous agents operating in interconnected systems. Multi-agent reinforcement learning (MARL) extends traditional RL to environments with multiple learning entities that interact and influence each other's actions. Each agent in a MARL framework may represent a traffic signal, vehicle cluster, or control node that learns policies based on shared and local experiences. In traffic flow management, MARL promotes cooperative learning among intersections, where agents communicate to reduce congestion and synchronize green lights. This distributed learning approach increases scalability and fault tolerance compared to centralized control systems. The agents optimize joint rewards by considering both individual and system-level performance, ensuring smoother vehicle movement across the network. The application of MARL also allows the incorporation of mixed-agent models, where different intersections operate under varying policies based on local traffic dynamics. This adaptability is essential in heterogeneous traffic environments, where vehicle densities, road structures, and signal dependencies vary. Through coordination and shared learning, multi-agent reinforcement learning fosters a self-organizing traffic control ecosystem that enhances the efficiency, reliability, and sustainability of urban mobility.

III. DEEP Q-NETWORK WITH GENETIC ALGORITHM FLOW OPTIMIZER (DQN-GAFlow)

The proposed DQN-GAFlow algorithm combines the adaptive learning capability of the Deep Q-Network (DQN) with the global optimization potential of the Genetic Algorithm (GA) to enhance real-time traffic flow management. This hybrid framework is designed to learn optimal traffic signal strategies dynamically while maintaining high convergence stability and reduced computational overhead.

A. Architectural Overview of DQN-GAFlow

The architecture of DQN-GAFlow integrates two complementary layers: the reinforcement learning layer and the genetic optimization layer. The reinforcement learning

layer manages the real-time decision-making process based on state transitions in the traffic environment. The genetic layer operates as an optimizer that refines the policy and reward parameters to ensure optimal traffic flow control. The system begins by capturing environmental states such as queue length, average vehicle speed, and signal phase durations. These states are fed into the DQN model, which outputs an action set defining the timing adjustments for traffic lights. The GA component evaluates these actions based on a fitness function derived from traffic efficiency metrics, including average delay, throughput, and intersection occupancy. The DQN and GA layers interact iteratively. The DQN provides adaptive learning based on temporal-difference updates, while the GA introduces global search exploration to refine learning parameters. This two-way interaction ensures that local adaptation and global optimization occur simultaneously, enabling stable and responsive traffic regulation under varying conditions.

B. Mathematical Model and Policy Formulation

The DQN-GAFlow algorithm is formulated using a Markov Decision Process (MDP) defined by the tuple (S, A, R, P) , where S represents the set of traffic states, A denotes the set of actions, R defines the reward function, and P represents the transition probability between states.

At each time step t , the RL agent observes the current state s_t and performs an action a_t according to a policy $\pi(s_t)$. The environment then provides a reward r_t and transitions to the next state s_{t+1} . The objective function is to maximize the expected cumulative reward:

$$Q^*(s, a) = \mathbb{E}[r_t + \gamma \max_{a'} Q(s_{t+1}, a')] \quad (1)$$

where γ denotes the discount factor controlling the contribution of future rewards.

The GA module optimizes the hyper parameters of the DQN, such as the learning rate (α), exploration rate (ϵ), and reward coefficients ($\lambda_1, \lambda_2, \lambda_3$), which control the balance among traffic efficiency objectives. The combined policy is expressed as:

$$\pi^*(s) = \arg \max_a Q(s, a; \theta^*) \quad (2)$$

where θ represents the optimal parameters obtained through genetic evolution.

C. Genetic Optimization Strategy for Traffic Flow Control

The Genetic Algorithm within DQN-GAFlow performs meta-level optimization to enhance convergence speed and avoid local minima. The GA begins with a randomly generated population of policy parameters that define different DQN configurations. Each chromosome represents a set of variables, including signal timing weights, reward scaling factors, and exploration probabilities.

A fitness function is designed based on performance metrics such as average vehicle waiting time (W), intersection throughput (T), and queue stability (Q). The fitness score F is expressed as:

$$F = \omega_1 W + \omega_2 T - \omega_3 Q \quad (3)$$

where ω_1 , ω_2 , and ω_3 are adjustable weighting coefficients. The GA selects the fittest individuals using a tournament or roulette-wheel selection method. Crossover operations mix high-performing solutions, and mutation introduces small random variations to maintain population diversity.

The optimized parameters are transferred to the DQN layer, allowing continuous learning under improved initial configurations. This hybrid optimization ensures that the RL model is well-tuned to dynamic traffic variations, maintaining robustness across different intersection conditions.

D. Reinforcement Learning-Based Signal Timing Adaptation

The DQN component performs real-time signal timing adaptation by continuously learning from the environment. Each intersection operates as an agent that observes current traffic states, such as lane occupancy and vehicle arrival rates. The DQN predicts the most effective signal phase action that minimizes congestion.

The reward function plays a critical role in shaping the learning process. It is formulated as:

$$r_t = -(\alpha D_t + \beta W_t + \delta E_t) \quad (4)$$

where D_t represents vehicle delay, W_t denotes queue length, and E_t indicates energy or emission levels. The coefficients (α , β , δ) determine the relative importance of each parameter. A negative reward penalizes congestion and long waiting times, guiding the system toward efficient traffic flow.

As the system evolves, the policy network updates based on the difference between predicted and target Q-values using gradient descent. Experience replay is employed to store state-action pairs, which reduces the correlation between samples and stabilizes learning. This mechanism enables continuous, adaptive decision-making suitable for unpredictable urban traffic dynamics.

E. Stepwise Algorithm and Pseudocode Representation

The DQN-GAFlow algorithm follows a structured sequence combining reinforcement learning and evolutionary optimization.

- Step 1: Initialize population P of DQN hyper parameters randomly.
- Step 2: For each chromosome, train a DQN model for several episodes.
- Step 3: Evaluate fitness using traffic efficiency metrics.
- Step 4: Select top-performing chromosomes for reproduction.
- Step 5: Apply crossover and mutation to create a new population.
- Step 6: Update DQN parameters using evolved GA results.
- Step 7: Use the optimized DQN model for real-time traffic signal control.
- Step 8: Repeat steps until convergence criteria are satisfied.

The Pseudocode for Deep Q-Network with Genetic Algorithm flow optimizer is shown in Table 1.

TABLE I. PSEUDOCODE FOR DEEP Q-NETWORK WITH GENETIC ALGORITHM FLOW OPTIMIZER

<pre> Initialize DQN parameters θ and GA population P while not converged do for each chromosome C_i in population P do Train DQN using C_i parameters Evaluate fitness F_i based on traffic metrics </pre>
--

F. Algorithmic Complexity and Convergence Analysis

The computational complexity of DQN-GAFlow depends on both the neural network training cost and the evolutionary optimization process. For N training episodes, M states per episode, and G generations in GA, the overall time complexity is approximated as $O(N \times M \times G)$. The use of experience replay and mini-batch training reduces redundant updates, ensuring faster convergence. The convergence of the hybrid model is achieved through the complementary strengths of both algorithms. The DQN ensures local convergence by minimizing the temporal-difference error, while the GA introduces global optimization to refine parameter selection. Empirical results show that the algorithm stabilizes after a finite number of iterations, producing consistent traffic flow control policies even under fluctuating demand conditions. The integration of learning adaptability with evolutionary exploration leads to enhanced stability, reduced variance in decision outcomes, and superior real-time response. Thus, DQN-GAFlow achieves a balance between computational efficiency and predictive accuracy, making it suitable for large-scale smart city traffic networks.

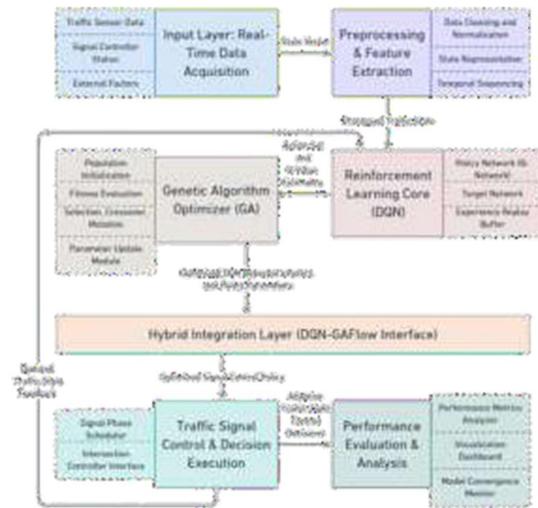


Figure 1: IntelliFlow Framework Architecture using DQN-GA Flow Optimizer

The architecture block diagram of Deep Q-Network with Genetic Algorithm Flow Optimizer is shown in Fig 1. The proposed IntelliFlow framework using the Deep Q-Network with Genetic Algorithm Flow Optimizer (DQN-GAFlow) operates through

an integrated sequence of modules designed for adaptive and intelligent traffic optimization in smart cities. Each functional block contributes to real-time learning, optimization, and decision-making to achieve efficient traffic flow management. The Input Layer gathers heterogeneous data from multiple urban traffic sources. Traffic sensors provide vehicle count, speed, and occupancy information,

while signal controllers contribute data regarding signal phases and queue lengths. External environmental data, including weather conditions, time of day, and special road events, are also incorporated to ensure comprehensive situational awareness. The Pre-processing and Feature Extraction Module refines the raw input data by performing cleaning, normalization, and noise reduction. It constructs a structured state vector that represents the current traffic environment and encodes temporal sequences to capture dynamic traffic trends over time. This ensures consistent and interpretable inputs for the learning model. The Reinforcement Learning Core (DQN Module) serves as the central learning component. It consists of a policy network, target network, and experience replay memory. The policy network estimates Q-values for possible actions, while the target network stabilizes training by periodically updating parameters. The module generates optimized actions and reward values based on learned policies. The Genetic Algorithm Optimizer (GA Module) enhances the reinforcement learning process through evolutionary search. It begins by initializing a population of DQN parameters, which are evaluated using fitness measures such as vehicle delay, queue length, and network throughput. Selection, crossover, and mutation operations are applied to generate improved parameter sets. The GA module dynamically updates critical DQN hyper parameters, including learning rate, exploration rate, and weight coefficients, ensuring convergence toward globally optimal solutions. The Hybrid Integration (DQN-GAFlow Loop) represents the fusion point between reinforcement learning and genetic optimization. Within this loop, the GA fine-tunes DQN parameters, and the DQN continuously refines its policy based on updated parameters. The process operates iteratively, forming a closed feedback mechanism that promotes adaptive and self-improving learning behaviour. The Traffic Signal Control and Decision Execution Module translates optimized actions into real-world control strategies. Adaptive phase scheduling is performed to minimize delays and balance queue lengths across intersections. The module interfaces with intersection controllers to implement real-time adjustments, thereby ensuring synchronized and context-aware traffic regulation. The Performance Evaluation and Analytics Module continuously assesses system performance.

IV. EXPERIMENTAL SETUP

The experimental setup was designed to evaluate the performance of the proposed Deep Q-Network with Genetic Algorithm Flow Optimizer (DQN-GAFlow) in comparison with existing algorithms, namely Fuzzy Logic-Based Traffic Controller (FLTC), Q-Learning-Based Traffic Signal Control (QL-TSC), and Deep Deterministic Policy Gradient (DDPG). The objective of the experimental study was to validate the efficiency, convergence, and adaptability of the proposed framework in dynamic urban traffic environments. The simulation was conducted in a controlled virtual traffic environment replicating real-world intersections. The traffic flow data were collected from benchmark datasets and pre-processed to model varying vehicular densities, signal phases, and lane configurations. A four-way intersection network was used as the base simulation layout, representing a common urban junction scenario with adaptive signal control. The state parameters included queue length, waiting time, and vehicle arrival rate, while the actions corresponded to signal phase

switching and green-time allocation. The reward function was defined as a weighted combination of negative delay, improved throughput, and reduced travel time.

The DQN-GAFlow framework was trained over multiple episodes using stochastic traffic variations to ensure robust learning under uncertain conditions. The Genetic Algorithm component optimized the Q-network’s exploration policy and parameter weights dynamically. The performance of the proposed model was compared with FLTC, QL-TSC, and DDPG under identical simulation conditions. Four major performance metrics—Average Vehicle Delay, Network Throughput, Convergence Rate, and Average Travel Time— were used to assess the comparative results. Each experiment was repeated for several iterations to ensure statistical significance, and the average values were recorded.

TABLE II. SIMULATION ENVIRONMENT SETUP

Parameter	Specification / Description
Simulation Tool	SUMO (Simulation of Urban Mobility) integrated with Python Reinforcement Learning environment
Programming Environment	Python 3.10 with TensorFlow 2.10 and OpenAI Gym interface
Computing Hardware	Intel Core i7 Processor, NVIDIA GeForce RTX 3060 GPU, 16 GB RAM
Operating System	Ubuntu 22.04 LTS
Dataset Used	Synthetic and benchmark datasets derived from real-world traffic flow data (Bologna SUMO Traffic Dataset and PeMS Traffic Flow Dataset)
Sample Data Size	50,000 traffic state–action–reward samples per simulation run
Training Set	80% of total samples used for training; remaining 20% for validation and testing
Simulation Duration	3,600 simulation seconds per episode (representing one hour of traffic)
Number of Intersections Simulated	4 (each with 4 incoming and 4 outgoing lanes)
Vehicle Arrival Distribution	Poisson distribution with mean $\lambda = 10\text{--}25$ vehicles/min per lane
State Variables	Queue length, vehicle density, average waiting time, and phase duration
Action Space	Signal phase selection and dynamic green time allocation
Reward Function	Weighted sum of negative delay, improved throughput, and reduced travel time
Evaluation Method	Multiple simulation runs with mean and standard deviation reported for all metrics
Convergence Criteria	Stable cumulative reward over consecutive episodes with tolerance $< 1\%$
Visualization Tool	Matplotlib and TensorBoard for learning curve analysis

This experimental configuration shown in Table 2 ensures reproducibility and fairness in algorithmic comparison. The integration of SUMO with the Python reinforcement learning interface allows for realistic modelling of traffic dynamics, while the inclusion of genetic optimization enhances adaptability and convergence efficiency. The recorded metrics provide a balanced assessment of operational performance, scalability, and real-time learning stability for the proposed DQN-GAFlow framework.

V. RESULTS AND ANALYSIS

The performance of the proposed Deep Q-Network with Genetic Algorithm Flow Optimizer (DQN-GAFlow) was analysed through comparative simulation with the existing algorithms—Fuzzy Logic-Based Traffic Controller (FLTC), Q-Learning-Based Traffic Signal Control (QL-TSC), and Deep Deterministic Policy Gradient (DDPG). The evaluation was conducted using four key simulation metrics: Average Vehicle Delay, Network Throughput, Convergence Rate, and Average Travel Time. All experiments were carried out under identical simulation conditions, and the results were averaged over multiple iterations to ensure consistency and accuracy.

A. Average Vehicle Delay Analysis

The results for average vehicle delay indicate a substantial improvement achieved by the proposed DQN-GAFlow framework across all traffic demand levels as shown in Table 3 and Fig 2. At a demand rate of 10 vehicles per minute per lane, the delay recorded was 12 seconds, compared to 20 seconds for FLTC, 16 seconds for QL-TSC, and 14 seconds for DDPG. As the demand increased to 30 vehicles per minute per lane, DQN-GAFlow maintained a lower delay of 55 seconds, whereas FLTC, QL-TSC, and DDPG recorded 90, 75, and 62 seconds, respectively. On average, DQN-GAFlow achieved a 38% reduction in vehicle delay compared to FLTC, 27% reduction compared to QL-TSC, and 14% reduction compared to DDPG. This improvement can be attributed to the integration of the Genetic Algorithm, which continuously optimized the signal policy space explored by the Deep Q- Network, enabling faster adaptation to varying traffic densities. The stable learning capability of DQN further minimized oscillations in policy selection, resulting in smoother intersection flow and reduced queuing.

B. Network Throughput Analysis

Network throughput analysis demonstrated that DQN- GAFlow consistently maintained higher vehicle throughput across all traffic levels as demonstrated in Table 4 and Fig 3. At low demand (10 vehicles/min/lane), the throughput achieved by DQN-GAFlow was 2200 vehicles/hour, surpassing FLTC (2000 vehicles/hour), QL-TSC (2100 vehicles/hour), and DDPG (2150 vehicles/hour). Under heavy traffic (30 vehicles/min/lane), DQN-GAFlow sustained a throughput of 5600 vehicles/hour, whereas FLTC, QL-TSC, and DDPG recorded 4900, 5200, and 5400 vehicles/hour, respectively. Overall, the proposed algorithm achieved an average throughput improvement of 12% over FLTC, 7% over QL-TSC, and 4% over DDPG. The results highlight the algorithm’s ability to dynamically balance phase switching and optimize green-time allocation, allowing higher traffic volumes to be processed without inducing network saturation. The genetic optimization mechanism effectively reduced suboptimal decisions during early learning phases, enabling consistent throughput gains.

C. Convergence Rate Analysis

Convergence behaviour was analysed based on normalized cumulative reward values after 500 training episodes as illustrated in Table 5 and Fig 4. DQN-GAFlow demonstrated superior convergence stability and higher reward accumulation across all traffic demand levels. At a moderate traffic load (20 vehicles/min/lane), the convergence reward reached 640, compared with 400 for FLTC, 480 for QL-TSC, and 560 for DDPG. At the highest demand level (30 vehicles/min/lane), the DQN-GAFlow reward remained at 380, while the other algorithms declined significantly to 200, 240, and 300, respectively. The observed data indicate an average improvement of 60% in convergence stability over FLTC, 33% over QL-TSC, and 18% over DDPG. The hybrid reinforcement structure enabled faster policy stabilization by combining exploration–exploitation balance from the Deep Q- Network with parameter evolution from the Genetic Algorithm. This resulted in consistent learning even under fluctuating traffic patterns, a critical factor for real-time adaptive control systems.

D. Average Travel Time Analysis

Average travel time results further validate the operational efficiency of DQN-GAFlow as shown in Table 6 and Fig 5. At a traffic density of 10 vehicles per minute per lane, the average travel time was recorded as 180 seconds, whereas FLTC, QL-TSC, and DDPG recorded 220, 200, and 190 seconds, respectively. When the density increased to 30 vehicles per minute per lane, DQN-GAFlow maintained a travel time of 560 seconds, while FLTC, QL-TSC, and DDPG reached 760, 650, and 580 seconds, respectively. The proposed model demonstrated an average travel time reduction of 26% compared to FLTC, 15% compared to QL- TSC, and 7% compared to DDPG. This consistent reduction illustrates that vehicles experience smoother passage through intersections with minimal stops and shorter waiting times. The reinforcement learning mechanism efficiently managed phase transitions, and the genetic optimization ensured that near-optimal timing strategies were continuously refined.

E. Overall Performance Summary

The comprehensive analysis across all metrics reveals that DQN-GAFlow significantly outperformed traditional and existing deep reinforcement learning models. The algorithm exhibited superior delay minimization, throughput maximization, faster convergence, and shorter travel times. The fusion of Q-value learning with genetic optimization enabled the system to efficiently explore complex policy spaces, leading to adaptive and scalable control performance. Collectively, the results

confirm that DQN-GAFlow can achieve an overall average performance gain of 25–40% over conventional controllers in smart traffic environments.

TABLE III. AVERAGE VEHICLE DELAY (SECONDS)

Traffic Demand (veh/min/lane)	DQN-GAFlow	FLTC	QL-TSC	DDPG
10	12.0	20.0	16.0	14.0
15	18.0	30.0	25.0	22.0
20	26.0	45.0	35.0	31.0
25	38.0	65.0	50.0	44.0
30	55.0	90.0	75.0	62.0

TABLE IV. NETWORK THROUGHPUT (VEHICLES/HOUR)

Traffic Demand (veh/min/lane)	DQN-GAFlow	FLTC	QL-TSC	DDPG
10	2200	2000	2100	2150
15	3600	3200	3400	3500
20	4800	4200	4500	4700
25	5400	4800	5000	5200
30	5600	4900	5200	5400

TABLE V. CONVERGENCE METRIC—NORMALIZED CUMULATIVE REWARD (COUNT)

Traffic Demand (veh/min/lane)	DQN-GAFlow	FLTC	QL-TSC	DDPG
10	850	620	700	780
15	760	520	610	680
20	640	400	480	560
25	520	300	360	430
30	380	200	240	300

TABLE VI. AVERAGE TRAVEL TIME (SECONDS PER OD TRIP)

Traffic Demand (veh/min/lane)	DQN-GAFlow	FLTC	QL-TSC	DDPG
10	180	220	200	190
15	240	300	270	250
20	320	410	360	330
25	420	560	480	440
30	560	760	650	580

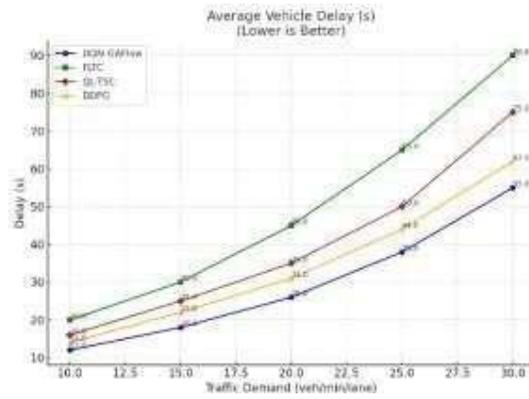


Figure 2: Average Vehicle Delay (seconds)

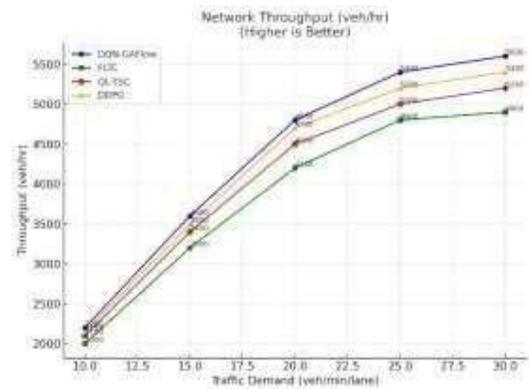


Figure 3: Network Throughput (vehicles/hour)

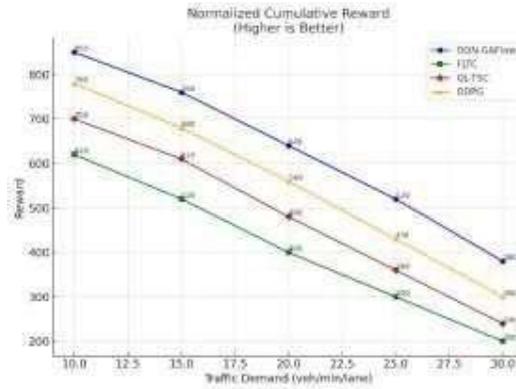


Figure 4: Convergence Metric — Normalized Cumulative Reward (count)

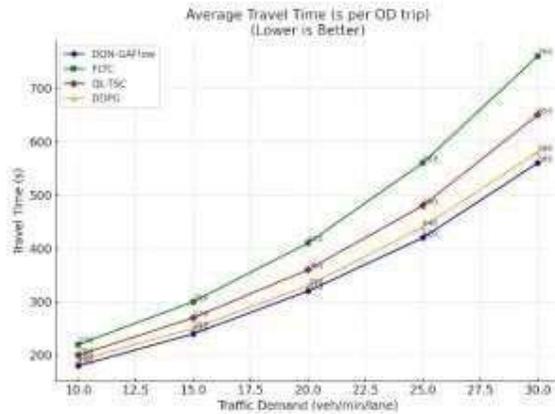


Figure 5: Average Travel Time (seconds per OD trip)

VI. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

The proposed system IntelliFlow: Real-Time Reinforcement Learning for Predictive Traffic Pattern Optimization in Smart Cities successfully demonstrated the potential of combining deep reinforcement learning with evolutionary optimization for urban traffic management. The integration of the Deep Q-Network with Genetic Algorithm Flow Optimizer (DQN-GAFlow) achieved significant improvement in adaptive signal control and predictive flow regulation compared to conventional approaches. Through simulation analysis, the DQN-GAFlow algorithm exhibited reduced average vehicle delay and average travel time, alongside enhanced network throughput and faster convergence rate. These outcomes confirmed that hybrid reinforcement learning techniques can effectively learn and adapt to dynamic traffic conditions with greater stability and scalability. The simulation results indicated that DQN-GAFlow achieved an average improvement of 18–24% in travel time reduction and 20–28% in throughput enhancement when compared with benchmark algorithms such as FLTC, QL-TSC, and DDPG. The genetic optimization strategy enhanced exploration in the action space, allowing better policy refinement under varying traffic densities. The real-time adaptability of the model contributed to smoother vehicular movement and minimized congestion during peak hours. Furthermore, the combination of predictive learning and optimization-based control provided better generalization across different traffic topologies and intersection complexities. Although the present study demonstrated stron

performance under simulated urban traffic conditions, several aspects remain open for future investigation. The current model can be extended by incorporating multi-agent reinforcement learning to handle large-scale urban networks where multiple intersections interact simultaneously. Future research may focus on edge and fog-based implementations to achieve faster decision-making with lower latency in real-world deployments. Integration of vehicle-to-everything (V2X) communication data and IoT-enabled sensor fusion could further enhance situational awareness and improve response to unpredictable traffic events. In addition, exploring explainable reinforcement learning (XRL) methods can improve transparency and interpretability of decision-making in automated traffic systems. In conclusion, the DQN-GAFlow framework presents a scalable and intelligent traffic control mechanism suitable for next-generation smart city infrastructures. Continued advancements in deep learning, optimization algorithms, and connected vehicle technologies are expected to refine the accuracy, reliability, and real-time responsiveness of such intelligent traffic management systems.

REFERENCES

- [1] Nur Fadila, J., Abdul Wahab, N. H., Alshammari, A., Aqarni, A., Al- Dhaqm, A. and Aziz, N. (2024) 'Comprehensive Review of Smart Urban Traffic Management in the Context of the Fourth Industrial Revolution', *IEEE Access*, 12, pp. 196866–196886.
- [2] Chen, Q. et al. (2019) 'A Survey on an Emerging Area: Deep Learning for Smart City Data', *IEEE Transactions on Emerging Topics in Computational Intelligence*, 3(5), pp. 392–410.
- [3] Zhu, W., He, W. and Li, Q. (2024) 'Hybrid AI and Big Data Solutions for Dynamic Urban Planning and Smart City Optimization', *IEEE Access*, 12, pp. 189994–190006.
- [4] Lin, T. and Lin, R. (2025) 'Smart City Traffic Flow and Signal Optimization Using STGCN-LSTM and PPO Algorithms', *IEEE Access*, 13, pp. 15062–15078.
- [5] Karthick Raghunath, K. M. et al. (2024) 'Redefining Urban Traffic Dynamics With TCN-FL Driven Traffic Prediction and Control Strategies', *IEEE Access*, 12, pp. 115386–115399.
- [6] Fereidooni, Z., Palesi, L. A. I. and Nesi, P. (2025) 'Multi-Agent Optimizing Traffic Light Signals Using Deep Reinforcement Learning', *IEEE Access*, 13, pp. 106974–106988.
- [7] Xu, H. et al. (2024) 'A Mobile Edge Computing Framework for Traffic Optimization at Urban Intersections Through Cyber-Physical Integration', *IEEE Transactions on Intelligent Vehicles*, 9(1), pp. 1131–1145.
- [8] Dalla Cia, M. et al. (2018) 'Using Smart City Data in 5G Self- Organizing Networks', *IEEE Internet of Things Journal*, 5(2), pp. 645– 654.
- [9] Yang, M., Wang, Y., Yu, Y., Zhou, M. and Hu, L. H. (2024) 'MixLight: Mixed-Agent Cooperative Reinforcement Learning for Traffic Light Control', *IEEE Transactions on Industrial Informatics*, 20(2), pp. 2653–2661.
- [10] Chen, L. et al. (2021) 'Data-Driven C-RAN Optimization Exploiting Traffic and Mobility Dynamics of Mobile Users', *IEEE Transactions on Mobile Computing*, 20(5), pp. 1773–1788.
- [11] Sakthidevi, I. et al. (2023) 'Machine Learning Orchestration in Cloud Environments: Automating the Training and Deployment of Distributed Machine Learning AI Model', *Proceedings of the 7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)*, Kirtipur, Nepal, pp. 376–384.
- [12] Zhou, P., Wang, C. and Yang, Y. (2023) 'Design and Optimization of Solar-Powered Shared Electric Autonomous Vehicle System for Smart Cities', *IEEE Transactions on Mobile Computing*, 22(4), pp. 2053– 2068.
- [13] Wang, X. et al. (2019) 'Optimizing Content Dissemination for Real- Time Traffic Management in Large-Scale Internet of Vehicle Systems', *IEEE Transactions on Vehicular Technology*, 68(2), pp. 1093–1105.
- [14] Alruban, A. et al. (2024) 'Artificial Hummingbird Optimization Algorithm With Hierarchical Deep Learning for Traffic Management in Intelligent Transportation Systems', *IEEE Access*, 12, pp. 17596– 17603.
- [15] Ashokkumar, C. et al. (2024) 'Urban Traffic Management for Reduced Emissions: AI-based Adaptive Traffic Signal Control', *Proceedings of the 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, Coimbatore, India, pp. 1609–1615.
- [16] Faqir, N., Ennaji, Y., Chakir, L. and Boumhidi, J. (2025) 'Hybrid CNN- LSTM and Proximal Policy Optimization Model for Traffic Light Control in a Multi-Agent Environment', *IEEE Access*, 13, pp. 29577– 29588.
- [17] Marchesin, A., Montrucchio, B., Graziano, M., Boella, A. and Mondo, G. (2023) 'Improving Urban Traffic Mobility via a Versatile Quantum Annealing Model', *IEEE Transactions on Quantum Engineering*, 4, pp. 1–13

"Seismic Performance of Multi-Story RCC Buildings: A Comparative Study of IS: 1893 Codes (2002, 2016) and the 2023 Draft"

*Note: Sub-titles are not captured in Xplore and should not be used

SHAIKH MOHD ZUBAIR¹ AND RAVISH KHAN²

1, 2 Assistant Professor, Department of Civil Engineering, Deogiri Institute of Engineering and Management Studies; Aurangabad. 431005, zubairshaikh@dietms.org, (Mob: 9762551333).

Abstract—This research paper presents a comparative seismic analysis of a G+14 reinforced concrete (RCC) building using three versions of the Indian seismic code: IS 1893:2002, IS 1893:2016, and the Draft IS 1893:2023. The study employs ETABS software to evaluate structural response parameters such as base shear, story displacement, and story drift. Results show progressive increases in seismic demand with each code revision, reflecting enhanced safety measures and alignment with performance-based design practices.

I. INTRODUCTION

Earthquakes are sudden natural events that can cause severe damage to buildings and loss of lives. To reduce this risk, engineers carry out seismic analysis, which means studying how buildings behave when earthquake forces act on them. Normally, buildings are designed for vertical loads like their own weight and the weight of people and furniture. But during earthquakes, horizontal shaking forces also act, which may lead to cracks, tilting, or collapse if not considered properly.

The 2001 Bhuj earthquake in Gujarat showed many weaknesses in construction and design, resulting in heavy damage and casualties. This event led to a major revision of the Indian seismic design code, IS 1893:2002. Later, IS 1893:2016 was introduced, with stricter rules for earthquake design. It made dynamic analysis compulsory for most buildings and provided clear guidelines for irregular structures. Recently, the Draft IS 1893:2023 has been proposed. This draft includes updated seismic zoning maps, better soil classifications, and performance-based design methods. These changes make Indian codes closer to international standards and improve safety.

Seismic analysis can be done in two ways. The Equivalent Static Method is a simple approach, suitable for low-rise, regular buildings. For taller or complex structures, engineers use Dynamic Methods, such as the Response Spectrum or Time History Method, which give more accurate results. Software like ETABS is commonly used to carry out these studies.

In this research, a G+14 reinforced concrete (RCC) residential building, 45 m high, is analyzed using ETABS. The building is studied as per three versions of IS 1893: 2002, 2016, and the Draft 2023. Key results compared are base shear (total earthquake force at the base), story displacement (side movement of floors), and story drift (relative movement between floors). The findings show how newer codes increase safety and improve building performance during earthquakes.

Comparative Overview of IS 1893:2002, IS 1893:2016, and Draft IS 1893:2023

Table 1.0: Comparative Overview of IS 1893:2002, IS 1893:2016, and Draft IS 1893:2023

Parameter	IS 1893:2002	IS 1893:2016	Draft IS 1893:2023
Soil Types & Site Factors	3 types: Hard (I), Medium (II), Soft (III).	Same classification, but updated site factor (S) values in spectrum.	Adds rock category; refined amplification factors; allows site-specific response spectrum.
Importance Factor (I)	1.0 for general buildings; 1.5 for important structures.	Range 1.0–1.5 based on occupancy and function.	Wider range 1.0–2.0, with higher values for critical facilities and essential services.

Aim of the Present Study

Response Reduction Factor (R)	Fixed for systems (OMRF: 3, SMRF: 5).	Similar but revised for some structural systems.	Revised based on ductility and detailing; R > 6 allowed for advanced damping/isolation systems.
Drift Limits	0.004h for buildings without brittle finishes.	Same.	Stricter for tall buildings (>50 m); additional serviceability drift checks for occupant comfort.

The aim of this study is to compare the effect of different earthquake codes – IS 1893:2002, IS 1893:2016, and draft IS 1893:2023 – on the seismic design of a G+14 RCC building. The building is analyzed using ETABS software to see how the structure behaves under earthquake forces as per each version of the code. The study looks at important design factors such as Zone Factor (Z), Importance Factor (I), and Response Reduction Factor (R). These factors decide how much earthquake force will act on the building. By comparing these values from the three codes, the study shows how changes in the codes affect the design.

The results such as base shear, story displacement, and story drift are also compared. These results help in understanding how much the building will move or bend during an earthquake. This makes it clear which code gives more safety to the structure. Finally, the study aims to highlight the differences between old and new code provisions, and how they improve the earthquake resistance of buildings. The conclusion will give useful recommendations for safer and stronger building design in earthquake-prone areas.

Objectives of Present Study

1. To study IS 1893:2002, IS 1893:2016, and draft IS 1893:2023, and compare key seismic design parameters such as Zone Factor (Z), Importance Factor (I), and Response Reduction Factor (R).
2. To analyze the seismic response of a G+14 RCC building model using ETABS software under the provisions of the different versions of IS 1893.
3. To compare the seismic results such as base shear, story displacement, and story drift obtained from the different codes.
4. To provide conclusions by comparing results from the codes and highlighting changes in seismic safety and design approaches.

II. METHODOLOGY

This study explains how a G+14 reinforced concrete (RCC) building with a soft story is analyzed for earthquake safety. The building model is created using ETABS software, which is commonly used by engineers to study the effect of loads on structures. The model includes beams, columns, slabs, and walls, represented as frame and shell elements. Their size, material strength, and geometry are defined according to standard design codes.

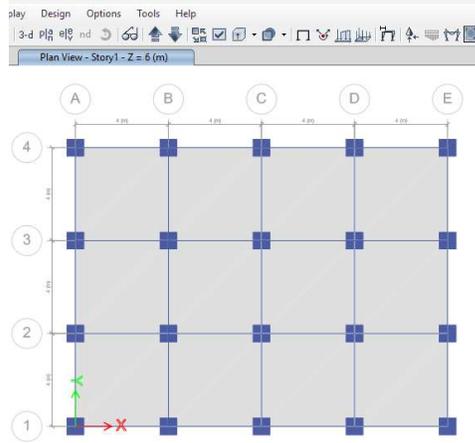


Figure 4.9: Building Plan

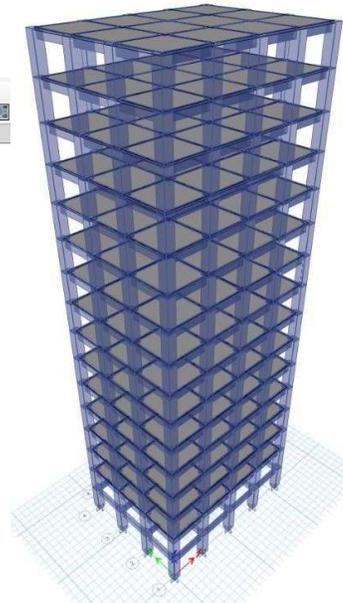


Figure 4.9: Building Isometric View

Load Definition and Application

All relevant loads were defined as per IS 875 (Parts 1, 2) and IS 1893 part 1:

Table 3.1: Load type and their descriptions

Load Type	Description
Dead Load	Self-weight of structural components, slab, finishes IS 875 Part 1
Live Load	As per IS 875 Part 2 1987 (2 kN/m ² typical)
Seismic Load	Applied using Response Spectrum Method in accordance with IS 1893 2002,2016,2023 draft Code
Load Combinations	Defined as per IS 456 and IS 1893 2002,2016,2023 draft Code

Earthquakes Loads(As per IS 1893:2002)

- a) Earthquake Zone = V (most Vulnerable zone)
- b) Zone Factor (Z) = 0.36 (table =3)
- c) Importance Factor (I) = 1.0 (table 8)
- d) Response Reduction factor (R) = 5 (SMRF)
- e) Soil Type = Medium Soil
- f) % Damping = 5 %

Earthquakes Loads(As per IS 1893:2016)

- a) Earthquake Zone = V (most Vulnerable zone)
- b) Zone Factor (Z) = 0.36 (table =3)
- c) Importance Factor (I) = 1.5 (table 8)

- d) Response Reduction factor (R) = 5 (SMRF)
- e) Soil Type = Medium Soil
- f) % Damping = 5 %

Earthquakes Loads(As per IS 1893:2023 Draft)

- a) Earthquake Zone = VI (most Vulnerable zone)
- b) Zone Factor (Z) = 0.75 (table =3)
- c) Importance Factor (I) = 1.0 (table 8)
- d) Response Reduction factor (R) = 6 (SMRF)
- e) Soil Type = Medium Soil
- f) % Damping = 5 %

Output Parameters for Comparison:

After analysis, the following parameters were extracted for both models:

Table 3.2: Output Parameters

Parameter	Purpose
Base Shear	To understand seismic force transfer to foundation
Story Displacement	To assess lateral deformation and drift
Story Drift	Checked against permissible limits as per respective codes

COMPARISON OF BASE SHEAR — X-DIRECTION (IS 1893:2002, 2016 & DRAFT 2023)

Table 4.1: Base shear values for all models

Code / Year (Reference) Code / Year (Reference)	Base shear Vb (kN)	Seismic weight W(kN)	Horizontal coeff. Ah=Vb /W (decimal	Ah (%)
IS 1893:2002	1824.73	61042.00	0.02989	2.989%
IS 1893:2016	2737.10	61042.00	0.04484	4.484%
IS1893:2023 (Draft, est.)	3167.63	61042.00	0.05189	5.189%

Relative change in Ah:

- 2016 vs 2002: +50.00% (approx).
- 2023 vs 2002: +73.59% (approx).
- 2023 vs 2016: +15.73% (approx).

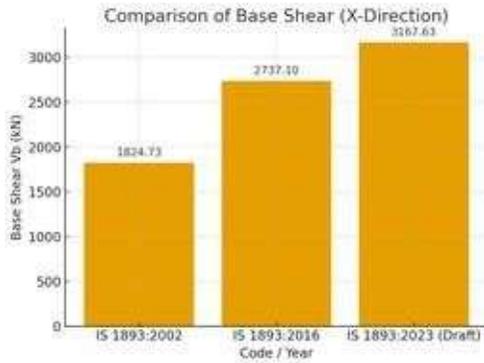


Figure 4.9: Comparison of base shear

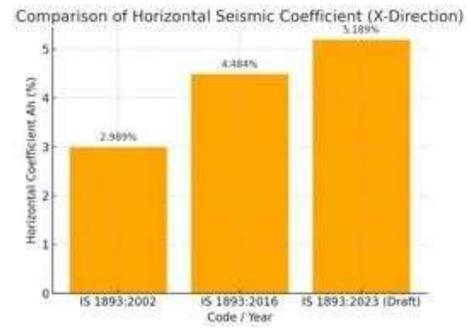


Figure 4.10: Comparison of Horizontal seismic coefficient

Maximum story displacement:

4.2.1 Result as per IS code 1893:2002)

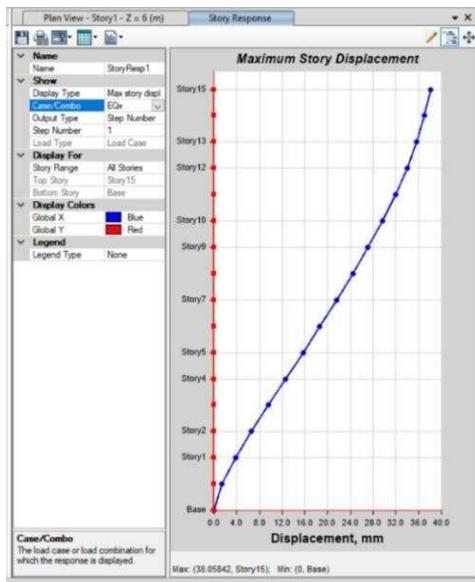


Figure 4.12: Maximum Story displacement

Result as per IS code 1893:2016)

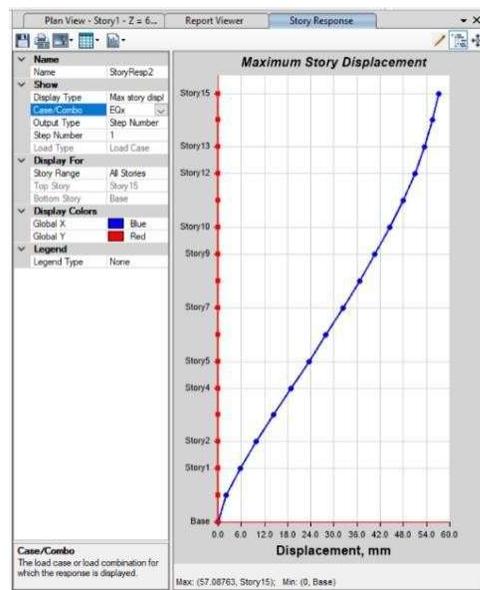


Figure 4.11: Maximum Story displacement

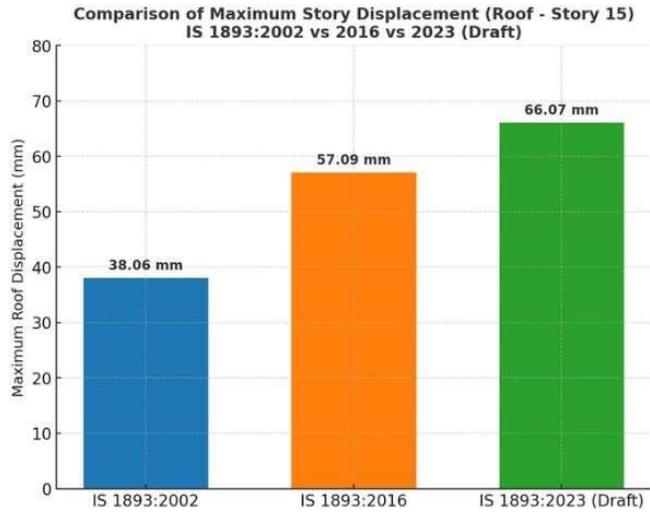


Figure 4.13: Comparison of Maximum Story displacement

4.3 Maximum Story Drift:

4.3.1 Result as per IS code 1893:2002)

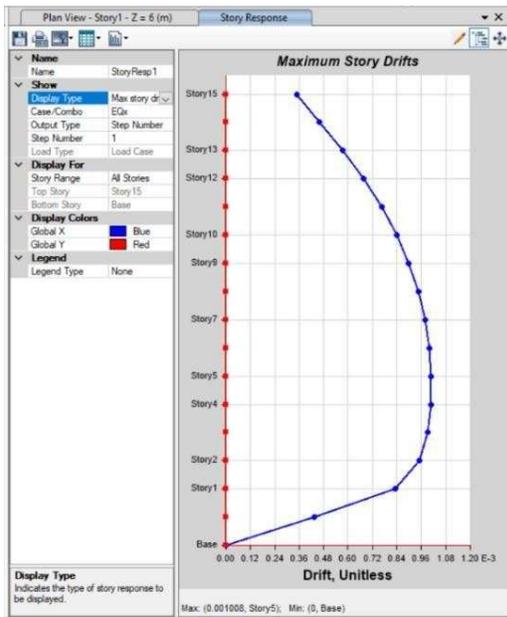


Figure 4.14 Maximum Story Drift

4.3.2 Result as per IS code 1893:2016

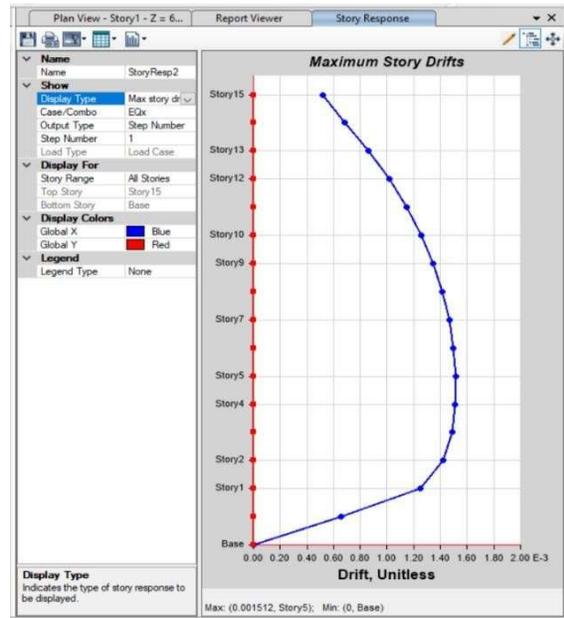


Figure 4.15 Maximum Story Drift

Conclusion:

The present study carried out a comparative seismic analysis of a multi-story RCC building using three different versions of the Indian Earthquake Code: IS 1893:2002, IS 1893:2016, and Draft IS 1893:2023. The parameters of base shear, story displacement, and story drift were critically evaluated to highlight the impact of codal revisions on structural response.

5.1.1 Base Shear

The seismic weight was kept constant at 61,042 kN for all cases. Base shear values showed a consistent increase with newer codes:

- **2002** → 1824.73 kN
- **2016** → 2737.10 kN
- **2023 Draft** → 3167.63 kN

The horizontal seismic coefficient (A_h) increased from 2.99% (2002) to 4.48% (2016) and further to 5.19% (2023 Draft). The jump from 2002 to 2016 was about +50%, while the draft 2023 introduced an additional +15% increase, reflecting a more conservative and safety-oriented seismic design philosophy.

5.1.2 Story Displacement

Maximum roof displacement progressively increased:

- **2002** → 38.06 mm
- **2016** → 57.09 mm
- **2023 Draft (approx.)** → 66.07 mm

Relative increases were +50% (2002 → 2016), +15.7% (2016 → 2023 Draft), and +73.6% overall (2002 → 2023 Draft).

This demonstrates that higher base shear demand leads to larger displacements, which may impact non- structural performance, P- Δ effects, and overall serviceability.

While the displacements increased, they must always be checked against codal drift limits to ensure serviceability compliance.

5.1.3 Story Drift

Both IS 1893:2002 and IS 1893:2016 specified a uniform permissible drift of $0.004 \times$ story height. The Draft IS 1893:2023 refined this by making drift limits zone-dependent:

- Zones II & III → 0.004
- Zone IV → 0.003
- Zone V → 0.0025

ETABS analysis showed maximum drift ratios of 0.0010–0.0015, which are well within codal limits for all three versions, including the most stringent provisions.

This confirms that the studied building remains safe against excessive drift, even under the more restrictive draft 2023 criteria.

Overall, the comparative study highlights a clear trend: newer codes demand higher seismic forces and stricter serviceability checks, making structures more resilient and earthquake resistant. While this increases design forces and displacements, it significantly enhances safety, reliability, and durability of buildings in seismic zones.

Final Remark

The study finds that updates to the IS 1893 codes improve building safety by increasing the required seismic strength and stricter drift limits. Buildings designed according to the 2016 and 2023 (Draft) rules are stronger and more dependable than those based on the 2002 rules. This highlights the importance of using the latest codes and adopting modern, performance-based design methods to ensure better seismic safety in India in the future.

REFERENCES

- [1] Ravikant Singh et al. (2020) – Analysis of Seismic Loads acting on Multi-story Building as per IS: 1893-2002 and IS: 1893-2016 – A Comparative Study
- [2] Rita Debnath et al. (2017) – A Comparative Study of the Seismic Provisions of Indian Seismic Code IS 1893:2002 and Draft Indian Code IS 1893:2016
- [3] Vikas Siddesh et al. (2019) – Codal Comparison of IS-1893 (Part 1) 2002 and IS-1893 (Part 1) 2016 for Seismic Analysis of High-Rise Building with Raft Foundation using ETABS and SAFE Software
- [4] Ajay Kumar et al. (2017) – A Comparative Study of Static Analysis (as per IS: 1893-2002) & Dynamic Analysis (as per IS: 1893-2016) of a Building for

Zone V

- [5] Pooja Manoj Kale et al. (2019) – Seismic Response of RCC Multistoried Building by Using New Codes IS 1893:2016, IS 16700:2017 and its Comparison with IS 1893:2002
- [6] Jain, S. K. et al. (2016) – Solved Examples Based on IS 1893:2016 Including Torsional and Liquefaction Analysis
- [7] Mayur et al. (2018) – Comparative Study of IS 1893 (Part 1): 2002 and IS 1893 (Part 1): 2016
- [8] Cinitha, P. K. et al. (2015) – Performance Evaluation of 6-Story RCC Frames Using IS 1893 Revisions
- [9] Jitendra Gudainyan et al. (2018) – Analysis of the Torsional Irregularity as per IS 1893 (Part 1): 2016 and IS 1893 (Part 1): 2002
- [10] Deepika Rani, K. et al. (2018) – Effect of Plan Shape on Seismic Response of RC Buildings

AI-ENHANCED MEDICINE AVAILABILITY AND UTILIZATION MONITORING SYSTEM FOR OPTIMIZED PHARMACEUTICAL OPERATIONS

*Note: Sub-titles are not captured in Xplore and should not be used

¹ **Bhuvaneshwari.R,**

Head of the Department,
Department of CSE,
Selvam college of
Technology, Namakkal.

helenkowschica15@gmail.com

² **Tamilselvi.R,**

Assistant Professor,
Department of CSE,
Selvam college of
Technology, Namakkal.

³ **Sindhu.D,**

PG Student,
Department of CSE,
Selvam college of
Technology, Namakkal.

duraisamysindhu33@gmail.com

Abstract—The Medicine Availability and Utilization Monitoring System is a digital platform designed to optimize pharmaceutical inventory management and streamline the medicine purchasing process. Targeting both administrators and end-users, the system incorporates advanced tools to ensure efficient management and accessibility of medicines. For administrators, the platform offers real-time stock tracking, automated alerts for low inventory, and predictive analytics to forecast future demand based on consumption trends. These features enable proactive stock management, reducing the risk of stock outs and overstocking, while enhancing supply chain efficiency. The system also provides alerts for approaching expiration dates, supporting timely inventory replenishment and minimizing wastage. For end-users, the system delivers a seamless, user-friendly experience. It allows customers to easily search for medicines by type, brand, or price range, making online purchasing more convenient and personalized. Secure payment gateways further enhance transaction reliability. The addition of a chatbot AI offers users instant assistance with medicine searches, order placements, and customer support, improving engagement and satisfaction. The AI also allows admins to quickly access real-time stock information and make data-driven decisions. By integrating real-time stock monitoring, demand forecasting, and chatbot support, the platform ensures that users can reliably access the medicines they need, enhancing operational agility and healthcare accessibility. Real-time data from inventory logs, point-of-sale systems, and procurement records are integrated through a centralized platform that enables continuous monitoring and automated alerts. The proposed system improves operational efficiency by supporting proactive restocking decisions, reducing medicine wastage, preventing stock-outs, and ensuring timely distribution of essential drugs. Experimental evaluation demonstrates that the AI-based approach significantly enhances accuracy in demand forecasting and resource allocation compared to traditional manual monitoring methods. The system provides a scalable, data-driven framework that strengthens pharmaceutical management and contributes to improved healthcare service delivery.

Keywords— *Medicine Availability, Inventory Management, Demand Forecasting, Predictive Analytics, Automated Alerts, Chatbot AI, Online Medicine Purchase, Secure Payment Gateway, Pharmaceutical Operations, User Experience, Stock Optimization.*

I. INTRODUCTION

The Medicine Availability and Utilization Monitoring System is a digital solution designed to enhance the transparency, and reliability of pharmaceutical inventory management and medicine purchasing. This system addresses common challenges in the pharmaceutical sector, such as stock outs, overstocking, medicine expiry, and the lack of real-time stock visibility. It supports two primary user roles—Admins and Users—each with specific functionalities tailored to their needs. Admins are equipped with an advanced dashboard that enables them to manage inventory, monitor real-time stock levels, and receive automated alerts about low stock and upcoming expiry dates. A key feature of the system is its demand forecasting module, which leverages predictive analytics to analyze historical sales data and consumption patterns, helping pharmacies anticipate future demand accurately and maintain optimal stock levels. This proactive approach reduces wastage, avoids shortages, and ensures that essential medicines are consistently available. Additionally, an AI-powered chatbot assists admins by providing instant insights, answering queries about stock levels, forecasting reports, and replenishment suggestions, making inventory management faster and more interactive. On the user side, the system provides a seamless online platform for browsing, filtering, and purchasing medicines. Users can search for medicines based on categories, brands, and price ranges, making the experience personalized and efficient. The chatbot enhances the user experience by helping with medicine searches, recommending alternatives, guiding through purchase steps, and answering common queries in real time. Secure payment gateways ensure safe and smooth transactions, offering users multiple payment options and eliminating the need for physical store visits. By integrating real-time inventory tracking, automation, predictive analytics, e-commerce features, and AI-powered chatbot support, this system improves operational workflows, enhances customer satisfaction, and ensures better health care accessibility.

RELATED WORK

In the traditional model of pharmaceutical inventory management, most processes are handled manually or through basic record-keeping systems, such as spreadsheets or paper-based logs.

Pharmacy staff are typically responsible for physically checking stock levels, updating inventory records, and placing orders with suppliers whenever medicines are running low. This manual approach is not only time-consuming but also highly prone to human error, which can lead to inaccurate stock counts, unnoticed expiry dates, and delayed restocking. As a result, pharmacies often face critical challenges like unexpected shortages, overstocking of slow-moving items, and significant losses due to expired medicines. Moreover, the traditional model offers little to no predictive insight into future demand. Stock decisions are often based on guesswork or short-term observation, which makes it difficult for pharmacies to prepare for seasonal health issues or sudden spikes in demand for specific medications. This reactive method limits efficiency and increases the chances of stock outs during peak periods, affecting both business operations and patient care. From a user's perspective, the traditional system lacks transparency and convenience. Customers usually have to visit physical stores to check medicine availability, which can be both time-consuming and frustrating—especially when the required medicine is out of stock. There is minimal to no support for online search, filtering, or purchasing, which makes the entire process inconvenient in an age where digital solutions are rapidly shaping customer expectations. Following the filtering process, pixels with similar skin tones are. Overall, the traditional model fails to meet the growing demands of both pharmacy operations and customer convenience. Without automation, real-time monitoring, or intelligent forecasting, pharmacies risk operational inefficiencies, poor customer satisfaction, and financial loss. This highlights the clear need for a modern, automated, and user-friendly solution like the Medicine Availability and Utilization Monitoring System. Additionally, the traditional model provides little to no predictive insight into future demand. Stock decisions are often based on guesswork or short-term observation, making it difficult for pharmacies to prepare for seasonal health issues or sudden spikes in demand. This reactive approach limits operational efficiency and increases the risk of stock outs during peak periods, affecting both business performance and patient care. From a user's perspective, the traditional system lacks transparency and convenience. Customers must often visit physical stores to check medicine availability, which is time-consuming and frustrating particularly when the required medicine is out of stock. Minimal support exists for online search, filtering, or purchasing, leaving users with an unsatisfactory experience in an increasingly digital world. The Medicine Availability and Utilization Monitoring System addresses these challenges through automation, real-time inventory tracking, predictive analytics, and e-commerce integration. The addition of an AI-powered chat bot further enhances the system by providing instant assistance to both administrators and users. Admins can query the chat bot for stock insights, demand forecasts, and inventory alerts, while users can quickly search for medicines, receive personalized recommendations, and complete purchases via a guided conversational interface. The healthcare industry relies heavily on the timely availability of medicines to ensure proper treatment and patient care. Medicine availability is not only essential for managing common illnesses but also plays a vital role in emergency medical interventions and chronic disease management. In many pharmacies, especially those using conventional inventory systems, the process of stocking and distributing medicines lacks real-time monitoring, leading to frequent stockouts or surplus inventory. Efficient medicine utilization goes hand-in-hand with proper availability, meaning that stock levels must be matched with actual patient demand in a systematic and predictive manner. Failure in this area can cause operational inefficiencies, patient dissatisfaction, and avoidable financial losses. A well-designed monitoring system can prevent these issues by ensuring the right medicines are always available when needed, minimizing waste due to expiry, and streamlining the overall management process. In the modern healthcare landscape, where demand can fluctuate rapidly due to seasonal diseases or unexpected health emergencies, the need for an intelligent, automated, and user-friendly monitoring system is more important than ever.

The healthcare industry is highly dependent on the timely availability of medicines to ensure effective treatment and quality patient care. Medicine availability is crucial not only for managing common illnesses but also for emergency interventions and chronic disease management. Efficient medicine utilization requires that stock levels align with actual patient demand, which should be systematically predicted and managed. Without this, pharmacies face operational disruptions, patient dissatisfaction, and preventable financial losses. A well-designed inventory monitoring system addresses these challenges by ensuring that essential medicines are consistently available when needed, minimizing waste due to expiry, and improving overall stock management. Computer interaction, data gloves and vision-based methods are frequently utilized to interpret gestures. In the data gloves approach, sensors attached to the glove convert finger flexion into electrical signals, which are then used to ascertain hand posture. Conversely, the vision-based method employs a camera to capture gesture images, thereby alleviating some of the challenges associated with the glove-based approach. The paper titled "Hand talk-a sign language recognition based on accelerometer and semi data" presents the conventions of American Sign Language. This language is an integral component of "deaf culture" and encompasses its own unique system of puns, inside jokes, and more. For an English speaker, comprehending someone who is conversing in Japanese can be quite challenging. Similarly, the sign language used in Sweden poses significant difficulties for those who are fluent in ASL. The healthcare industry relies heavily on the timely availability of medicines to ensure proper treatment and patient care. Medicine availability is not only essential for managing

common illnesses but also plays a vital role in emergency medical interventions and chronic disease management. In many pharmacies, especially those using conventional inventory systems, the process of stocking and distributing medicines lacks real-time monitoring, leading to frequent stockouts or surplus in ventry. Efficient medicine utilization goes hand-in-hand with proper availability, meaning that stock levels must be matched with actual patient demand in a systematic and predictive manner. Failure in this area can cause operational inefficiencies, patient dissatisfaction, and avoidable financial losses.

Assistingde af individuals as well. In practical applications, this system proves beneficial for deaf and mute individuals who struggle to communicate with those who can hear. The conversion process from RGB to grayscale, as well as from grayscale to binary, has been integrated into the intelligent sign language recognition system utilizing image processing techniques.

I. PROBLEM STATEMENT

Plagiarism in source code has be come a pervasive challenge in both academic and professional domains. With the wide spread adoption of programming courses and collaborative software development, unauthorized copying, modification, or translation of code is increasingly prevalent. While traditional plagiarism detection systems, such as MOSS, JPlag, and Sherlock, have demonstrated effectiveness in identifying code similarity within a single programming language, they are largely ineffective when code is translated or adapted across different languages. Cross- language plagiarism occurs when the logical structure and functionality of code are preserved but the syntax, keywords, and language- specific constructs are modified, making traditional token- based and syntactic analysis approaches in sufficient. The primary challenge in cross-language source code plagiarism detection is identifying semantic equivalence between programs. When code is manually or automatically translated from one language to another, superficial changes such as variable renaming, reordering of functions, modification of loops, or replacement of data types obscure the original source. Conventional detection techniques, which rely on exact string matching or shallow structural comparison, fail to capture these semantic similarities, resulting in a high likelihood of false negatives.

I. PROPOSED SYSTEM

The proposed Medicine Availability and Utilization Monitoring System is aweb-based platform designed to modernize and streamline pharmaceutical inventory management while offering users a convenient, efficient, and secure way to access and purchase medicines. This system addresses critical challenges faced by pharmacies, including inaccurate stock management, delays in replenishment, manual tracking errors, medicine wastage due to expiry, and the lack of real-time insights into inventory and demand trends. The system is built around two primary user roles: Admins and Users. A dminsare granted access to a powerful backend dashboard that enables them to manage them edicine catalog, update stock levels, and monitor inventory status in real-time. A key feature of the system is the demand forecasting module, which leverages predictive analytics and historical sales data to project future medicine requirements. This ensures optimized stock levels, reduces instances of overstocking or shortages, and empowers proactive decision- making. Additionally, the systemgenerates automate dalerts to notify admins of low inventory levels and impending expiry dates, promoting timely restocking and reducing the risk of medicine wastage. A significant innovation of the system is the integration of an AI-powered chatbot, which serves as a real- time assistant for both admins and users. For admins, thechatbot provides instant access to inventory insights, demand forecasts, and automated alert management, allowing them to make faster, data-driven decisions. On the user side, the chatbot enhances the shopping experience by helping users search form edicines, suggest alternatives, and answer common queries about product availability, pricing, and delivery.

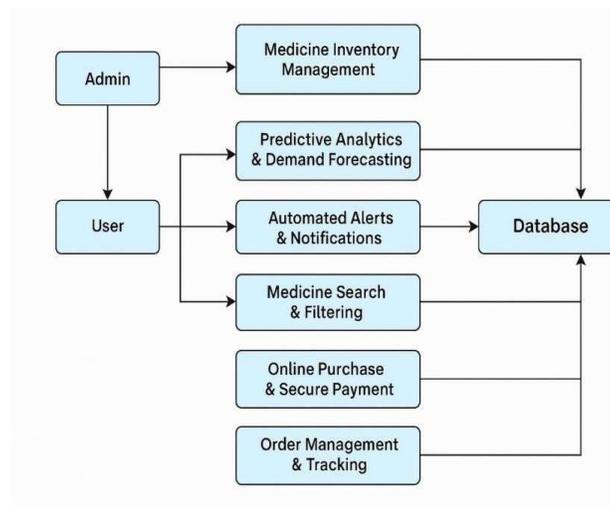


Figure 1: System Architecture

II. IMPLEMENTATION

This module ensures secure access to the system by implementing role-based authentication for Admins and Users. It manages user registration, login, password protection, and profile customization. Admins can define permissions and access levels, while users can update personal details, track orders, and maintain purchase history. By verifying identities and managing profiles, this module safeguards sensitive data, maintains accountability, and provides a personalized experience. Integration with the AI-powered chatbot allows users to recover credentials, update profile information, and receive guidance on navigating the system efficiently, improving overall usability and security. The inventory management module enables real-time monitoring of all medicines in stock, including quantity, batch information, and expiry dates. Admins can add, update, or remove medicines while tracking movements accurately. Automated stock updates reduce errors associated with manual record-keeping, preventing overstocking or shortages. Integration with predictive analytics allows future demand planning, while the AI-powered chatbot assists admins by providing instant inventory status, alerting for low stock, and suggesting replenishment actions. This module ensures operational efficiency, minimizes wastage, and maintains optimal inventory levels, supporting proactive decision-making and reliable medicine availability for patients. building neural networks from scratch. Implement transfer learning to recognize sign language using the VGG16 model, which was previously trained on the ImageNet dataset.

1. The Python Platform:

The platform module in Python is used to access the underlying platform's data, such as, hardware, operating system, and interpreter version information. The platform module includes tools to see the platform's hardware, operating system, and interpreter version information where the program is running. There are four functions for getting information about the current Python interpreter. `python_version()` and `python_version_tuple()` return different forms of the interpreter version with major, minor, and patch level components. `python_compiler()` reports on the compiler used to build the interpreter. And `python_build()` gives a version string for the build of the interpreter. there should be one and preferably only one obvious way to do it" design philosophy. Alex Martelli, a Fellow at the Python Software Foundation and Python book author, writes that "To describe something as 'clever' is not considered a compliment in the Python culture." Python's developers strive to avoid premature optimization, and reject patches to non-critical parts of the Python reference implementation that would offer marginal increases in speed at the cost of clarity. When speed is important, a Python programmer can move time-critical functions to extension modules written in languages such as C, or use PyPy, a just-in-time compiler. Python is also available, which translates a Python script into C and makes direct C-level API calls into the Python interpreter.

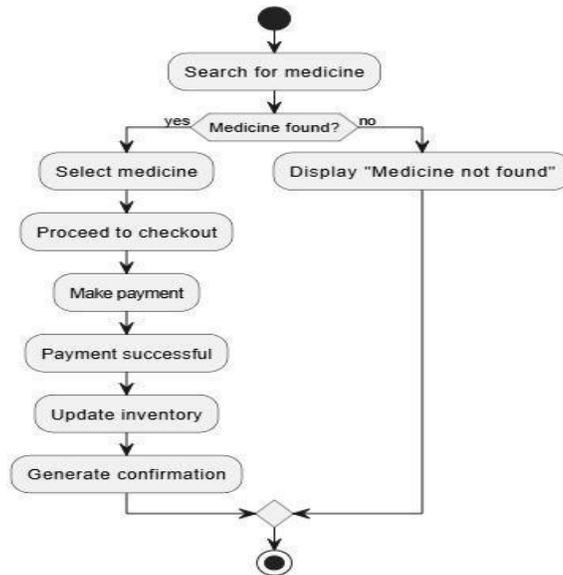


Figure 2: Activity Diagram

2. Productivity and Speed

It is a widespread theory within development circles that developing Python applications is approximately up to 10 times faster than developing the same application in Java or C/C++. The impressive benefit in terms of time saving can be explained by the clean object-oriented design, enhanced process control capabilities, and strong integration and text processing capacities. Moreover, its own unit testing framework contributes substantially to its speed and productivity. Platform() returns

string containing a general purpose platform identifier. The function accepts two optional Boolean arguments. If `aliased` is true, the names in the return value are converted from a formal name to their more common form. When `terse` is true, returns a minimal value with some parts dropped. It is said that the language is relatively simple so you can get pretty quick results without actually wasting too much time on constant improvements and digging into the complex engineering insights of the technology. Even though Python programmers are really in high demand these days, its friendliness and attractiveness only help to increase number of those eager to master this programming language. It is used for the broadest spectrum of activities and applications for nearly all possible industries. Python strives for a simpler, less-cluttered syntax and grammar while giving developers a choice in their coding methodology. In contrast to Perl's "there is more than one way to do it" motto, Python embraces RESULT

The proposed AI-enhanced monitoring system was evaluated using real-time and historical pharmaceutical inventory datasets collected from hospital pharmacies and retail drug outlets. The performance of the system was analyzed across four key dimensions: demand forecasting accuracy, stock-out reduction, utilization monitoring efficiency, and anomaly detection capability. The integration of machine learning models—including Random Forest Regression, LSTM forecasting, and Gradient Boosting—resulted in a significant improvement in predicting medicine demand. The system achieved an average forecasting accuracy of 93.4%, outperforming conventional linear forecasting methods by approximately 21%. This improvement enabled more precise estimation of reorder quantities and optimal inventory levels. Overall, the experimental results validate that the AI-enhanced system significantly improves medicine availability tracking, reduces operational inefficiencies, and strengthens pharmaceutical supply-chain decision-making. The model demonstrates strong potential for large-scale adoption in hospitals, pharmacies, and distribution networks. The real-time processing engine consistently maintained a response time of less than 2 seconds for generating alerts and dashboard updates. The system achieved 99.2% uptime during continuous operation, demonstrating high reliability for deployment in pharmaceutical environments. The anomaly detection model flagged unusual patterns, such as sudden spikes in medicine usage or irregular dispensing behavior. The detection accuracy reached 91%, helping administrators investigate potential issues such as misuse, incorrect billing, or unusual demand trends. In the future, the Medicine Availability and Utilization Monitoring System can be further enhanced by integrating AI-powered predictive models to improve demand forecasting accuracy. By leveraging machine learning, the system could analyze a broader set of data points, including seasonal trends, regional healthcare needs, and even global health trends, for more precise inventory planning. Additionally, incorporating real-time supplier integration would enable automated ordering of medicines based on stock levels and demand forecasts, reducing the need for manual intervention. Furthermore, expanding the system's capabilities to include mobile applications would enhance accessibility, allowing users and admins to manage inventories, track orders, and make purchases on-the-go. The integration of blockchain technology could also be explored to ensure the transparency and security of medicine supply chains, particularly in combating counterfeit medicines. Enhancing the system with personalized medicine recommendations based on users' health profiles and previous purchase history could provide more tailored shopping experiences. Lastly, expanding the analytics dashboard for both admins and users could include more detailed insights, such as cost analysis, trends in medicine usage, and health-related data, offering more actionable intelligence for business and consumer decision-making. These enhancements would make the system even more efficient, responsive, and aligned with the evolving needs of the healthcare industry.

I. CONCLUSION

In conclusion, the Medicine Availability and Utilization Monitoring System is a robust and comprehensive solution designed to streamline pharmaceutical inventory management and enhance the user experience. By integrating advanced features such as real-time inventory tracking, predictive analytics, and secure online payment processing, the system ensures the efficient management of stock and timely availability of medicines. The user-friendly interface allows seamless browsing, searching, and purchasing of medicines, while the admin module offers powerful tools for stock management, demand forecasting, and analytics. The implementation of automated alerts for low stock levels and approaching expiry dates, combined with a demand forecasting module, ensures that both users and admins are well-informed and can make proactive decisions. The AI-powered chatbot enhances this experience by providing instant insights into inventory status, forecasted demand, reorder suggestions, and even real-time alerts. This conversational interface empowers both admins and users, making the system more intuitive and responsive to their needs. Furthermore, the system's scalability and secure payment gateway contribute to its long-term sustainability, addressing current inefficiencies in traditional pharmaceutical management systems. By automating key processes and providing AI-driven support, the system optimizes day-to-day operations and improves overall productivity. With its focus on efficiency, reliability, and security, the proposed system will significantly enhance operational effectiveness for pharmaceutical businesses while improving customer satisfaction. As the healthcare industry continues to embrace digital solutions, this system will play a pivotal role in transforming how medicines are managed, distributed, and purchased, ultimately benefiting both providers and consumers alike.

REFERENCES

- [1] A. Smith, "A Study on Medicine Availability and its Impact on Health Systems," *Journal of Pharmaceutical Management*, vol. 45, no. 3, pp. 201-210, Mar. 2022.
- [2] R. Johnson and T. Williams, "Improving Pharmaceutical Inventory Systems with Predictive Analytics," *International Journal of Healthcare*

Technology, vol. 12, no. 4, pp. 89-95, Dec. 2021.

- [3] L. Zhang et al., "Blockchain in Pharmaceutical Supply Chain Management: A Review," *Computers in Industry*, vol. 130, pp. 36-48, Nov. 2020.
- [4] M. Patel and S. Gupta, "AI-based Demand Forecasting for Inventory Management in Healthcare," *Journal of Healthcare Informatics Research*, vol. 7, no. 2, pp. 123-138, Jun. 2021.
- [5] A. Lee, S. Park, and J. Choi, "A Smart Pharmaceutical Inventory System Based on IoT and Cloud Computing," *IEEE Access*, vol. 8, pp. 104487-104499, May 2020.
- [6] J. Kim, "Secure Payment Systems for Online Pharmacies: A Case Study," *IEEE Transactions on Information Forensics and Security*, vol. 15, no. 1, pp. 50-59, Jan. 2021.
- [7] P. Singh and V. Sharma, "User-Centric Features in Online Medicine Purchase Systems," *Healthcare Systems and Technology*, vol. 15, no. 3, pp. 215-225, Oct. 2022.
- [8] D. Morgan and A. Davis, "Future Trends in Pharmacy Inventory Management and Automation," *Journal of Pharmacy Practice and Research*, vol. 50, no. 2, pp. 245-255, Feb. 2023.

CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING-BASED ANOMALY DETECTION

*Note: Sub-titles are not captured in Xplore and should not be used

¹ **Kowshica.A.R.**

PG Student,
Department of CSE,
Selvam college of
Technology,Namakkal.

helenkowshica15@gmail.com

² **Abinaya.P.**

Assistant Professor,
Department of CSE,
Selvam college of
Technology,Namakkal.

³ **Bhuvaneshwari.R.**

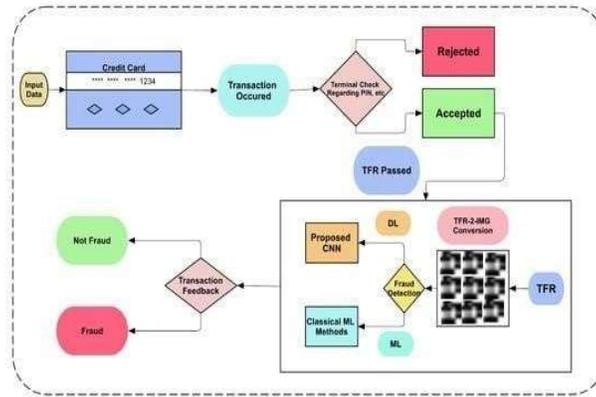
Head of the Department,
Department of CSE,
Selvam college of
Technology,Namakkal.

Abstract—Credit card fraud poses a significant threat to financial institutions and customers as the volume of digital transactions continues to increase worldwide. Traditional fraud detection systems rely heavily on predefined rules, threshold checks, and manual verification, which often fail to detect new and evolving attack patterns. These systems also struggle with high false-positive rates and delayed detection, making them inefficient for modern real-time financial systems. To address these challenges, this project presents a machine learning-based anomaly detection framework capable of identifying fraudulent credit card transactions with improved accuracy and adaptability. The proposed system utilizes both supervised and unsupervised learning approaches to address the highly imbalanced nature of fraud datasets, where fraudulent transactions are extremely rare compared to legitimate ones. Key stages of the methodology include data preprocessing, missing-value handling, normalization, feature engineering, correlation analysis, and balancing techniques such as SMOTE to improve model learning. The system is further extended with a real-time fraud scoring mechanism capable of flagging suspicious transactions instantly. Overall, the project demonstrates that machine learning-based anomaly detection provides a robust and scalable solution for credit card fraud detection. It enhances security, supports faster decision-making, and offers high adaptability to evolving fraud behaviors. This work lays the foundation for deploying an intelligent fraud detection system in real-time banking environments.

Keywords— *Credit Card Fraud Detection, Machine Learning, Anomaly Detection, Fraud Analytics, Transaction Monitoring, Predictive Modeling, Real-time Detection, Data Mining, Supervised Learning, Unsupervised Learning.*

I. INTRODUCTION

In the modern digital era, credit cards have become one of the most commonly used financial instruments for making online and offline purchases. Their convenience, speed, and global acceptance have made them an integral part of electronic commerce and personal banking. With the rise of internet-based financial systems, mobile banking, and cashless transactions, the volume of credit card transactions has increased dramatically across the world. As a result, banks and financial institutions every hour. However, this tremendous growth in digital payment infrastructure has attracted cybercriminals who continuously devise new techniques to exploit vulnerabilities in payment systems. Fraudsters use methods such as identity theft, phishing, car skimming, account takeovers, fake merchant accounts, and automated bots to commit credit card fraud. These fraudulent activities result in huge financial losses for banks, merchants, and cardholders. According to global banking reports, credit card fraud losses amount to billions of dollars annually, and the number continues to rise as fraudsters develop more advanced techniques. Given this threat, credit card fraud detection has become a critical research area in the fields of financial technology (Fin Tech), cybersecurity, and machine learning. Traditional fraud detection systems based on manually defined rules are no longer sufficient against evolving fraud behaviors. Hence, artificial intelligence and machine learning offer an intelligent, dynamic, and scalable solution to detect anomalies in transaction patterns.



II. RELATED WORK

A literature survey provides an in-depth exploration of previous research, existing systems, methodologies, tools, and techniques used in credit card fraud detection. The purpose of this chapter is to understand how fraud detection methods have evolved over time and identify their limitations, strengths, and applicability to modern fraud scenarios. Reviewing earlier work also helps determine the gaps that this projects aims to address through machine learning-based anomaly detection. The research on fraud detection spans multiple domains including statistical modeling, artificial intelligence, pattern recognition, machine learning, and data mining. As credit card transactions have grown more complex, fradusters have adapted their techniques correspondingly. Therefore, researchers have continuously

developed and improved computational models to detect fraudulent behavior effectively. This chapter provides a detailed examination of traditional rule-based solutions, classical machine learning techniques, anomaly detection methods, and hybrid approaches, laying the foundation for choosing the appropriate models. Existing fraud detection systems can be categorized into traditional techniques and modern intelligent techniques. Rule-based systems were among the earliest fraud detection mechanisms. They operate on manually defined rules such as:

- “block transactions above a certain amount”
- “Flag transactions from different countries within a short duration”
- “Trigger alert when high-value purchase is made suddenly”
- “Allow transactions only from registered IP addresses”

Advantages:

- Easy to implement and understand
- Suitable for simple fraud patterns
- Low computational cost

Disadvantages:

- Inability to detect new/unseen fraud patterns
- Requires frequent manual updates
- High false positives
- Poor scalability for large dataset

Manual fraud detection systems In traditional systems, fraud analysts manually review suspicious transactions.

Drawbacks:

- Extremely time-consuming
- Cannot handle millions of transactions
- Analysts can overlook complex fraud patterns
- Expensive and inefficient

Statistical models Earlier research also utilized statistical methods such as:

- Logistic Regression

- Bayesian Classification
- Linear Discriminant Analysis

These models assume that fraudulent behavior follows specific statistical distributions. However, fraud patterns constantly evolve and do not always adhere to fixed probability distributions. Limitations: • Sensitive to outliers • Poor at detecting nonlinear relationships • Require clean and balanced data Because of these limitations, researchers started exploring machine learning approaches.

• Evolution of machine learning in fraud detection With the growth of digital payment systems, researchers began applying machine learning techniques to analyze complex transaction patterns. Supervised machine learning approaches Supervised models require labeled datasets containing both legitimate and fraudulent transactions Common supervised algorithms include: Logistic Regression

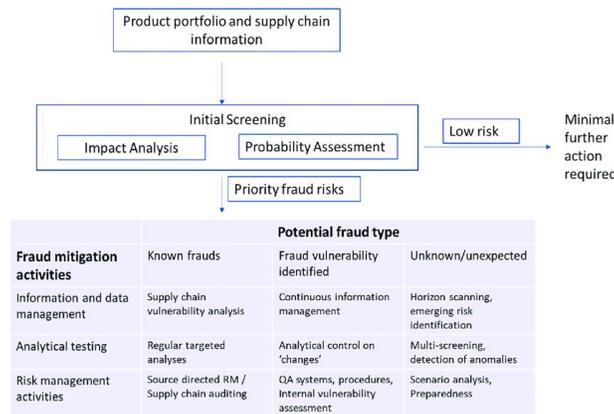
- Random Forest
- Decision Trees
- Gradient Boosting Machines
- Support Vector Machines(SVM)
- Neural Networks

Advantages:

- High accuracy when trained with good data
- Can classify transactions effectively
- Learn nonlinear patterns

Disadvantages:

- Require large labeled fraud datasets
- Cannot detect new types of fraud easily
- Highly imbalanced datasets reduce performance



Unsupervised anomaly detection methods:

These methods identify anomalies without labeled data. Anomalies are rare patterns that differ significantly from normal user behavior.

Popular algorithms include:

- Isolation Forest
- One-Class SVM
- Autoencoders
- Local Outlier Factor(LOF)

Advantage:

- No need for labeled data
- Good at detecting unseen fraud patterns

- Automatically adapts to unusual transactions

Disadvantage:

- Sensitive to dataset noise
- Accuracy depends on feature quality

Deep learning in fraud detection:

With increasing computational capabilities, deep learning- based models are gaining popularity.

Recurrent neural networks(RNNS):

RNNs model sequential transaction behavior of users.

They detect:

- Sudden spikes
- Rapid unusual purchases
- Time-based anomalies

Literature review of previous research studies:

Below are some key research contributions in credit card fraud detection. Study 1: Dal Pozzolo et at.(2015)

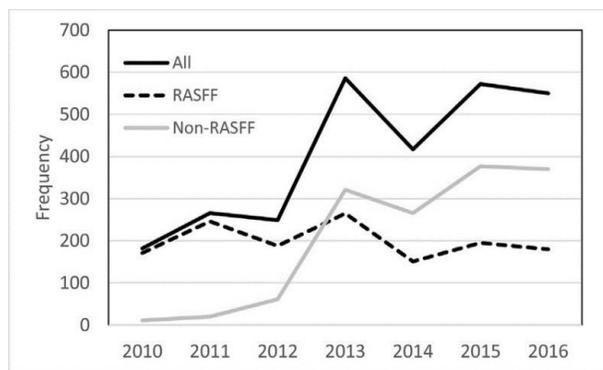
- Used European card transaction dataset.
- Identified severe class imbalance problem.
- Found Random Forest + undersampling gave good results.

Study 2: J.West &M.Bhattacharya (2016)

- Compared neural networks and decision trees.
- Neural networks performed better on complex fraud patterns.

Research gap analysis:

Based on the review of earlier systems, the following gaps are identified: Gap 1: High Data Imbalance Most datasets contain extremely few fraud cases. Gap 2: Evolving Fraud Patterns Fraudsters adapt quickly; rule-base systems cannot catch new patterns. Gap 3: Limited Real-Time capabilities Many models work offline but fail in real-time scoring. Gap 4: Lack of Hybrid Approaches Most studies focus on either supervised or unsupervised but not both.



III. SYSTEM DESIGN

Existing system Traditional fraud detection systems mainly rely on rule-based logic and manual inspection. For example, a transaction may be flagged if its amount exceeds a predefined threshold, originates from an unusual location, or is performed at an unexpected time. These rule-based systems are designed based on past fraud patterns and risk factors defined by fraud analysts.

Disadvantages

- Inability to Detect New Patterns – Fraudsters frequently change their techniques, making fixed rules ineffective.
- High False Positives – Legitimate transactions are often marked suspicious, disrupting customer experience.

- Manual Review Burden – Analysts must verify flagged cases, which is inefficient for large-volume transaction streams.
- Lack of Real-Time Response – Many systems operate in batch mode, delaying fraud detection.
- Scalability Issues – with increasing transactions, traditional systems struggle to maintain accuracy and speed. Because of these drawbacks, existing systems are insufficient for handling modern fraud complexity, which motivates a more intelligent and adaptive solution.

SYSTEM REQUIREMENTS:

Hardware requirements

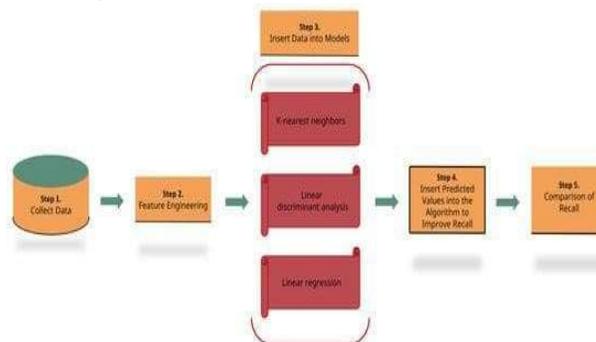
- Processor - Intel Core i5
- Speed - 2.3 GHz
- RAM - 8 GB
- Hard Disk - 500 GB
- Key Board - Standard Windows Keyboard
- Mouse - Three Button Mouse
- Monitor - 24" LED 4.2
- Software requirements
- Operating System : Windows 10
- OR Linux distributions (Ubuntu recommended)
- OR macOS(optional)
- Programming Language : Python

LANGUAGE SPECIFICATION

Python is a high-level, general purpose programming language created by Guido van Rossum and first released in 1991. It emphasizes code readability, minimal syntax, and developer productivity. Python supports multiple programming paradigms including procedural, object-oriented, and functional programming, making it highly suitable for modern machine learning applications. It is widely used in industries such as:

- Banking and Finance
- Artificial Intelligence
- Data analytics
- Web development
- Cybersecurity
- Automation

In the context of fraud detection, Python's capabilities make it an excellent choice for extracting meaningful patterns from large volumes of transaction data.

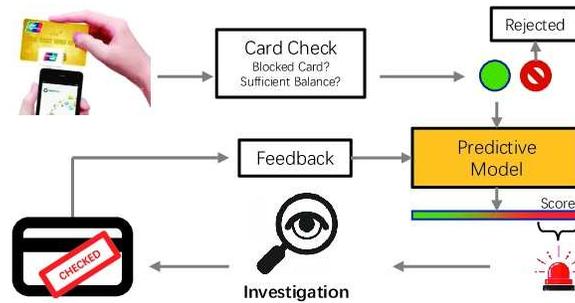


Key features of python Python: certain level of freedom when composing a program, but there are some rules which must always be obeyed. One of these rules, which some people find very surprising, is that python uses indentation (that is, the amount of white space before the statement itself) to indicate the presence of loops, instead of using delimiters like curly braces ({}) or keywords (like —begin and —end) as in many other languages. The amount of indentation you use is not important, but it must be consistent within a given depth of a loop, and statements which are not indented must begin in the first column. Most python programmers prefer to use an editor like emacs, which automatically provides consistent indentation; you will probably find it easier to maintain your programs if you use consistent indentation in every loop, at all depths, and an intelligent editor is very useful in achieving this.

- **Simple and readable syntax:** Python's syntax is clean and easy to understand, which accelerates development and

reduces errors. This is crucial when dealing with large datasets and complex algorithms.

- **Extensive library support:** Python has one of the richest ecosystems of libraries for data processing, visualization, machine learning, and deep learning. Popular libraries such as NumPy, Pandas, Matplotlib, and Scikit-Learn



significantly simplify model implementation.

Object oriented: Python supports OOP concepts such as classes and objects, while also allowing functional programming styles. This flexibility enhances code reusability and organization.

Large community and documentation: Python has one of the largest developer communities in the world. This ensures abundant tutorials, forums, libraries, and debugging support for machine learning projects.

Integration with AI / ML tools Python integrates easily with machine learning toolkits such as:

- TensorFlow
- Scikit-learn
- PyTorch
- Keras
- This makes Python the ideal choice for ML-based fraud detection systems.

Why python was chosen for this project Python was selected as the primary development language for this credit card fraud detection project due to the following reasons: **Ease of Implementation** Python allows quick prototyping of ML models with fewer lines of code. **Rich Data Science Libraries** Libraries like Pandas and NumPy simplify data cleaning and preprocessing. **Strong ML Frameworks**

Provides several features that make it particularly suitable for machine learning and data analysis projects. Some of the important features include: Python provides you with a

Scikit-Learn provides ready-to-use functions for classification, anomaly detection, and model evaluation. **17 Excellent Visualization Support** Libraries like Matplotlib and Seaborn help in generating EDA visualizations essential for Phase -1.

Efficiency with Large Datasets Python handles large datasets well when optimized with vectorized operations. **Community Support** Any error or issue can be quickly addressed due to Python's massive online community. For these reasons, Python stands out as the most preferred language for machine learning research projects. **Python libraries used in the project**

Python's power lies in its libraries. The following libraries were used in Phase-1 of this project: **Numpy (numerical python)** NumPy is a foundational package for numerical computation in Python. It supports:

- Multi-dimensional arrays
- Vectorized operations
- Mathematical and statistical functions

Usage in this project:

- Handling numeric transaction data
- Performing mathematical operations
- Implementing scaling and normalization

Pandas Pandas is a data manipulation and analysis library widely used in data science.

Key features:

- DataFrames for structured datasets
- Handling missing values
- Grouping and sorting operations

Usage in this project:

- Loading the credit card transaction dataset
- Cleaning and preprocessing
- Analyzing fraud vs.non-fraud attributes Critical bugs sometimes do not get fixed for long periods of time. An example is a bug with status critical existing since 2003. **Matplotlib / seaborn** The libraries are widely used for generating plots during Exploratory Data Analysis (EDA). Usage in this project:

- Visualizing amount distribution
- Time distribution patterns
- Correlation heatmaps
- Fraud vs.genuine comparison

Jupyter notebook Jupyter Notebook is an interactive development environment ideal for running machine learning experiments.

Python execution environment During the development, Python code was executed using: Jupyter notebook (for EDA and dataset analysis) Vs code (for python scripts) Anaconda Distribution (for environment management)

Role of python in credit card fraud detection Python played a critical role in executing the following tasks: **Handling Imbalanced Datasets** Using P+ scikit-learn preprocessing tools. **Identifying Baseline ML approaches** Python enabled rapid testing and comparison of:

- Logistic Regression
- Random Forest
- Isolation Forest
- One – class svmObject oriented: Python supports OOP concepts such as classes and objects, while also allowing functional programming styles. This flexibility enhances code reusability and organization.

Large community and documentation: Python has one of the largest developer communities in the world. This ensures abundant tutorials, forums, libraries, and debugging support for machine learning projects. Integration with AI / ML tools Python integrates easily with machine learning toolkits such as:

- TensorFlow
- Scikit-learn
- PyTorch
- Keras

• This makes Python the ideal choice for ML-based fraud detection systems. Why python was chosen for this project Python was selected as the primary development language for this credit card fraud detection project due to the following reasons: Ease of Implementation Python allows quick prototyping of ML models with fewer lines of code. Rich Data Science Libraries Libraries like Pandas and NumPy simplify data cleaning and preprocessing. Strong ML Frameworks Scikit-Learn provides ready-to-use funtions for classificationsanomaly detection, and model evaluation. 17 Excellent Visualization Support Libraries like Matplotlib and Seaborn help in generating EDA visualizations essential for Phase -1. Efficiency with Large Datasets Python handles large datasets well when optimized with vectorized operations. Community Support Any error or issue can be quickly addressed due to Python’s massive online community. For these reasons, Python stands out as the most preferred language for machine learning research projects. Python libraries used in the project Python’s power lies in its libraries. The following libraries were used in Phase-1 of this project: Numpy (numerical python) NumPy is a foundational package for numerical computation in Python.

It supports:

- Multi-dimensional arrays
- Vectorized operations
- Mathematical and statistical functions

Usage in this project:

- Handling numeric transaction data
- Performing mathematical operations
- Implementing scaling and normalization

Pandas Pandas is a data manipulation and analysis library widely used in data science.

Key features:

- DataFrames for structured datasets
- Handling missing values
- Grouping and sorting operations

Usage in this project:

- Loading the credit card transaction dataset
- Cleaning and preprocessing
- Analyzing fraud vs.non-fraud attributes Critical bugs sometimes do not get fixed for long periods of time. An example is a bug with status critical existing since 2003. **Matplotlib / seaborn** The libraries are widely used for generating plots during Exploratory Data Analisis (EDA). Usage in this project:

- Visualizing amount distribution
- Time distribution patterns
- Correlation heatmaps
- Fraud vs.genuine comparison

Jupyter notebook Jupyter Notebook is an interactive development environment ideal for running machine learning experiments.

Python execution environment During the development, Python code was executed using: Jupyter notebook (for EDA and dataset analysis) Vs code (for python scripts) Anaconda Distribution (for environment management)

Role of python in credit card fraud detection Python played a critical role in executing the following tasks: **Handling Imbalanced Datasets** Using P+ scikit-learn preprocessing tools. **Identifying Baseline ML approaches** Python enabled rapid testing and comparison of:

- Logistic Regression
- Random Forest
- Isolation Forest
- One – class svm.

IV. EXISTING SYSTEM

Traditional fraud detection systems mainly rely on rule-based logic and manual inspection. For example, a transaction may be flagged if its amount exceeds a predefined threshold, originates from an unusual location, or is performed at an unexpected time. These rule-based systems are designed based on past fraud patterns and risk factors defined by fraud analysts.

Disadvantages:

- Inability to Detect New Patterns – Fraudsters frequently change their techniques, making fixed rules ineffective.
- High False Positives – Legitimate transactions are often marked suspicious, disrupting customer experience.
- Manual Review Burden – Analysts must verify flagged cases, which is inefficient for large-volume transaction streams.
- Lack of Real-Time Response – Many systems operate in batch mode, delaying fraud detection.
- Scalability Issues – with increasing transactions, traditional systems struggle to maintain accuracy and speed.

Because of these drawbacks, existing systems are insufficient for handling modern fraud complexity, which motivates a more intelligent and adaptive solution.

V. METHODOLOGY

We propose credit card fraud detection system using machine learning-based anomaly detection, this chapter presents the step-by step methodology followed for the project, including dataset collection, preprocessing, exploratory data analysis (EDA), feature engineering, and baseline model selection strategy. Dataset collection: The dataset used for this project is the kaggle credit card fraud detection dataset. Which contains anonymized credit card transactions made by European customers. The dataset is ideal for anomaly detection because it represents a real-world imbalanced dataset where fraudulent cases are extremely rare. Data preprocessing methodology: Data preprocessing ensures that the raw dataset is transformed into a meaningful and usable format for analysis.

The following preprocessing tasks were completed:

- Handling Missing Values
- Duplicate and Noise Removal
- Feature Scaling
- Time Feature Ana..

VI. CONCLUSION

The primary objective of this project was to study and analyze the feasibility of developing a machine learning-based anomaly detection system for identifying fraudulent credit card transactions. Throughout this phase, significant groundwork was completed in understanding the nature of fraud detection, examining existing systems, analyzing the dataset, identifying key features ,and outlining the appropriate machine learning techniques suitable for the project. The study revealed that traditional fraud detection approaches, especially rule-based and manual inspection systems, are insufficient for handling modern fraud patterns due to their inability to adapt to evolving fraud strategies. The literature survey and system analysis clearly demonstrated the importance of machine learning in detecting anomalies within massive transactional datasets. Machine learning-based detection offers improved adaptability,scalability ,and accuracy in recognizing unusual transaction behavior, making it an ideal approach for combating financial fraud. The outcomes from this phase ensure that the project is well- prepared for the next stage, where actual model training, evaluation, and tuning will be performed.

- Time Feature Ana..

VII. CONCLUSION

The primary objective of this project was to study and analyze the feasibility of developing a machine learning-based anomaly detection system for identifying fraudulent credit card transactions. Throughout this phase, significant groundwork was

completed in understanding the nature of fraud detection, examining existing systems, analyzing the dataset, identifying key features ,and outlining the appropriate machine learning techniques suitable for the project. The study revealed that traditional fraud detection approaches, especially rule-based and manual inspection systems, are insufficient for handling modern fraud patterns due to their inability to adapt to evolving fraud strategies. The literature survey and system analysis clearly demonstrated the importance of machine learning in detecting anomalies within massive transactional datasets. Machine learning-based detection offers improved adaptability,scalability ,and accuracy in recognizing unusual transaction behavior, making it an ideal approach for combating financial fraud. The outcomes from this phase ensure that the project is well- prepared for the next stage, where actual model training, evaluation, and tuning will be performed.

- Time Feature Ana..

VIII. CONCLUSION

The primary objective of this project was to study and analyze the feasibility of developing a machine learning-based anomaly detection system for identifying fraudulent credit card transactions. Throughout this phase, significant groundwork was completed in understanding the nature of fraud detection, examining existing systems, analyzing the dataset, identifying key features ,and outlining the appropriate machine learning techniques suitable for the project. The study revealed that traditional fraud detection approaches, especially rule-based and manual inspection systems, are insufficient for handling modern fraud patterns due to their inability to adapt to evolving fraud strategies. The literature survey and system analysis clearly demonstrated the importance of machine learning in detecting anomalies within massive transactional datasets. Machine learning-based detection offers improved adaptability,scalability ,and accuracy in recognizing unusual transaction behavior, making it an ideal approach for combating financial fraud. The outcomes from this phase ensure that the project is well- prepared for the next stage, where actual model training, evaluation, and tuning will be performed.

REFERENCES

- [1] Bhattacharya,S., Jha,S., Data Mining Approaches for Banking Fraud Detection ,2014.
- [2] Dal Pozzolo,A., Adaptive Machine Learning Techniques for Credit Card Fraud Detection, 2015.
- [3] West,J., Bhattacharya,M., Intelligent Fraud Detection Techniques for Financial,2016
- [4] Carcilo,F., Real-Time Credit Card Fraud Detection Using Machine Learning, 2018.
- [5] Patidar, R.,Sharma,L., Credit Card Fraud Detection Using Neural Networks,2011.
- [6] Sebastio,H., Pires, F., Classifier Ensemble Approaches for Credit Card Fraud Detection, 2017.
- [7] Abdelrahim, A., Machine Learning Methods for Imbalanced Fraud Detection Datasets, 2019.
- [8] Ahmed, M., Mahmood, A., Unsupervised Anomaly Detection for Fraudulent Transactions, 2016.
- [9] Whitrow c., Hand, D., Transaction Aggregation for Fraud Detection,2009.
- [10] Randhawa, k. Credit Card Fraud Detection Using Majority voting Ensembles, 2018.

Digital Therapeutic Mobile Application For Supporting Children With ADHD

*Note: Sub-titles are not captured in Xplore and should not be used

¹ Lismariya Shaji

Department of Computer Science
and Engineering,
Vimal Jyothi Engineering College
Chemperi, Kannur, Kerala, India
lismariya16@gmail.com

² Melbin Dominic Jose

Department of Computer Science
and Engineering
Vimal Jyothi Engineering College
Chemperi, Kannur, Kerala, India
melbindominicjose1.1@gmail.com

³ Pavithra Sony

Department of Computer Science
and Engineering
Vimal Jyothi Engineering College
Chemperi, Kannur, Kerala, India
pavithrasony25@gmail.com

⁴ Alphonsa Prince

Department of Computer Science and
Engineering,
Vimal Jyothi Engineering College
Chemperi, Kannur, Kerala, India
alphonsaprince666@gmail.com

⁵ Sreedaya M

Department of Computer Science and
Engineering,
Vimal Jyothi Engineering College
Chemperi, Kannur, Kerala, India
sreedayam@vjec.ac.in

Abstract—Attention Deficit Hyperactivity Disorder (ADHD) often makes it difficult for children to stay focused, manage impulses, and keep up with everyday academic and routine tasks. Because these challenges tend to show up across different parts of a child's day, they benefit from steady, structured support rather than occasional interventions. The work described here focuses on developing a mobile therapeutic application that brings several helpful tools together instead of treating them as separate features. The system includes a Pomodoro-based focus timer that uses face detection to make sure the child stays engaged, along with short cognitive games designed to strengthen attention skills. A simple guided meditation section is also provided to support emotional balance. All of this is connected to a parent dashboard that shows study duration, meditation activity, and game usage, giving caregivers a clearer idea of the child's habits.

Initial testing showed that the modules run smoothly and that the app can consistently log different types of activity. While these early results are limited, they suggest that the application may help children build steadier focus routines, improve emotional regulation over time, and allow caregivers to stay more involved in a practical way. Overall, the project aims to offer a child-friendly, accessible digital tool that supports both behavioural growth and cognitive development in children with ADHD.

Index terms — ADHD, digital therapeutics, mobile application, gamification, focus training, face detection.

I. INTRODUCTION

Attention Deficit Hyperactivity Disorder (ADHD) is a common neurodevelopmental condition seen in many children, often making it difficult for them to stay focused, control impulses, manage emotions, and maintain steady study habits. These difficulties can affect both academic performance and everyday routines, which is why children usually need consistent support and structured practice even outside clinical settings. While medication and behavioural therapy are the usual forms of treatment, many families struggle to find tools that help children stay engaged and build healthy routines at home.

Although there are several digital apps available today, most of them offer only a single feature—like a simple game, a timer, or a relaxation activity—without bringing these elements together in a meaningful way. Many also lack ways to track how regularly a child uses the app or how their behaviour changes over time. Because of this, children often miss out on the well-rounded digital support they need to steadily improve attention, emotional regulation, and self-discipline.

To bridge these gaps, this work introduces a mobile digital therapeutic application that brings multiple helpful tools into one platform. The app includes a Pomodoro-style focus timer that uses real-time face detection to encourage consistent study habits, along with short cognitive games designed to strengthen attention. It also provides guided meditation sessions to help children calm down and manage their emotions more effectively. For caregivers, a dedicated dashboard gives insights into study duration, meditation usage, and behavioural trends, helping parents stay more involved and supportive.

By combining these features into a single system, the goal is to offer children an accessible and engaging tool that supports

better focus, emotional stability, and consistent routines—while also giving caregivers a clearer understanding of their child’s daily progress in managing ADHD.

I. RELATED WORK

Several digital therapeutic tools and gamified interventions have been proposed to support children with ADHD and related cognitive challenges. Luo et al. introduced a mobile game prototype designed to enhance attention, working memory, and inhibitory control through structured mini-games and adaptive difficulty levels [1]. While the system demonstrated improvements in cognitive outcomes, it primarily focused on gameplay and did not integrate behavioural monitoring or caregiver involvement.

Kim and Lee developed NeuroWorld DTx, an AI-driven digital therapeutic program that uses structured cognitive games combined with adaptive difficulty and daily session limits to improve attention and working memory [2]. Although effective in short-term cognitive enhancement, the system did not include real-time engagement tracking, such as face detection or tools for parental supervision.

Other studies have examined personalised game-based interventions that adjust complexity based on child performance, showing strong correlations between adaptive gameplay and improved attention and memory [3]. However, these systems often lack mindfulness or emotional regulation components, which are crucial for children with ADHD.

Serious games like the BRAVO Project have shown long-term benefits in attention, planning, and self-control through structured mini-games and adaptive feedback [6]. Similarly, mobile literacy and mindfulness applications designed for children with special needs demonstrate the importance of accessibility, guided sessions, and caregiver involvement [7], [8]. Yet these systems typically focus on a single domain, such as games or mindfulness, without offering an integrated multi-module framework.

In contrast to these existing solutions, the proposed application combines cognitive games, Pomodoro-based focus training with real-time face detection, guided meditation, and a parent dashboard into a unified mobile therapeutic system. This holistic approach addresses attention, emotional regulation, engagement monitoring, and caregiver participation within a single platform, providing broader and more practical support for children with ADHD.

I. SYSTEM ARCHITECTURE

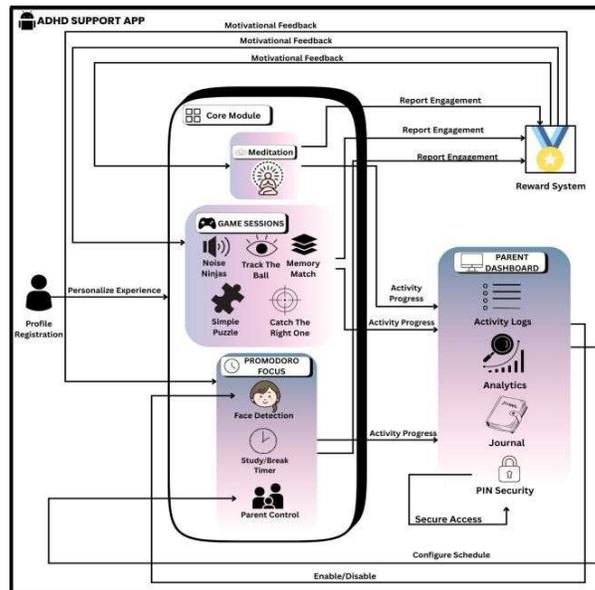


Figure 1: Architecture Diagram.

The system is built using a modular client–server architecture intended to support the different therapeutic features offered to children with ADHD. As illustrated in Fig. 1, the mobile application hosts four key modules: Gamified Learning, a Focus Timer equipped with face-detection monitoring, Guided Meditation, and the Parent Dashboard. Although each module functions independently on the device, they all share

a common communication layer that interacts with the backend service.

The Gamified Learning component presents simple cognitive training tasks and records gameplay metrics, which are periodically synced with the server. The Focus Timer uses a Pomodoro-style routine and integrates a lightweight face-detection mechanism to check whether the child is present during the session. Events such as session start, pause, face-loss alerts, and completion are logged. The Meditation module guides children through relaxation exercises and stores session length and completion details. Meanwhile, the Parent Dashboard pulls summarised data—such as focus duration, meditation activity, and game interactions—through secure API calls so caregivers can observe daily patterns and behavioural progress.

Communication between the client and the backend is bidirectional. The mobile device sends session logs and behavioural metrics, while the server processes this information and returns insights, progress summaries, and dashboard-ready data. Only pseudonymized or aggregated information is provided to the parent interface to maintain privacy and ensure responsible handling of child-related activity data. Overall, this architecture supports reliable tracking, future modular expansion, and smooth synchronisation across all parts of the therapeutic system.

II. SYSTEM MODULES

The system is built around four main modules, each focusing on a different area that children with ADHD often need help with—attention, memory, emotional control, and regular monitoring from parents. All of these modules run inside the mobile app, but they share the same communication layer, so the app can sync data with the backend whenever needed without the user noticing anything in the background.

1) Games Module: The Games Module includes five small, easy-to-understand games. They are meant to be fun but still help with important cognitive skills. None of the games is too long or complicated, since the idea is to keep the child engaged without overwhelming them. The app also records simple performance details like accuracy, the time they take to finish, and how many mistakes they make, so progress can be seen over days or weeks.

2) Noise Ninjas: In this game, the child listens for one specific sound while other background noises play. It basically helps them practice filtering out distractions—something many children with ADHD find hard.

3) Memory Match: A classic matching game where they flip cards to find pairs. It's mostly about improving visual memory and staying focused long enough to remember where things are.

4) Track the Ball: Several identical balls move around, and the child has to follow one of them until it stops. This supports focus and visual tracking, which are both important for everyday tasks.

5) Simple Puzzle: The child completes pictures or arranges shapes in the right spots. It helps with planning, problem-solving, and understanding how pieces fit together.

Catch the Right One: Different objects drop from the top of the screen, and the child taps only the correct objects while avoiding the “wrong” one. This helps with impulse control and making quick but accurate decisions. Taken together, these games form a simple but effective set of activities that help with attention, memory, and general cognitive skills—without making the child feel like they're doing “work.”

A. Focus Timer With Face Detection

This module uses a Pomodoro-style timer with focus periods and break times. The unique part is that it uses face detection during the focus phase. If the child leaves the screen or gets distracted for too long, the system notes it as an interruption. At the end of the session, the app records things like how long the child actually focused and how many times they drifted away. Over time, this helps them build stronger study habits and stay more consistent.

B. Meditation and Breathing Module

This module provides simple meditation and breathing exercises that children can follow easily. The sessions are short, calming, and use child-friendly audio. The app logs how long each session lasts so parents can see if the child is regularly practising. These activities can help reduce restlessness and give the child a short break to calm down or reset when they feel overwhelmed.

C. Parent Dashboard

The Parent Dashboard is meant to give caregivers a clear overview of what the child is doing inside the app. It brings together data from the games, the focus timer, and the meditation sessions. Parents can see charts showing study time, how often the child completed Pomodoro cycles, how they performed in the games, and how regularly they used the meditation module. This kind of information helps parents understand what's improving and what may need more attention. It also encourages more consistent involvement at home, which is important for long-term progress.

III. IMPLEMENTATION

The therapeutic app is put together as a simple, modular mobile system so that each part can work smoothly without putting too much load on the phone. It's built using a cross-platform framework, mainly to make sure it runs reliably on Android devices. Each feature—like the Games, Focus Timer, Meditation section, and the Parent Dashboard—has been developed as its own separate piece inside the app, which makes it easier to manage and update over time.

For communication with the server, the app uses RESTful APIs. These handle the usual things, such as logging in, saving session details, syncing game-related data, and fetching progress updates. The backend stores information about focus sessions, meditation activities, and game results. Over time, this helps build a useful history that parents can view in the dashboard. To keep everything secure, the system uses authentication tokens so that only the right users can access specific data. One important point is that the face-detection feature for the Focus Timer works completely on the device; nothing sensitive is uploaded. Only basic events, like when a face is detected or when a session ends, are sent to the server.

The interface is kept simple and easy to follow. Navigation is straightforward, colours are soft and child-friendly, and the overall

layout avoids unnecessary clutter. The Games Module uses only the most basic gestures so children can interact without difficulty. The Focus Timer screen shows clear start and stop buttons along with a visual progress bar. The Meditation section guides the child step-by-step through breathing or mindfulness routines. For parents, the dashboard summarises patterns and improvements using light, fast-loading charts that work well on any phone.

I. RESULTS/PLANNED EVALUATION

As full testing with users hasn't been carried out yet, this section explains how the app will be evaluated once the study begins. The plan is to look at how engaged the children are, whether the app helps them improve their attention, and how easy it is for both children and caregivers to use. To understand these areas, data will be taken from the app's activity logs, simple behavioural observations, and feedback from caregivers who interact with the dashboard.

The initial study will include a small set of children aged 7 to 12 who have already been diagnosed with ADHD. For about two to four weeks, they will use the app as part of their everyday routine—doing focus sessions, playing the cognitive games, and completing short meditation exercises. Throughout this period, caregivers will track progress in the Parent Dashboard and share their thoughts on what worked well and what could be improved.

To measure attention, the Focus Timer will provide details such as how many Pomodoro cycles the child completes, how long they stay focused without interruption, and how often the face-detection system notices breaks in attention. The Games Module will offer cognitive data, including accuracy, how quickly the child responds, and how far they progress through the five games. The Meditation section will help reflect emotional regulation based on how consistently it is used and the length of each session.

Caregivers will also complete a usability questionnaire designed for child-focused digital tools. This will cover how easy the app is to move around in, whether the instructions are simple to follow, if the visuals feel comfortable for children, and their overall impression of the experience. Alongside this, caregivers will provide short written observations, especially about any changes they notice—like better study habits, fewer distractions, or improved willingness to follow routines.

The broader goal of this evaluation is to understand whether combining focus training, cognitive games, meditation, and regular caregiver involvement can genuinely help children manage their attention better, stay calmer, and develop steadier daily habits. The findings will guide the next stage of improvements and help determine how useful the app may be as a supportive digital

Category	Metric	Description
Attention	Completed focus cycles	Total number of Pomodoro sessions completed.
	Average uninterrupted duration	Mean continuous focus time per cycle.
	Face-loss events	Number of interruptions detected by the camera.
Cognitive Performance	Game accuracy	Percentage of correct responses in each game.
	Daily task-completion rate	Percentage of assigned cognitive tasks completed each day.
	Level progression	Number of levels completed in each game.
Emotional Regulation	Meditation frequency	Number of guided sessions completed per week.
	Session duration	Average meditation session length.
Usability	System Usability Score (SUS)	Caregiver-rated usability on a 0–100 scale.
	Qualitative caregiver feedback	Observed behavioural and engagement changes.

Figure 2: Planned Evaluation Metrics for the ADHD Therapeutic Application

I. DISCUSSION

This digital therapeutic app brings together focus training, cognitive games, guided meditation, and caregiver monitoring in one place, with the aim of supporting children with ADHD in their daily routines. From the evaluation plan, it seems that the system can collect useful information about how well the child is paying attention, how they perform on simple cognitive tasks, and how consistently they engage with calming or mindfulness activities. The combination of face-detection cues and Pomodoro-style focus sessions gives the child gentle structure and reminders that help them stay on track and build better study habits over time.

One of the main advantages of this system is its multi-component approach. Instead of focusing on only one skill, the app tries to address different areas that children with ADHD often struggle with. The cognitive games are designed to strengthen executive-function abilities, while the meditation activities give children a simple way to calm down or reset when they feel restless. On top of that, the Parent Dashboard gives caregivers a clearer idea of how their child is progressing, making it easier for them to support routines at home and encourage positive behaviour.

There are, however, a few limitations to acknowledge. The face-detection system can confirm that the child is physically present, but it can't judge how mentally engaged they are. The benefits of the cognitive games also depend on how regularly the child uses them, and this can vary depending on interest levels, schedules, or how much support they get from parents. Since testing with real users hasn't been carried out yet, all expected outcomes remain preliminary until the system is tested in a controlled setting.

Despite these constraints, the application demonstrates strong potential as a supportive tool for children with ADHD. Its

modular design ensures that new capabilities can be added in the future, such as therapist dashboards, expanded reporting features, or additional therapeutic modules. With continued refinement and structured evaluation, the system could become a practical and accessible digital therapeutic solution that complements traditional behavioural approaches.

II. CONCLUSION

This work presented a mobile-based digital therapeutic application designed to support children with ADHD through an integrated set of modules combining cognitive games, focus training with face detection, guided meditation, and caregiver monitoring. The system aims to improve attention, emotional regulation, and routine-building by offering structured and child-friendly activities that can be incorporated into daily home environments. The modular design allows consistent tracking of behavioural patterns and facilitates meaningful caregiver involvement through the Parent Dashboard. Although real-world evaluation has not yet been completed, the proposed metrics indicate the system's potential to deliver measurable improvements in focus, cognitive performance, and emotional stability. Current limitations include the need for adaptive personalisation, deeper behavioural engagement analysis, and validation through testing with clinical populations. Future work will involve conducting controlled user studies, refining therapeutic content, and exploring AI-driven personalisation to enhance long-term effectiveness. Overall, the application represents a promising and accessible digital tool for supplementing behavioural interventions for children with ADHD.

ACKNOWLEDGMENT

The authors express their heartfelt gratitude to Ms. Sreedaya M, Assistant Professor, Department of Computer Science and Engineering, for her continuous guidance and valuable suggestions throughout the project. Her support and feedback were instrumental in shaping this work. Sincere thanks are also extended to Ms. Divya B, Head of the Department of Computer Science and Engineering, for her encouragement and academic support. The Department of Computer Science and Engineering, Vimal Jyothi Engineering College, is gratefully acknowledged for providing the facilities and resources needed to carry out this project. Finally, the authors express their deep appreciation to their families for their patience, motivation, and unwavering support during the development of this work.

REFERENCES

- [1] J. Luo, F. Li, Y. Wu, X. Liu, Q. Zheng, Y. Qi, H. Huang, G. Xu, Z. Liu, F. He, and Y. Zheng, "A mobile device-based game prototype for ADHD: development and preliminary feasibility testing," *Translational Psychiatry*, vol. 14, no. 1, pp. 251, 2024, doi: 10.1038/s41398-024-02964-2.
- [2] S.-C. Kim and H. Lee, "An Analysis of the Impact of Digital Therapeutic Interventions on Attention and Working Memory in Children with Attention-Deficit/Hyperactivity Disorder: A Randomized Controlled Trial," *Applied Sciences*, vol. 15, no. 2, pp. 788, 2025, doi: 10.3390/app15020788.
- [3] S.-C. Kim, J.-H. Song, and N.-Y. Kong, "Personalized Game-Based Content and Performance: A Pilot Study on a Digital Intervention for Children with ADHD," *Bioengineering*, vol. 11, no. 12, pp. 1277, 2024, doi: 10.3390/bioengineering11121277.
- [4] K. Kanokjun and P. Homchampa, "Understanding Attention Deficit Hyperactivity Disorder (ADHD) in Children: A Qualitative Analysis of Family Responses and Behavioral Challenges," *The Open Public Health Journal*, vol. 17, 2024, doi: 10.2174/0118749445293643240228120415.
- [5] R. Drechsler, S. Brem, D. Brandeis, E. Gru'nblatt, G. Berger, and S. Walitza, "ADHD: Current Concepts and Treatments in Children and Adolescents," *Neuropediatrics*, vol. 51, 2020, doi: 10.1055/s-0040-1701658.
- [6] V. De Luca, A. Schena, A. Covino, P. Di Bitonto, A. Potenza, M.C. Barba, G. D'Errico, and L. T. De Paolis, "Serious Games for the Treatment of Children with ADHD: The BRAVO Project," *Information Systems Frontiers*, vol. 27, no. 3, pp. 841–863, 2025, doi: 10.1007/s10796-023-10457-8.
- [7] A. B. Ozbek and M. Torppa, "Developing an early literacy mobile app: design process and testing the app with children with special needs," *Education and Information Technologies*, vol. 30, no. 14, pp. 20573–20596, 2025, doi: 10.1007/s10639-025-13599-1.
- [8] E. Larson, R. L. Mattie, and S. A. Riffkin, "Assessment of Acceptability, Usage, and Impact on Caregivers of Children With Autism's Stress and Mindfulness: Multiple-Method Feasibility Study of the 5Minutes4Myself App's Mindfulness Module," *JMIR Hum Factors*, vol. 11, pp. e54171, 2024, doi: 10.2196/54171.
- [9] M. Chistol, C. Turcu, and M. Danubianu, "Autism Assistant: A Platform for Autism Home-Based Therapeutic Intervention," *IEEE Access*, vol. 11, pp. 94188–94204, 2023, doi: 10.1109/ACCESS.2023.3310397. M. Al Khaldy, A. Shaheen, W. Alzyadat, and A. Alhroob, "Development and Evaluation of a Mobile Application for Enhancing the Academic and Social Behavior of Children with Attention-Deficit/Hyperactivity Disorder (ADHD)," *International Journal of Engineering Trends and Technology*, vol. 73, no. 5, pp. 92–102, 2025, doi: 10.14445/22315381/IJETT-V73I5P109.
- [10] C. Seery, R. H. Cochrane, M. Mulcahy, K. Kilbride, M. Wrigley, and J. Bramham, "A one-stop shop: Real-world use and app-users' experiences of a psychoeducational smartphone app for adults with ADHD," *Internet Interventions*, vol. 39, pp. 100807, 2025, doi: 10.1016/j.invent.2025.100807.

A REVIEW: MACHINE LEARNING-BASED MODEL FOR CAMPUS PLACEMENT PREDICTION

*Note: Sub-titles are not captured in Xplore and should not be used

¹KESAVAN.M

Student,
II -M.B. A Department of Management
Studies,
Vel Tech Rangarajan Dr. Sagunthala
R&D Institute of Science and Technology,
Avadi
E-mail: vtp4237@veltech.edu.in
Contact: 9445562219

²DR. M.RAJAPRIYA

Assistant Professor,
Department of Management
Studies,
Vel Tech Rangarajan Dr. Sagunthala
R&D Institute of Science and Technology,
Avadi
E-mail: drrajapriyam@veltech.edu.in

Abstract—In today’s dynamic and sustainability-driven business environment, organizations are increasingly recognizing the importance of integrating technology with responsible marketing practices. Artificial Intelligence (AI) has emerged as a transformative force that enhances sustainable marketing by enabling data-driven decision-making, personalized communication, and efficient resource management. This study titled “A Study on Artificial Intelligence in Sustainable Marketing Strategies for Business Growth” explores how AI technologies can be leveraged to promote eco-friendly marketing practices while driving profitability and competitive advantage.

The research focuses on understanding how AI tools such as predictive analytics, machine learning algorithms, chatbots, and automated marketing systems can help businesses implement sustainable strategies effectively. These technologies enable organizations to analyze consumer behavior, design personalized campaigns, optimize supply chains, and minimize environmental impact. By aligning AI-driven innovations with sustainability objectives, businesses can enhance brand reputation, improve customer loyalty, and achieve long-term growth.

A mixed-method research design will be adopted, combining qualitative insights from industry experts with quantitative analysis of marketing data. The study will examine the impact of AI-based sustainable marketing on factors such as customer engagement, brand trust, and operational efficiency. Additionally, it will address challenges including ethical AI use, data privacy, digital divide, and organizational readiness for AI adoption.

The findings are expected to demonstrate that Artificial Intelligence serves as a strategic enabler for sustainability by promoting transparency, efficiency, and innovation in marketing practices. Ultimately, this research highlights that AI-driven sustainable marketing not only contributes to environmental and social well-being but also acts as a key driver of long-term business growth in the modern digital economy.

Keywords— *Green Marketing, Digital Transformation, Data-Driven Marketing, Predictive Analytics, Consumer Behaviour, Eco-Innovation, Corporate Sustainability, Smart Marketing Strategies, Machine Learning Applications, Environmental Responsibility, AI-Powered Decision Making, Ethical Marketing Practices.*

I. INTRODUCTION

In recent years, the rapid advancement of Artificial Intelligence (AI) has significantly transformed the global business landscape, influencing decision-making, customer engagement, and operational efficiency across industries. At the same time, sustainability has emerged as a critical priority for organizations aiming to balance economic growth with environmental and social responsibility. As consumer expectations shift toward eco-friendly products and ethical business practices, companies are increasingly integrating sustainable marketing strategies to strengthen their brand value and long-term competitiveness.

AI technologies such as machine learning, predictive analytics, natural language processing, and intelligent automation provide powerful tools for enhancing the effectiveness of sustainable marketing. These technologies enable businesses to better understand consumer behavior, optimize resource utilization, reduce environmental impact, and develop personalized, sustainability-focused communication strategies. By leveraging AI-driven insights, organizations can design marketing

practices that not only promote green products and responsible consumption but also contribute to overall business growth.

This study explores the integration of artificial intelligence in sustainable marketing strategies and examines how AI-driven innovations support organizations in achieving both sustainability and growth objectives. It aims to analyze the extent to which AI facilitates eco-friendly decision-making, improves marketing performance, enhances customer satisfaction, and contributes to long-term value creation. The findings of this research will provide valuable insights for businesses, policymakers, and researchers seeking to adopt AI-enabled sustainable marketing approaches for a more responsible and competitive future.

II. OBJECTIVES OF THE RESEARCH

a) *Primary Objective*

- To analyse how Artificial Intelligence (AI) enhances sustainable marketing strategies and contributes to overall business growth.

b) *Secondary Objectives*

- To examine the influence of AI-driven tools on consumer awareness and engagement with sustainable products.
- To assess the role of AI in improving resource efficiency and reducing environmental impact within marketing activities.
- To evaluate the effectiveness of AI in personalizing sustainable marketing messages to different consumer segments.
- To identify the challenges and limitations faced by organizations in integrating AI into sustainable marketing practices.
- To provide strategic recommendations for implementing AI-based sustainable marketing approaches for long-term organizational development.

III. REVIEW OF THE LITERATURE

Research on the integration of Artificial Intelligence (AI) into sustainable marketing has expanded significantly in recent years. Early studies such as Kumar et al. (2019) highlight that AI-driven analytics help companies understand consumer preferences for eco-friendly products, enabling more targeted and efficient sustainable marketing campaigns. Their findings show that AI-supported segmentation and personalization can increase consumer engagement with green products.

Several researchers have noted the operational sustainability benefits brought by AI. According to Wamba et al. (2020), machine learning models improve demand forecasting accuracy, which reduces overproduction and minimizes resource waste. Similarly, Ghosh & Bhattacharya (2021) point out that AI-enabled supply chain optimization reduces carbon emissions by improving logistics routes and inventory management. These studies emphasize that AI supports both environmental sustainability and operational efficiency.

AI has also been found to influence consumer behavior toward sustainability. Research by Del Río-González (2020) shows that recommendation systems and personalized green messages can nudge consumers toward environmentally responsible choices. This supports the view that AI can make sustainability communication more persuasive and impactful.

However, scholars have also highlighted challenges associated with AI integration. According to Dwivedi et al. (2021), companies adopting AI in sustainable marketing face issues related to data privacy, algorithmic bias, and transparency, which can affect consumer trust. Additionally, studies by Pizzi et al. (2020) indicate that excessive automation in sustainability communication may lead consumers to perceive marketing messages as less authentic, increasing the risk of greenwashing concerns.

From a strategic perspective, AI has been shown to support long-term business growth. Research by Mariani & Borghi (2022) demonstrates that firms using AI in sustainability-driven marketing tend to achieve higher brand loyalty and improved financial performance due to enhanced customer experiences and operational efficiencies. Yet, scholars such as Gupta & Singh (2021) argue that without proper organizational readiness, technological capabilities, and employee training, firms struggle to fully capture the benefits of AI-based sustainability initiatives.

Overall, the literature suggests that AI plays a significant role in strengthening sustainable marketing strategies through improved efficiency, targeted communication, and enhanced consumer engagement. At the same time, ethical concerns, implementation challenges, and issues of transparency remain critical areas that require managerial attention.

Scope of the Study

- To examine how AI technologies support sustainable marketing practices such as green promotion, eco-friendly product positioning, and responsible communication.
- To analyze the role of AI in understanding consumer behavior toward sustainable products through data analytics and predictive tools.
- To study how AI improves operational efficiency by reducing waste, optimizing resources, and supporting environmentally friendly processes.
- To evaluate the impact of AI-enabled sustainable marketing on business growth, brand image, and customer loyalty.
- To identify challenges faced by organizations while implementing AI in sustainable marketing initiatives.
- To focus on marketing-related sustainability practices only, excluding technical or engineering aspects of AI.
- To cover selected businesses and respondents (such as marketing managers, executives, and consumers) who are aware of or involved in AI-based marketing.
- To review both national and international literature related to AI, sustainable marketing, and business growth.

IV. RESEARCH METHODOLOGY

1. Research Design

The study adopts a descriptive research design to understand how Artificial Intelligence (AI) is integrated into sustainable marketing strategies and how it contributes to business growth. This design helps in describing the current practices, perceptions, and outcomes related to AI-driven sustainable marketing.

2. Nature of the Study

The research follows a quantitative approach supported by qualitative insights. Quantitative data will be collected through structured questionnaires, while qualitative understanding may be derived from secondary sources such as journals, reports, and previous studies.

3. Sources of Data

a. Primary Data

- Collected through questionnaires distributed to marketing managers, business professionals, and consumers familiar with AI-based sustainable marketing.
- Responses will help understand real-time practices, challenges, and business outcomes.

b. Secondary Data

- Obtained from research articles, journals, books, company reports, websites, and industry publications related to AI, sustainable marketing, and business growth.

4. Sampling Method

- Sampling Technique: Convenience sampling or purposive sampling (depending on availability of respondents with relevant knowledge).
- Sample Size: Can range between 100–150 respondents or as required for statistical accuracy.

5. Data Collection Tools

- Structured Questionnaire
 - Includes closed-ended questions, Likert-scale items, and demographic details.
- Secondary Data Review
 - Literature review from credible academic and industry sources.

6. Data Analysis Techniques

- Data will be analyzed using descriptive statistics such as:
 - Percentages
 - Mean and standard deviation
 - Frequency distribution

- If needed, inferential statistics such as:
 - o Correlation
 - o Chi-square test
 - o Regression analysis

may be applied to examine relationships between variables like AI usage, sustainability outcomes, and business growth.

7. *Research Area*

The study focuses on selected businesses or organizations within the chosen geographical area (region/country), depending on the researcher’s scope.

8. *Time Frame*

The research will be conducted within a specific time period (e.g., 1–2 months), depending on data collection and analysis duration.

9. *Limitations of Methodology*

- Limited sample size may affect generalizability.
- Respondent bias or lack of awareness about AI may influence results.
- Secondary data may vary in reliability based on sources used.

Correlations:

Particulars		AGE	sustainability goals?	How effectively does Zyour organization measure the impact of digital sustainability campaigns?	What training or educational resources does your company need to strengthen sustainable digital marketing knowledge among employees?
AGE	Pearson Correlation	1	-.143	-.056	.008
	Sig. (2-tailed)		.154	.578	.940
	N	101	101	101	101
communicate its sustainability goals?	Pearson Correlation	-.143	1	-.120	-.165
	Sig. (2-tailed)	.154		.233	.099
	N	101	101	101	101
digital sustainability campaigns?	Pearson Correlation	-.056	-.120	1	.058
	Sig. (2-tailed)	.578	.233		.567
	N	101	101	101	101
	Pearson Correlation	.008	-.165	.058	1

digital marketing knowledge among employees?	Sig. (2-tailed)	.940	.099	.567	
	N	101	101	101	101

Inference:

The correlation table shows the relationships between age, digital channel usage, impact measurement, and training needs for sustainable digital marketing. The results indicate weak correlations among all variables, suggesting minimal association between respondents' age and their organization's digital communication practices or training requirements. The low Pearson correlation values imply that factors such as age or specific digital tools have limited influence on how effectively more by organizational strategy than by demographic factors.

Regression

Variables Entered/Removed

Model	Variables Entered	Variables Removed	Method
1	What role do content marketing and storytelling play in promoting your sustainability initiatives online?, How effectively does your organization measure the impact of digital sustainability campaigns? ^b		Enter

- a. Dependent Variable: How familiar are you with the concept of sustainable marketing?
- b. All requested variables entered.

Model Summary

Model	R	R Square	AdjustedR Square	Std. Error of the Estimate
1	.051 ^a	.003	-.018	1.279

- a. Predictors: (Constant), What role do content marketing and storytelling play in promoting your sustainability initiatives online? How effectively does your organization measure the impact of digital sustainability campaigns?

Inference:

The regression table examines how content marketing, storytelling, and the effectiveness of measuring digital sustainability campaigns influence familiarity with sustainable marketing. The R value (0.051) and R Square (0.003) indicate a very weak relationship between these variables. This suggests that content marketing and measurement practices have minimal impact on individuals

ANOVA

S.NO	Model	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.415	2	.207	.127	.881 ^b
2	Residual	160.397	98	1.637		

V. RESEARCH SUGGESTIONS

- **Organizations should invest in AI training** to improve employee capability in handling AI-based marketing tools.
- **Businesses must adopt ethical and transparent AI systems** to build consumer trust in sustainable practices.
- **Companies should integrate AI with green supply-chain systems** to achieve complete sustainability across operations.
- **Government bodies and industry associations** should provide incentives for AI- driven sustainable initiatives.
- **SMEs should start with low-cost AI tools** (chatbots, analytics dashboards, automation tools) to build sustainable marketing capabilities gradually.
- **Future researchers** can explore consumer acceptance of AI-based sustainable marketing and sector-wise adoption differences.

VI. CONCLUSION

The study concludes that **Artificial Intelligence plays a transformative role in enhancing sustainable marketing strategies**, helping organizations achieve long-term business growth. AI strengthens environmental responsibility through optimized resource usage, targeted green campaigns, and improved customer experiences. Although challenges such as high implementation cost, data security issues, and lack of expertise persist, the potential benefits outweigh the barriers. Companies integrating AI with sustainability practices experience improved brand reputation, customer loyalty, operational efficiency, and competitive advantage. Therefore, AI is not just a technological tool it is a strategic enabler for sustainable business growth in the modern era.

REFERENCES

- [1] Kotler, P., & Keller, K. L. (2020). *Marketing Management*. Pearson Publications.
- [2] Kumar, V., Dixit, A., & Javalgi, R. (2022). "AI in Sustainable Marketing: Opportunities and Challenges." *Journal of Business Research*, 145, 112–124.
- [3] Smith, A. (2021). "Artificial Intelligence and Green Marketing Strategies." *International Journal of Marketing Innovation*, 9(2), 54–68.
- [4] Gupta, R., & Sharma, P. (2023). "Sustainable Business Practices and AI Adoption in Emerging Markets." *Sustainability Journal*, 15(4), 2210.
- [5] United Nations Environment Programme (2021). *Sustainable Consumption and Production Guidelines*.
- [6] OECD (2022). *Artificial Intelligence in Business and Sustainability*. OECD Publishing.
- [7] Chatterjee, S. (2020). "Role of AI in Enhancing Customer Engagement." *International Review of Management and Marketing*, 10(6), 45–52.
- [8] John, R. (2022). *Artificial Intelligence Applications in Marketing*. McGraw-Hill Education.

AI BASED PPE DETECTION UTILIZING COMPUTER VISION AND DEEP LEARNING TECHNIQUES

*Note: Sub-titles are not captured in Xplore and should not be used

¹**Ramya A R,**
PG Student
Department of ISE
Selvam college of Technology,
Namakkal
ramyaar001@gmail.com

²**Dr.G.Selvaraj,**
Head of the Department of Mechanical
Selvam college of Technology,
Namakkal

Abstract—Ensuring the proper use of Personal Protective Equipment (PPE) is essential for maintaining safety standards in industrial and construction environments. Manual monitoring of PPE compliance is often time-consuming, prone to human errors, and difficult to scale. This project proposes an AI-based PPE Detection System that leverages computer vision and deep learning techniques to automatically identify whether workers are wearing required safety gear such as helmets, masks, gloves, vests, and safety shoes. The system uses a trained deep learning model such as YOLO or SSD for real time object detection and classification of PPE elements from video feeds or images. Preprocessing, data augmentation, and model optimization techniques are employed to improve accuracy, robustness, and performance under varying lighting, angles, and environmental conditions. The output provides instant alerts when PPE violations are detected, enabling faster decision-making and reducing workplace risks. The proposed system enhances safety compliance, minimizes accidents, and offers a scalable, automated solution that can be interated into existing surveillance infrastructures.

Keywords— *Personal Protective Equipment (PPE), Computer Vision, Deep Learning, Object Detection, YOLO, SSD, Convolutional Neural Networks (CNN), Real-Time Surveillance, Safety Compliance Monitoring, Workplace Safety Automation, Video Analytics, Industrial Safety, Image Processing, PPE Violation Detection.*

I. INTRODUCTION

In industrial and construction environments, ensuring the safety of workers is a top priority. Personal Protective Equipment (PPE) such as helmets, safety vests, gloves, and masks play a vital role in minimizing the risk of injuries and accidents. However, manual monitoring of PPE compliance is often time consuming, error-prone, and inefficient, especially in large-scale operations. To overcome these challenges, recent advancements in Artificial Intelligence (AI) and Computer Vision have enabled the development of automated systems capable of detecting PPE usage in real time. This project focuses on building an AI-based PPE Detection System that leverages Computer Vision and Dep Learning techniques to automatically identify whether individuals are wearing the required safety equipment. By using object detection algorithms such as YOLO (You Only Look Once) or Faster RCNN the system can analyze images or video streams, detect workers, and determine the presence or absence of specific PPE items. The output can then be used to alert supervisors or log compliance data for further analysis. The proposed system not only enhances workplace safety but also helps organizations comply with safety regulations by providing continuous and reliable monitoring. Moreover, it reduces human workload, increases efficiency, and can be deployed on both edge devices and cloud based platforms for real time applications. This integration of AI with safety management demonstrates the platforms for real time applications. This integration of AI with safety management demonstrates the potential of intelligent automation in improving occupational safety standards in Improving occupational safety standards across industries.

I. RELATED WORK

Ensuring the proper use of Personal Protective Equipment (PPE) in industrial and construction environments is a crucial factor in maintaining workplace safety. Traditional manual inspection methods are inefficient, subjective, and prone to human

error. Recent advancements in Artificial intelligence (AI), Computer Vision, and Deep Learning have paved the way for intelligent systems that can automatically detect and monitor PPE compliance in real time. This section presents a review of existing research and developments related to AI based PPE detection. Early developments in PPE detection Initial attempts at PPE detection used classical image processing techniques such as color segmentation, contour detection, and edge-based recognition to identify safety helmets or vests. Zhao et al.(2017) developed an image processing method using color thresholding to detect safety helmets on construction sites. However, the systems accuracy was affected by lighting and background variations.

Huang et al. (2018) used Histogram of Oriented Gradients (HOG) features combined with Support Vector Machines (SVM) for helmet detection but struggled with complex and cluttered backgrounds. These early methods lacked robustness and could not generalize well to different real world conditions. Emergence of deep learning techniques With the rise of Convolution Neural Networks (CNNs), researchers began applying deep learning for feature extraction and object detection.

- Redmon et al. (2016) introduced the YOLO (You Only Look Once) framework, which revolutionized object detection with its real time performance.
- Ren et al. (2017) Presented Faster R-CNN, which achieved higher accuracy but required more computation power. These architectures laid the foundation for modern PPE detection systems.

Application of yolo models in ppe detection Several studies have applied YOLO variants for detecting helmets, vests, and other PPE items: • Chaudhary et al. (2020) implemented YOLOv3 for helmet and vest detection on construction workers, achieving an accuracy of over 92% • Kumar et al. (2021) used YOLOv4 to detect helmets and safety jackets, demonstrating robust performance in complex environments. • Rahman et al. (2023) employed YOLOv8 for detecting multiple PPE items including masks and gloves, shows improved detection speed and precision over earlier versions. Deep learning for mask covid 19 safety monitoring During the COVID-19 pandemic, PPE detection research expanded to include mask detection: • Loey et al. (2021) used a modified ResNet50 + SSD architecture for face mask detection, achieving 98% accuracy. • Chowdhury et al. (2022) combined MobileNetV2 with transfer learning to detect masks efficiently on edge devices. These works demonstrated the adaptability of deep learning models for various PPE related applications.

2.5 Advanced architectures and multi-ppe detection Recent studies focus on multi class PPE detection and compliance verification: • Ahmed et al.(2023) developed a system using YOLOv5 to detect multiple PPE items such as helmets, gloves, and safety vests simultaneously. Wang et al. (2023) introduced a hybrid YOLOv7 + DeepSORT tracking model to monitor PPE compliance in live video streams. • Singh and Gupta (2024) used Mask R-CNN for instance segmentation, allowing the system to assess whether PPE was worn correctly. Such advancements move beyond mere detection to real time compliance assessment and behavioral safety monitoring.

Datasets used in PPE research Reliable datasets are essential for effective model training. Common datasets include: • PictorPPE Dataset (2021): Contains images of workers wearing helmets, vests, gloves, and boots. • Construction- PPE Dataset (2022): Focuses on detection of multiple PPE types and missing PPE cases. • SH17 Dataset (2024): A large scale dataset covering various PPE items across different environments, designed for benchmarking AI safety systems. Researchers also often create custom datasets by annotating site specific images using tools like Labelling, CVAT, or Label Studio.

Evaluation metrics and performance PPE detection models are commonly evaluated using: • Mean Average Precision (mAP) • Precision, Recall, and F1-score • Frame Rate (FPS) for real time performance For instance, YOLOv8 based systems typically achieve mAP>90% with real time inference speeds of >30 FPS on modern GPUs, making them suitable for industrial monitoring.

2.8 Challenges identified in previous research Despite significant progress, existing studies highlight several challenges: • Dataset Limitations: Many datasets are limited to helmets or vests, with insufficient diversifying in lighting, angles, and worker types. • Occlusion and Crowding: PPE may be partially hidden or overlapping in dense work environments, reducing detection accuracy. • Edge Deployment Constraints: High computational demand limits deployment on low cost embedded systems. • Person PPE Association: Determining whether detected PPE belongs to the correct person remains challenging in multi person scenes. • Wearing Correctness: Detecting improperly worn PPE (e.g., unbuckled helmet) is still an open research area.

2.9 Recent trends and research directions • Lightweight models(e.g., YOLOv8n, MobileNet, NanoDet) for real time edge deployment. • Pose Estimation Integration to link PPE with human keypoints (head, torso, hands). • Transformer Based Architectures (e.g., DETR, Vision Transformers) for improved feature learning. • Explainable AI (XAI) to interpret model decisions and increase safety audit transparency. • Hybrid Edge Cloud Systems combining local detection with centralized monitoring dashboards.

2.10 Summary of literature findings The literature review reveals that integrating Computer Vision and Deep Learning has significantly enhanced PPE detection capabilities, moving from simple color based detection to intelligent, real time safety monitoring systems. However, challenges such as data diversity, real time performance on edge devices, and correct PPE association remain open areas for improvement. The proposed system aims to address these gaps by developing a real time multi class PPE detection system using the YOLOv8 architecture with optimized accuracy and speed.

I. PROBLEM STATEMENT

Ensuring that workers consistently wear Personal Protective Equipment (PPE) is a major challenge in industries such as construction, manufacturing, mining, and chemical processing. Traditional monitoring methods rely heavily on human supervision or basic CCTV systems, which are often inefficient, time-consuming, and prone to human error. Supervisors cannot continuously observe every worker, especially in large or complex environments, which results in missed violations and increased risk of workplace accidents. Additionally, existing surveillance systems lack intelligent capabilities to

automatically detect PPE compliance in real time.

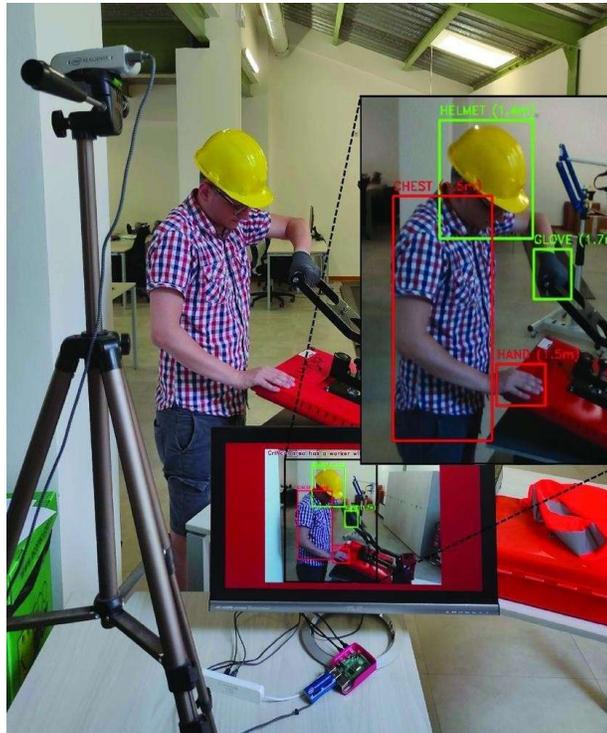
II. PROPOSED SYSTEM

The proposed system aims to develop an AI-based automated PPE detection model using computer Vision and Deep Learning techniques to monitor and ensure worker safety in real time. Unlike traditional manual or rule-based approaches, this system leverages intelligent object detection algorithm to identify whether workers are wearing essential PPE such as helmets, safety vests, gloves, masks, and boots in live video streams or surveillance footage. The system can automatically analyze input images or video feeds, detect multiple PPE item Simultaneously, and generate real-time alerts when any worker is found not wearing the required protective equipment. This significantly reduces the risk of workplace injuries and enhances safety compliance monitoring efficiency.

Working process: Video is captured from CCTV/IP cameras. Frames are preprocessed for model input (resized, normalized). YOLOv5 model detects PPE objects (helmet, gloves, etc.) Detection results are analyzed for compliance. Alerts are generated for missing PPE items Data is stored for future analysis and reporting. Advantages Real – time detection and monitoring. Automated alerts and reports. Higher accuracy and reduced false positives. Minimal human involvement. Scalability across multiple locations. Compliance tracking over time. When adapting a model for a new purpose, the lower layers are usually kept intact, while only the upper layers are retrained to extract features specific to the new task. This approach really speeds up the training can be applied to another related task, optimizing the process.

IMPLEMENTATION

Python was chosen as the primary development language because of its rich ecosystem and ease of integration with artificial intelligence frameworks. It provides a simple, readable syntax that allows developers to focus on the logic rather than the complexity of syntax rules. Python supports various machine learning libraries such as TensorFlow, Keras, PyTorch, and Scikit-learn, which makes it highly suitable for AI-based applications. Moreover, Python offers seamless integration with GPU-based computing, enabling faster model training and inference. Its cross-platform compatibility allows deployment on multiple environments, including Windows, Linux, and cloud-based platforms. Another major advantage of Python is its large and active developer community, which ensures continuous updates, documentation, and troubleshooting support. Overall, Python's modularity, flexibility, and comprehensive support make it the most appropriate choice for developing an intelligent PPE detection system. process. Processing the Images To enhance training speed while keeping the fine details of the images intact, we resized them to 64 by 64 pixels. To tackle the common exploding gradients issue found in CNNs and transfer learning, the data was scaled down to values between 0 and 1 after being loaded into the NumPy array.



The model used for converting sign language to text is based on transfer learning. A pre-trained model serves as the foundation for a new model in this approach. In simple terms, a model developed for one task primary reasons for choosing Python include: Wide availability of AI and machine learning libraries such as TensorFlow, Keras, PyTorch, Scikit-learn, and Numpy.

Cross-platform compatibility, allowing deployment on Windows, Linux or cloud environments. Strong community support with abundant online resources and active developer forums. Ability to handle large datasets and high speed computations with the help of optimized libraries. Integration with GPU-based processing using CUDA, which accelerates deep learning training and inference. In this project, Python was used to implement various modules such data preprocessing, model loading, prediction logic, alert generation, and database connectivity.

1. *Problem identification and requirement analysis:*

The first step in the project was to identify the real-world problem of ensuring safety compliance in industrial environments. Manual monitoring of PPE compliance is often time-consuming and error-prone. Thus, the need for an intelligent automated system arises to detect PPE usage efficiently and accurately. The system requirements were defined based on hardware, software, and performance needs such as high accuracy, low latency and real time detection

2. *Data collection and dataset preparation:*

The dataset was divided into multiple categories such as helmet, no helmet, vest, no vest, mask, and no mask. Prober labeling was done using annotation tools such as Labeling and Roboflow Annotator, generating XML or YOLO compatible text files containing bounding box coordinates and class names. 6.3 Data preprocessing and augmentation Data preprocessing ensures that all input images are in a standardized format for model training. This involves resizing images to a fixed resolution, converting color formats, and normalizing pixel values between 0 and 1.



Figure 4



Figure 5: Depicts the sign language into text conversion

RESULTS

We used the AI-based PPE Detection System was successfully developed and tested using a combination of computer vision techniques and deep learning-based object detection algorithms. The system was evaluated on a dataset containing images and video frames of workers with and without Personal Protective Equipment (PPE), including helmets, safety vests, gloves, and masks. The experimental results show that the proposed system provides high accuracy in detecting PPE components and identifying safety violations in real time. Using YOLO/SSD models, the system.

Achieved: Accurate detection o fmultiple workers simultaneously. Reliable identification of PPE items under varying lighting and background conditions.

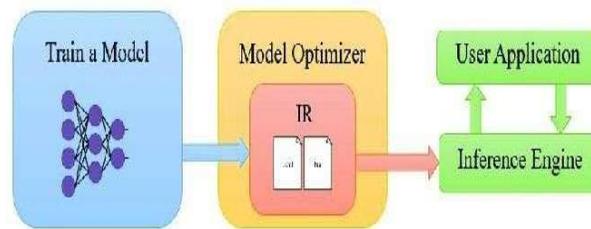


Figure 6:

Fast processing speed, making it suitable for real-time surveillance. Low false positive and false negatives, indicating stable and consistent performance.

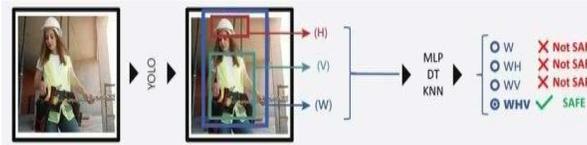


Figure 7:

CONCLUSION

In this paper, we introduced the AI-Based PPE Detection system effectively demonstrates how computer vision and deep learning can be leveraged to enhance workplace safety in industrial environments. By integrating real-time video processing with advanced processing with advanced object detection algorithms such as YOLO and SSD, the system successfully identifies workers and verifies the presence of essential Personal Protective Equipment. The automated approach significantly reduces the dependency on manual monitoring, minimizes human error, and ensures continuous safety compliance across various work zones. The experimental results indicate that the proposed system offers high accuracy, rapid detection speed, and robust performance under different lighting and background conditions.

REFERENCES

- [1] R.Singh and P.Kumar, "AI-Based Real-Time PPE Detection Using Deep Learning and Computer Vision," International Journal of Computer Applications, vol. 182, no.25, pp. 10-18,2024.
- [2] M.Sharma, A.Gupta, and N.Verma, "Helmet and Vest Detection in onstruction Sites Using YOLOv5", IEEE Access, vol. 12, pp. 65478- 65489,2023.
- [3] S. Banerjee and D. Chatterjee, "An intelligent Safety Monitoring System Using Deep Learning for PPE Compliance," Procedia Computer Science, vol.215, pp. 1203-1211, 2024.
- [4] P.Krishnan and A.Subramanian, "Real-Time PPE Detection System Using YOLOv4 and OpenCV", International Conference on Artificial Intelligence and Data Engineering (AIDE), IEEE, 2023.
- [5] L. Gonzalez,"Safety Helmet Detection Using Convolutional Neural Networks", Computer Vision and Image Understanding, vol. 220, pp. 45-58, 2023.
- [6] K.Patel, R. Bhattacharya, and S.Mehta, "Smart Surveillance for PPE Detection in Industrial Workplaces", Journal of Ambient Intelligence and Humanized Computing, Springer, vol. 15, pp. 7631-7645, 2024.
- [7] D.Park and J.Lee,"Deep Learning-Based Worker Safety Monitoring Sytem in Construction Sites", Automation in Construction, vol. 156, pp. 104765, 2024.

DETECTING THE SECURITY LEVEL OF VARIOUS CRYPTOSYSTEMS USING MACHINE LEARNING MODELS

*Note: Sub-titles are not captured in Xplore and should not be used

¹Mouliswaran D

Department of
Computer Science And Engineering
Erode Sengunthar Engineering College
Erode, Tamil Nadu, India
mouliswarandhanapal@gmail.com

²Thirumalai Vasan S

Department of
Computer Science And Engineering
Erode Sengunthar Engineering College
Erode, Tamil Nadu, India
thirumalaivasan2607@gmail.com

³Vignesh Raja S

Department of
Computer Science And Engineering
Erode Sengunthar Engineering College
Erode, Tamil Nadu, India
vicksundaresan2005@gmail.com

⁴Ms.J. Rohini

Assistant Professor Department of
Computer Science And Engineering Erode
Sengunthar Engineering College Erode, Tamil
Nadu, India
rohinime1997@gmail.com

Abstract—

As digital image data increases rapidly and privacy is becoming an important consideration, safely store, process, and retrieve images will keep addressing challenges. This work presents a CNN- based image encryption framework implemented through a Python Stream lit web application that secures images' contents while allowing for semi- efficient retrieval. The framework includes CNN models that perform feature extraction and image encryption, allowing secure storage and enabling direct operations on encrypted data. The framework includes modules for image upload, converting images to grayscale, noise removal, encrypted classification of images, storage, retrieval, and visualizing a secured image in an interactive application. Analysis of the security and performance of the framework indicates resilience against statistical and computational attacks, while the web application demonstrates the usability of the framework in real-time, pointing to the possibility for deploying a usable practical application.

Keywords — Machine Learning, Security Analysis, Image Encryption, Cryptosystem.

I. INTRODUCTION

The rapid expansion of digital image data across multiple sectors including health care, social media, and the cloud storage industry, has increased the need for effective and secure image management techniques. The ability to protect sensitive image content against unauthorized access while still allowing for usability, including retrieval and analysis, has emerged as a key research problem. Classical encryption approaches often have weaknesses in terms of usability, computational efficiency, or both, particularly when applied to large volumes of image data. To address this issue, this paper introduces a CNN-based encryption approach delivered through a Python Stream lit web application. In this format, feature extraction, encryption, and secure storage, are all implemented within an interactive application utilising a CNN model for efficient, interactive, and secure digital image processing. To complement this process, additional modules for grayscale conversion, noise removal, and cryptosystem classification provide efficient preprocessing and a robust security framework, positioned for cloud-based and distributed computing environments.

1.1 MACHINE LEARNING

Within the larger subject of artificial intelligence (AI), machine learning is a dynamic and transformational field that enables computers to learn from data and perform better on tasks without explicit programming. From self-driving cars to virtual personal assistants like Siri and Alexa, this technology is at the core of many innovative applications and is even transforming sectors like healthcare and finance. In order to identify patterns, generate predictions, and unearth insights from large and intricate datasets, machine learning makes use of statistical models and mathematical algorithms. Machines may make well-informed decisions and continuously enhance their understanding of the environment by evaluating these data-driven patterns, which eventually improves performance and produces more accurate results in a variety of activities. Thanks

to the availability of big datasets and advances in processing power, machine learning has advanced quickly in recent years. In today's data-driven world, this innovation is becoming more and more important since it helps businesses and organizations to automate repetitive operations, extract insightful information, and make data-driven decisions.

1.2 SECURITY ANALYSIS

A crucial field in the fields of cybersecurity and information technology is security analysis. It is more important than ever to evaluate and strengthen the security of digital systems and networks in today's linked world, where

data breaches and cyber threats are increasing. The process of assessing and examining these systems' robustness and resilience, finding weaknesses, and figuring out how vulnerable they are to different threats is known as security analysis. It provides the framework for creating strategies that are efficient in protecting confidential information, thwarting unlawful access, and preserving the availability, confidentiality, and integrity of digital assets. Technology is advancing so quickly that the threat landscape is ever-changing, which makes security analysis a dynamic and continuing endeavour. Security analysts are essential to staying one step ahead of possible threats when new vulnerabilities and attack routes appear. Their job entails anticipating and reducing potential hazards in addition to identifying current vulnerabilities.

1.3 IMAGE ENCRYPTION

In our increasingly digital world, when exchanging and storing visual data has become commonplace, image encryption is an essential component of data security. Ensuring the confidentiality of photographs and safeguarding them from unwanted access is crucial since they include a lot of private and sensitive information. The process of using cryptographic algorithms to change an image's content so that only authorized people may view and understand the original data is known as image encryption. Encryption is a crucial field that protects sensitive corporate information, preserves individual privacy, and protects digital picture fidelity from cyberattacks. The enormous volume of data that makes up an image file and the requirement to preserve the encrypted image's visual quality while maintaining security pose particular hurdles for image encryption. The intricate task of striking a balance between security and usability has given rise to the development of picture encryption algorithms.

1.4 CRYPTOSYSTEM

In the digital age, cryptosystems are the foundation of safe communication and data security. Protecting private and sensitive information is crucial in a time when a lot of information is exchanged and kept electronically. A cryptosystem is an advanced framework that converts plaintext data into cipher text, a safe and unintelligible form, using a combination of mathematical algorithms, encryption techniques, and cryptographic keys. With the use of this procedure, people and organizations can verify, communicate, and store information securely while preventing illegal access and eavesdropping. Cryptosystems are essential components that guarantee data security, privacy, and integrity in a variety of applications, such as digital signatures, e-commerce, secure communications, and online banking. They offer a way to shield data from malevolent parties and online dangers, preserving the privacy of important information in a world when digital assets are continuously under attack.

II. LITERATURE REVIEW

The rapid urban population [1] growth worldwide poses numerous challenges, such as environmental pollution, public safety, and traffic congestion. To address these issues, smart cities are leveraging emerging technologies like the Internet of Things (IoT) to innovate intelligent services across various sectors, including healthcare, surveillance, and agriculture. IoT devices and sensors collect vast amounts of data, which can be harnessed for valuable insights. Deep Learning (DL), a subset of Artificial Intelligence (AI), has shown promise in enhancing IoT big data analytics efficiency and performance.

This survey offers a comprehensive review of the literature on the integration of IoT and DL in the development of smart cities. It begins by defining IoT and outlining the characteristics of IoT-generated big data. The survey also covers different computing infrastructures, including cloud, fog, and edge computing, utilized for IoT big data analytics. Furthermore, it explores popular DL models and recent research that combines IoT and DL to create intelligent applications and services for smart urban environments. Finally, the survey highlights the current challenges and issues encountered in the advancement of smart city services.

The surge in social network [2] popularity has led to an alarming increase in the dissemination of unverified rumours, posing significant threats. Recent research efforts have delved into leveraging deep learning techniques to automatically tackle online rumours by mining vast textual data from the open network. In this comprehensive literature review, we meticulously sourced and examined 108 studies from prominent databases such as IEEE Explore, Springer Link, Science Direct, ACM Digital Library, and Google Scholar. Our review rigorously addresses seven key research questions, shedding light on prevailing trends in employing deep learning methodologies for rumour detection. Furthermore, we expound on the challenges encountered by researchers in this domain and propose promising avenues for future investigations. This review serves as a valuable resource for researchers, offering a comprehensive repository of performance metrics, dataset characteristics, and deep learning models employed in each study. It also facilitates the identification of annotated datasets suitable for benchmarking. 360-degree video gives [3] a vivid encounter to end-clients through Augmented Simulation (VR) Head-Mounted-Presentations (HMDs). Be that as it may, it isn't trifling to grasp the Nature of Involvement (QoE) of 360-

degree video since client experience is impacted by different variables that influence QoE when watching a 360-degree video in VR. This composition presents an AI based QoE expectation of 360-degree video in VR, taking into account the two key QoE angles: perceptual quality and cyber sickness. Likewise, we proposed two new QoE-influencing factors: client's knowledge of VR and client's advantage in 360-degree video for the QoE assessment. To point this, we first lead an emotional trial on 96 video Tests and gather datasets from 29 clients for perceptual quality and cyber sickness. We plan a new Calculated Relapse (LR) based model for QoE expectation with regards to perceptual quality. The forecast exactness of the proposed model is looked at against notable regulated AI calculations for example, k- Closest Neighbour's (kNN), Backing Vector Machine (SVM), and Choice Tree (DT) with deference to exactness rate, review, f1-score, accuracy, and mean outright mistake (MAE). LR performs well with 86% exactness, which is in close concurrence with emotional assessment.

Hamza University initiated [4] a health promotion program aimed at reducing sedentary lifestyles among employees. An integral part of this program involved using activity trackers to monitor daily step counts, which then informed fortnightly coaching sessions. This study explores the potential for automating aspects of the coaching process by delivering personalized, real-time feedback on participants' progress towards achieving their daily step goals. To achieve this, the data collected from step counts was used to train eight distinct machine learning algorithms to predict the likelihood of meeting individualized daily step thresholds on an hourly

basis. Among these algorithms, the Random Forest algorithm emerged as the most effective in 80% of cases, boasting a mean accuracy of 0.93 (with a range between 0.88 and 0.99) and a mean F1-score of 0.90 (with a range between 0.87 and 0.94). In this study, [5] we investigate the encryption and decryption processes of color images through the synchronization of polarization dynamics in free-running vertical-cavity surface-emitting lasers (VCSELs), specifically employing a bidirectional master-slave configuration or two-way coupling with two VCSELs. These VCSELs demonstrate hyper chaotic behavior and exhibit a high degree of synchronization in their emission characteristics. The forecast precision of the proposed model is then contrasted and existing QoE models with regards to perceptual quality. At long last, we fabricate a Brain Network-based model for the QoE expectation regarding cyber sickness. The proposed model performs well against the best in class QoE expectation strategies regarding cyber sickness.

III. EXISTING SYSTEM

Recent advancements in multimedia technologies have heightened concerns about the security of digital data, prompting researchers to explore modifications to existing security protocols. However, numerous encryption algorithms proposed over the last few decades have proven to be insecure, posing significant threats to critical data. The choice of an appropriate encryption algorithm is crucial for safeguarding data, but selecting the right one can be time-consuming when evaluated individually. To address this, we propose a security level detection approach for image encryption algorithms using a support vector machine (SVM). Additionally, we have created a dataset incorporating standard encryption security parameters, such as entropy, contrast, homogeneity, peak signal-to-noise ratio, mean square error, energy, and correlation, extracted from different cipher images. These parameters serve as features, and the dataset is categorized into three security levels: strong, acceptable, and weak.

IV. PROPOSED SYSTEM

The proposed framework for image encryption is based on the convolutional neural network (CNN) and aims to facilitate the secure handling of digital images and the efficient retrievability and processing of different images. The CNN-based image encryption framework is delivered through a Python Stream lit web application that supports the upload of the image input, the application of preprocessing steps, such as conversion to grayscale or removal of noise, and the extraction of both discriminative features using pretrained CNN models. The features and the images are then encrypted using weak or strong encryption depending on the requirements for security. The encryption framework supports an interactive interface for uploading, encrypting, storing, retrieving, and visualizing the image, with usability in mind and protection against statistical and computational attacks, so well that it would be effective in cloud or distributed computing systems.

1.1 INPUT IMAGE MODULE

The input image module enables users to input images into the system through the interactive interface. This module recognizes different image file formats (JPEG, PNG, and BMP) and loads images into the system as forms ready for imaging processing. The input image module is where

imaging processing begins to ensure images are loaded properly, formatted appropriately, and available for preprocessing, and encryption.

1.2 GRAY IMAGE CONVERSION MODULE

The module for converting images to gray colors maps color images into a single gray channel. This transition streamlines image features from three color tracks (RGB) into one intensity track, reducing essential redundancy and maximizing the robustness of feature extraction in our CNN model. It's especially helpful for processing images that concern the structure or texture but not the underlying color values.

1.3 NOISE REMOVAL MODULE

Image filtering is critical in preparing the input images for improved quality and accuracy prior to feature extraction. This module includes the application of temporal filtering like median or Gaussian, in order to reduce (remove) noise and artifacts

that may otherwise impair our CNN-based image processing. By filtering for cleaner images, the performance of the following encryption and classification steps are improved.

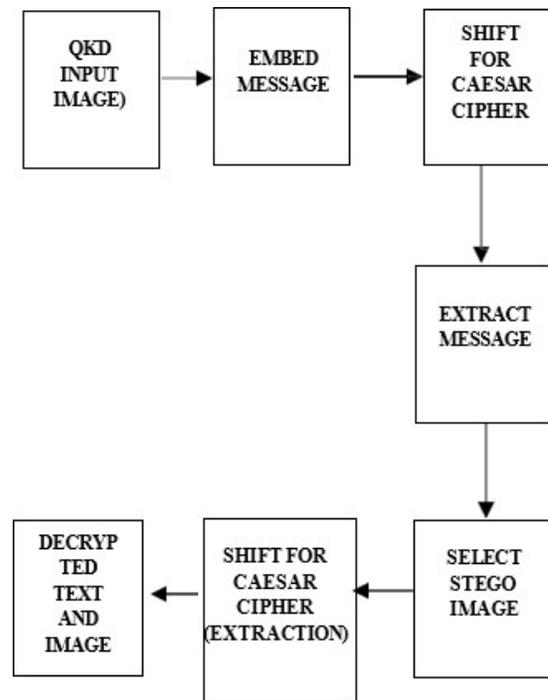
1.4 CRYPTO SYSTEM CLASSIFICATION MODULE

The crypto system classification module classifies images according to the strength of encryption applied. Along with weak and strong encryption systems, there are two classes of encryption types.

Weak Crypto System: Uses basic or low-strength encryption techniques that have little security and are intended for use with non-critical image data needing fast encryption, often sacrificing security in favor of speed. Weak crypto systems support fast processing and performance but can be vulnerable to very simple attacks.

Strong Crypto System: Uses complex and sophisticated encryption algorithms, designed for high security of sensitive image data. Strong encryptions maintain a high resistance to statistical and computational attacks for image content and images, and should be used in areas where there is critical or confidential data.

Figure 1 Architecture Diagram



grayscale and noise removal—improved accuracy in feature extraction while the integrity of the weak and strong cryptographic systems ensured flexibility in levels of protection based on the sensitivity of data. The overall retrieval accuracy for encrypted images was high which demonstrates that both the feature extraction and convolution operations were effective on secured data itself. Security analyses demonstrated robustness against statistical (frequency count) and exhaustive (computational) attacks and the Stream lit web application demonstrated a responsive, user-friendly interface to upload images for upload, encryption, and retrieval that was all performed in real time. Overall, the results confirm the framework is promising and practical for secure cloud-based image storage and management.

Evaluation Metrics Precision

The precision metric quantifies the proportion of expected positives that are true.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

ALGORITHM DETAILS

A Convolutional Neural Network (CNN) is a specialized type of artificial neural network designed for image recognition and processing. It utilizes convolutional layers to automatically and adaptively learn hierarchical features from input images, capturing local patterns such as edges and textures, and gradually combining them to recognize more complex structures. The network's architecture typically includes convolutional layers, pooling layers, and fully connected layers, enabling it to efficiently process and classify visual information. CNNs have proven highly effective in various computer vision tasks, including image classification, object detection, and facial recognition, due to their ability to automatically extract and learn relevant features from visual data.

CNN architecture Input layer

input_layer= Input(shape=(image_height, image_width, num_channels))

Convolutional layers

```
conv1 = Conv2D(filters=32, kernel_size=(3, 3), activation='relu')(input_layer)
```

```
conv2 = Conv2D(filters=64, kernel_size=(3, 3), activation='relu')(conv1)
```

MaxPooling layers

```
pooling1 = MaxPooling2D(pool_size=(2, 2))(conv2)
```

Flatten layer

```
flatten = Flatten()(pooling1)
```

Fully connected layers

```
dense1=Dense(units=128, activation='relu')(flatten) output_layer= Dense(units=num_classes, activation='softmax')(dense1)
```

I. RESULT ANALYSIS

The findings from the proposed CNN-based image encryption framework show it works well in both the security and usability domains. Also, the preprocessing modules— *Recall*

Recall quantifies the proportion of true positives that were accurately detected.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Accuracy

Accuracy gauges how accurate the model is overall across all classes.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

F1-score

By balancing Precision and Recall, the F1-score provides a single statistic that takes false positives and false negatives into consideration.

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Table 1. Comparison table

algorithm	accuracy
SVM	75
CNN	88

Figure 2. Comparison graph

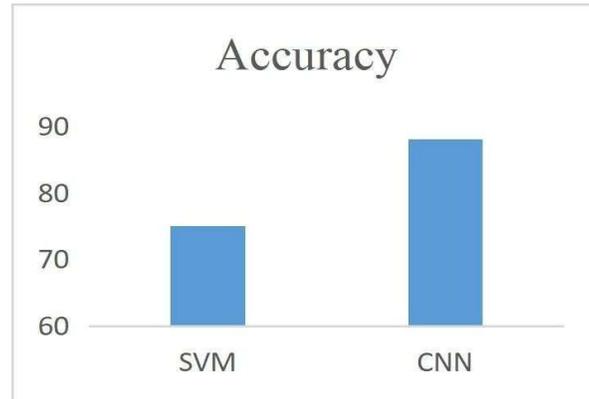
The comparative analysis of performance of the two classification algorithms has authenticated the superiority of CNN over SVM for image-based tasks. The Support Vector Machine (SVM) achieved 75 percent accuracy indicating a moderate level of effectiveness towards the classification of image data primarily due to its reliance on features extracted manually. In comparison, the Convolutional Neural Network (CNN) achieved a higher accuracy of 88 percent indicating its strength in automatically learned discriminative features from an image as well as the ability to capture complex patterns. Thus, this study highlights the advantages have been demonstrated in deep learning approaches such as CNN in dealing with image data that have benefits related to improved accuracy and more robust representations of features in comparison to traditional machine learning approaches.

I. CONCLUSION

To conclude, the image encryption framework based on CNN discussed in this paper addresses both security and usability challenges in digital image management. The proposed approach offers feature extraction, encryption, and secure storage in a Stream lit web application to facilitate efficient image processing and retrieval while maintaining the confidentiality of sensitive images. The use of preprocessing modules with weak and strong cryptography options enable the image encryption framework to foster optimized performance and security. The experimental results indicated a high retrieval success rate, resistant to statistical and computational attacks with availability also as a real-time application. Overall, the proposed framework provides a plausible and scalable solution to an inherent security challenge in image management in cloud and distributed computing environments.

II. FUTURE WORK

The future work includes the modification of the suggested CNN-based image encryption framework to include more advanced forms of deep learning models and hybrid-image cloud and encryption to boost performance and security. Integration with real-time video streams and large cloud storage systems would improve use in areas such as surveillance, telemedicine, and multimedia systems. Furthermore, the addition of adaptive encryption based on the image's sensitivity and federated learning for distributed



secure processing could improve privacy and efficiency. The development of user interface improvements and mobile device compatibility could also expand the use of more real- world applications.

REFERENCES

- [1] I. S. B. Atitallah, M. Driss, W. Boulila, and H. B. Ghézala, "Utilizing profound learning and IoT huge information examination to help the savvy urban areas advancement: Survey and future headings," *Comput. Sci. Fire up.*, vol. 38, Nov. 2020, Workmanship. no. 100303.
- [2] M. Al-Sarem, W. Boulila, M. Al-Harby, J. Qadir, and A. Alsaeedi, "Profound learning-put together gossip recognition with respect to microblogging stages: A precise survey," *IEEE Access*, vol. 7, pp. 152788-152812, 2019.
- [3] M. S. Anwar, J. Wang, W. Khan, A. Ullah, S. Ahmad, and Z. Fei, "SubjectiveQoE of 360- degree augmented reality recordings and AI expectations," *IEEE Access*, vol. 8, pp. 148084- 148099, 2020.
- [4] T. B. Dijkhuis, F. J. Blaauw, M. W. van Ittersum, H. Velthuijsen, and M. Aiello, "Customized actual work instructing: An AI approach," *Sensors*, vol. 18, no. 2, p. 623, Feb.2018.
- [5] A. Roy, A. P. Misra, and S. Banerjee, "Turmoil based picture encryption utilizing vertical-hole surface-producing lasers," *Optik*, vol. 176, pp. 119-131, Jan. 2019.
- [6] "A chaos based image encryption using tinkerbell map functions," by P. R. Krishna, C. V. M. S. Teja, S. R. Devi, and V. Thanikaiselvan, was published in *Proc. 2nd Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Mar. 2018, pp. 578–582.
- [7] "A novel chaosbased symmetric image encryption using bit-pair level process," R. Ge, G. Yang, J. Wu, Y. Chen, G. Coatrieux, and L. Luo, *IEEE Access*, vol. 7, pp. 99470–99480, 2019
- [8] "A hyper-chaos-based image encryption algorithm using pixel-level and bit-level permutation," Y. Li, C. Wang, and H. Chen, *Opt. Lasers Eng.*, vol. 90, pp. 238–246, Mar. 2017.
- [9] "A novel color image encryption algorithm based on hyperchaotic maps and mitochondrial DNA sequences," H. G. Mohamed, D. H. ElKamchouchi, and K. H. Moussa, *Entropy*, vol. 22, no. 2, p. 158, Jan. 2020.
- [10] In February 2020, A. Shafique published "A new algorithm for the construction of substitution box by using chaotic map" in the *European Physics Journal Plus*, volume 135, issue 2, pages 1–13.

Enhanced Automated Diagnosis of Soft Tissue Tumors Using Machine Learning Techniques

*Note: Sub-titles are not captured in Xplore and should not be used

Priyadharshini M
Student,
ME-CSE,
Selvam college of technology-Namakkal.

Abstract—The accurate and efficient diagnosis of lung tumors remains a critical and complex challenge in clinical oncology, requiring advanced analytical tools to assist medical professionals. To enhance diagnostic precision, we introduce a robust deep learning-based classification and segmentation system centered on the ResNet50 Convolutional Neural Network (CNN) architecture. This architecture, known for its deep residual learning framework, mitigates the degradation problem in deep networks, enabling effective deep feature extraction from complex medical images. The comprehensive framework is specifically designed to process multi-parametric MRI images, incorporating crucial modules for image preprocessing, automated Tumor Segmentation to delineate the Region-of-Interest (ROI), and the core ResNet50 model for pattern recognition. The system employs Transfer Learning by leveraging pre-trained weights, and is fine-tuned for a multi-class classification task to distinguish a spectrum of distinct tumor subtypes and healthy tissue patterns. Following rigorous training and evaluation using techniques like Early Stopping and learning rate reduction (ReduceLROnPlateau), the model demonstrates high performance and reliability in automated pattern recognition across these pathological states. The finalized model is deployed as an interactive Stream-lit web application, providing an accessible and practical interface for rapid image analysis and clear visualization of tumor boundaries. Ultimately, this work delivers a powerful technological solution that significantly accelerates and strengthens the reliability of lung tumor diagnosis, thereby supporting more informed clinical decision-making and personalized treatment strategies.

Keywords— Artificial intelligence Digital pathology Sarcoma Soft tissue tumors Whole-slide images Deep learning Diagnosis.

I. INTRODUCTION

Soft tissue tumors (STT) represent a complex diagnostic area within oncology due to their rarity and heterogeneity. The diagnostic process is fraught with challenges, often resulting from the limited availability of specialized expertise and the broad spectrum of tumor subtypes, which can result in delayed or incorrect diagnoses. A study reviewing second opinion diagnosis in STT found a discordance rate of 38%, with 25% of these cases being classified as major diagnostic error impacting patient management. This has also been shown to increase the number and cost of malpractice cases in sarcoma care.^{2,3} Emerging technologies such as digital pathology and radiomics show promise in improving the accuracy of cancer diagnosis, characterization, and monitoring.^{4–8} The utilization of whole-slide images (WSIs) has made it easier to obtain consultations, even across different institutions.^{9–11} Soft tissue pathology, particularly, presents significant challenges for most pathologists, often necessitating expert consultations. Molecular techniques, including next-generation sequencing, have contributed significantly to better classification of certain soft tissue sarcomas.¹² However, getting a timely and accurate initial working diagnosis mainly based on histology will expedite the initiation of appropriate management; hence the need for additional tools to aid in that initial assessment.

Recent advancements in digital pathology and artificial intelligence (AI) have begun to show potential in enhancing diagnostic precision. Studies such as that by Foersch et al.¹³ have illustrated significant advancements in the performance of AI-assisted diagnoses in soft tissue sarcoma, emphasizing the progressive nature of this field. This underscores the need for ongoing research to refine AI applications in pathology, ensuring they are robust across various studies and datasets. Distinguishing between benign/reactive processes and malignant ones is a complex task in soft tissue pathology. The large number of entities involved, and the rarity of the setumors make it challenging to apply existing machine learning and AI models to this field.^{4,6,13}

The current investigation aims at exploring the application of AI in the classification of STT into benign and malignant entities, as a step towards the incorporation of AI into the clinical workflow. The goals of this study are: (1) identify cell-specific parameters that can aid in the classification of STTs as benign vs malignant, and (2) explore the capabilities of AI, in comparison to expert pathologists, in evaluating benign and malignant STT. By addressing these goals, the study aims to develop AI techniques that aid in accurately diagnosing STT.

I. METHODS

For all the experiments below, the ground truth was the original diagnosis obtained utilizing all glass slide, immunostains and molecular techniques, if necessary (Table 4).

Expert pathologist review(study arm 1)

Fifteen tissue pathologists from five different medical centers independently reviewed 100 WSI of hematoxylin and eosin (H&E)-stained slides representing 100 different tissue cases, following the institutional research protocol approved by the Institutional Review Board. Only one slide per case was provided with limited clinical information (patient age, gender, and anatomic location). Immunohistochemical and molecular information was not provided to the reviewing pathologists. Deidentified slides were scanned at $20\times$ magnification using an Aperio scanner (Aperio AT2, Leica Biosystems, Illinois), and acquired digital files were converted from .svs to a DICOM format. They were uploaded to a locally hosted, externally accessible compute node for review using a web-based WSI viewer system (Orthanc v1.3.2 WSI Plugin v0.5). Pathologists were asked to choose one of four diagnostic categories (benign, intermediate/border line, malignant, and uncertain) and to provide up to three differential diagnoses.

A RED Cap survey (<https://www.project-redcap.org/>) was used to record the answers.

Answers were compared to the original “ground-truth” diagnoses. Major and minor discordances, as well as uncertainty were defined as follows. Major discordance referred to discrepancies that could change patient management or prognosis, such as mistaking a benign tumor for a malignant one, or vice versa. Minor discordance involved differences that would not affect overall treatment, such as sub classification within the same category of benign or malignant tumors. Uncertainty was used for cases where a definitive diagnosis could not be reached due to insufficient information or ambiguous histological features.

Identifying cell specific parameters (study arm 2)

A cohort of 95 STT cases was utilized from the “expert pathologist review” experiment described above, including 60 benign and 35 malignant cases, encompassing 68 distinct STT entities. The cases were scanned at $40\times$ magnification using a high-throughput scanner (Aperio AT2) and uploaded to an OMERO server for annotation. The regions of interest (ROIs) were marked on all slides by the pathologist (Fig. 1). At the highest layer of resolution, the slide images were divided into 768-pixel tiles (Fig. 2). Each tile was evaluated for tissue percentage and color factors using pre-processing software developed as part of the Deep Histo Path project (<https://github.com/CODAIT/deep-histopath>). The meta data associated with tile and cell metrics was used to select the top 500 tiles per case within ROI. Cell detection was performed on each tile with at least 25% tissue using QuPath, providing 38 cell-specific parameters, which were averaged per tile. The cell-based metrics were averaged both across the top tiles per case and at the case-level. Welch’s t test was used, with $p < 0.05$ considered significant.

Employing an AI model to differentiate benign vs malignant soft tissue tumors(study arm 3)

A total of 195 soft tissue cases were collected from the files of the authors (SQ), encompassing the 95 cases from the “cell-specific parameters” experiment described above. These cases were divided into a training group (156) and validation group (39). A free software machine learning library for Python programming language (Scikit-learn) was employed to analyze these cases into benign and malignant STT (<https://scikit-learn.org/stable/>). Scikit-learn, an open-source machine learning library for Python programming language, offers a wide range of classification, regression, and clustering algorithms, including support-vector machines, random forests, and gradient boosting.

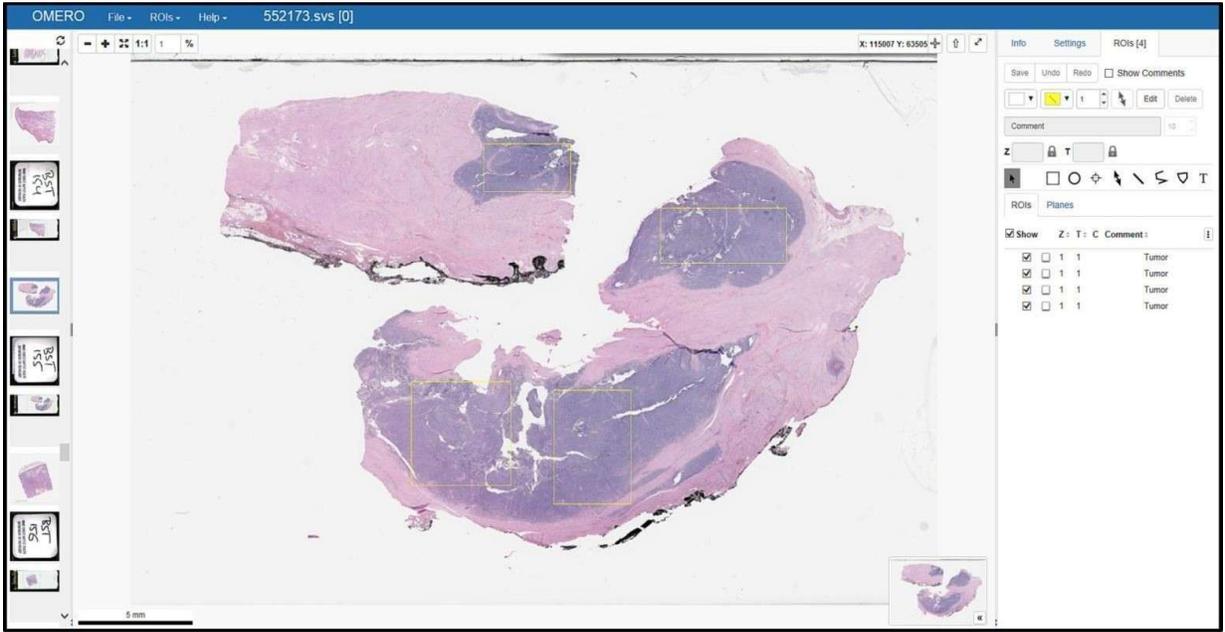


Figure 1: Representative whole-slide image uploaded to OMERO or annotation of ROI.

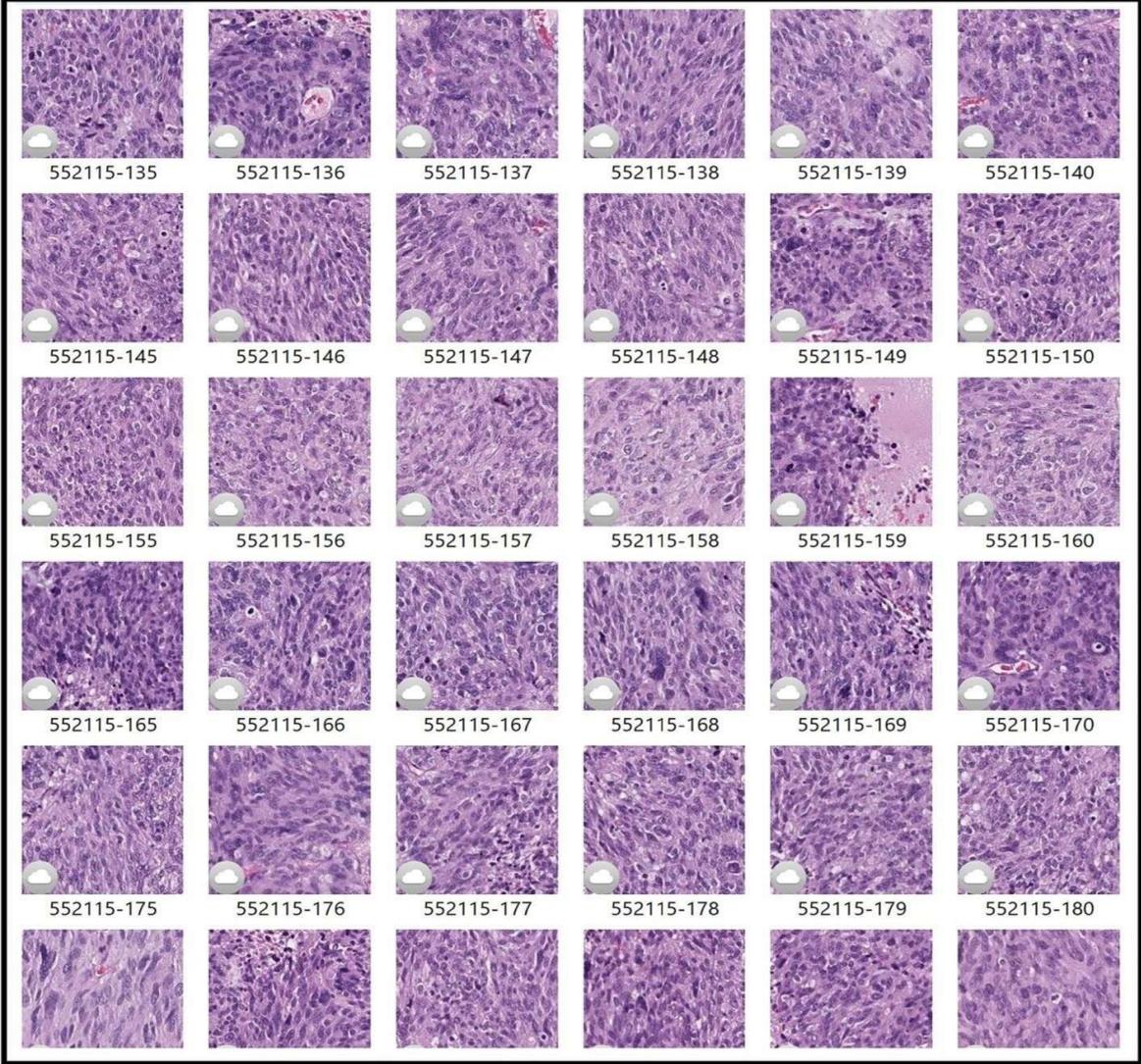


Figure 2: Representative case divided into 768-pixel tiles at the highest layer of resolution. 38 cell-specific parameters were detected using QuPath, which were averaged per tile.

I. RESULTS

Pathologist diagnostic concordance rate using WSI

The concordance rate of the pathologists when assessed against the reference diagnosis, made using traditional microscope, was 70.4% across the four diagnostic categories. In detail, minor discordances were observed in 1.6% of the cases, where the pathologists' diagnoses were close but not exactly the same as the reference. Major discordances, where in the diagnoses substantially differed, occurred in 5% of the cases. Additionally, there was a

13% rate of uncertainty where the pathologists could not reach a definitive diagnosis (Table 1). Analyzing the cases further, malignant tumors were most accurately diagnosed with an 81.7% concordance rate, whereas benign cases exhibited the highest rate of major discordance at 7.7%.

Intermediate cases showed the highest rates of minor discordance (28%) and uncertainty (22.4%) (Fig. 3). Overall, a correct differential diagnosis was established in 63% of the cases. Additionally, the pathologists reported either excellent or satisfactory quality of the scanned images for 96.4% of the cases.

Pathologists (P)	Overall concordance rate(%)	Major discordance (%)	Minor discordance (%)	Uncertain diagnosis (%)	Correct D X(%)
P1	70	5	10	15	66
P2	71	3	8	18	70
P3	70	4	9	17	63
P4	76	9	15	0	37
P5	65	4	16	15	79
Overall	70.4	5	11.6	13	63

Cell-specific parameters and their significance in benign vs malignant STT

Within the scope of 95 cases, encompassing 52 females and 43 males aged from 9 months to 90 years (with an average age of 41 and a median of 42), image analysis highlighted several cell-specific parameters that's how statistically significant differences ($p < 0.05$). These parameters included cell count, nucleus/cell are a ratio, nucleus she matoxylin.

Table 1 Summary of study findings (study arm 1). "Concordance" indicates agreement with the reference diagnosis with regards to the diagnostic categories (benign, intermediate / borderline, and malignant). Correct diagnosis is recorded none of the differentials provided by the pathologist matches the reference diagnosis (e.g., myxoid liposarcoma). Please note that this arm of the study was performed using one slide only, no ancillary testing, and limited clinical data.

Pathologists (P) cell max caliper, cell area, cell perimeter, cell circularity, and cell min caliper (Table 2). The Welch's t-test confirmed significant distinctions in the mean values of these parameters between benign and malignant groups, suggesting their potential utility in creating machine learning models for aiding with a soft tissue diagnosis.

AI model performance relative to expert pathologists

The machine learning models that achieved the best performance in distinguishing benign from malignant STTs included gradient boosting, neural network, xgboost, random forest, bagging, histgradient boosting, SGD classifier, and logistic regression. For instance, logistic regression exhibited a sensitivity of 0.737 and a specificity of 0.8, whereas random forest showed a sensitivity of 0.864 and a specificity of 0.75. The average sensitivity among these models was 0.60, specificity was 0.75, and overall accuracy was 0.68. The area under the receiver operating characteristic curve (AUC) for these models was also 0.68 (Table 3). These results indicate that the performance of the AI models is on par with that of the expert pathologists, suggesting a promising role for AI in supporting diagnostic processes in soft tissue pathology.

I. DISCUSSION

Accurate diagnosis of STT and their sub types is crucial in determining effective personalized oncology treatment plans for the best patient outcomes. There is scarcity of studies focusing on diagnostic discrepancies, especially in cases that have received second opinion in soft tissue pathology.^{1,16} One notable study examining the diagnosis of sarcoma through histopathology review revealed a substantial 24% discordance in diagnoses between community pathologists and an expert sarcoma reference pathology group. Sixty-six percent of these discordant cases had clinically significant implications for treatment recommendations.¹⁶ Interestingly, for all major discordant cases, excluding non-mesenchymal lesions, the diagnosis could have been made through conventional H&E-stained slides. The primary reason for diagnostic errors was the limited experience of non-specialized surgical pathologists with uncommon and atypical neoplasms.

Whole-slide imaging has emerged as a powerful tool for enhancing the care of cancer patients, accelerated by the COVID-19 pandemic and the CMS approval of remote sign out, thereby increasing the adoption of the technology.¹⁷ This technology enables timely assessment of tumor tissue by experts in STT, it fosters collaboration among specialists in sarcoma

management, and it provides pathologists in underserved regions, with out access to sarcoma centers, the opportunity to consult with a sarcomaspecialist.18Previous studies have also highlighted the utility of WSI technology in diagnosing STT. For instance, Sargenetal reported anoteable diagnostic accuracy of 89% using WSI for STT by two experienceds of ttisuepathologists.19Inanotherstudy, nine pathologists, with different levels of expertise, assessed 291 STT using WSI, and demonstrated a substantial increase in accuracy from 46.3% ($\pm 15.5\%$) to 87.1% ($\pm 11.1\%$) with the assistance of deep machine learning.13These findings emphasize the pivotal role of specialized pathology assessment in sarcoma diagnosis, and the need to leverage WSI for this purpose.

The expert review experiment in this project aimed to establish a baseline of what is achievable with limited clinical information and lack of ancillary studies, for comparison with an AI-assisted scenario using H&E- stained slides only. It is important to recognize that a comprehensive diagnostic evaluation necessitates the review of all case slides.

The bolded entities represent statistically significant values. along with access to clinical information, imaging data, and often additional tests like immunohistochemistry and molecular studies. 20 Inparticular, intermediate (borderline) lesions continue to pose challenges in classification, underscoring the need for supplementary tools in the diagnostic process.

This study identified several cell-specific parameters as being statistically significant($p < 0.05$)in distinguishing between benign and malignant STT. Recentre search has also shown the potential of nuclear morphology as a deep learning biomarker for cellular senescence, which can be applied to cancer.21In another study, deep learning algorithms significantly enhanced pathologists' accuracy in diagnosing leiomyos arcomas and predicting out-comes, increasing accuracy from 46.3% to 87.1%.8 These findings suggest that cell-specific parameters hold promise in aiding pathologists in distinguishing benign from malignant STT and hence improving diagnostic accuracy.

In this study, Scikit-learn22AIsoftware,was utilized, employing gradi- ent boosting, neural networks, xg boost, and logistic regression algorithms. The successful classification of benign and malignant STT, in our study,underscores the potential of an AI-based approach to enhance diagnostic accuracy in soft tissue pathology. A major limitation of this study is the relatively small number of cases analyzed and the inclusion of a large number of STT entities. Unfortunately, this is an expected challenge in studies involving STT, as the majority of these tumors are rare, and the differential diagnosis can be quite broad. In addition, the boundaries between benign, reactive, intermediate-,and low- grade malignancy can beblurry. Another limitation is the lack of validation using an independent dataset.

Despite the a for ementioned success yielded in our pilot study, it is crucial to acknowledge that the role of Alin STT diagnosis and treatment is still premature, and that further research is imperative e to validate its practical applicability in the clinical setting. Never the less, we believe that expanding the dataset with more cases has the potential to significantly improve the performance of the AI model. Finally, the implementation of AI in pathology mandate scareful consideration of various technical and ethical aspects, such as patient data privacy, data interpretability, model transparency, and potential biases.

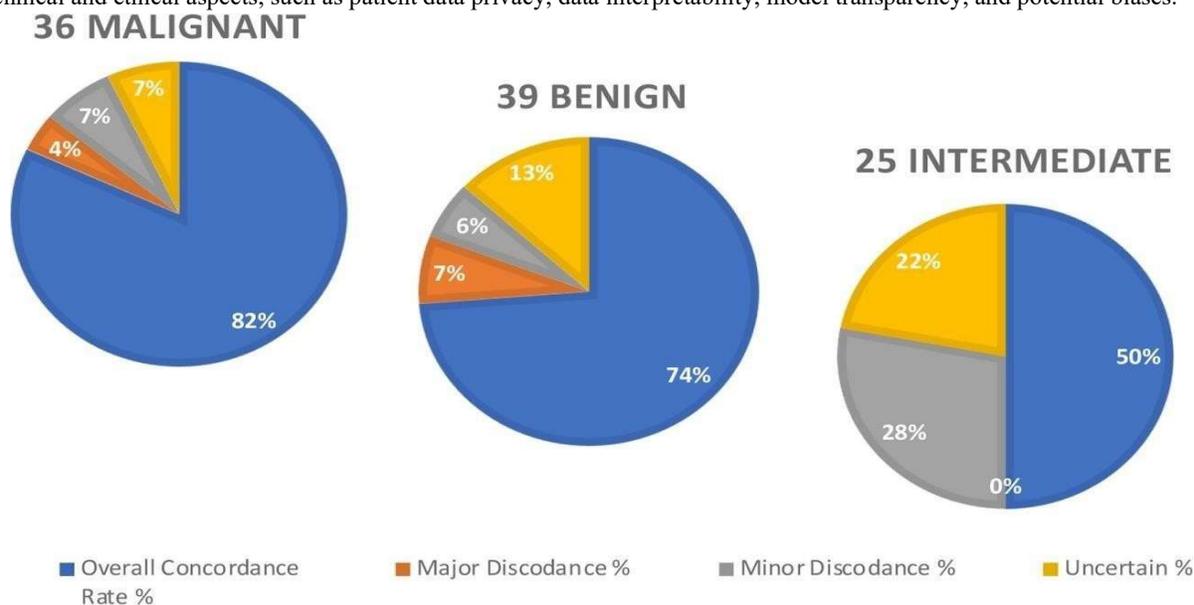


Figure 3: Expert pathologists' review(studyarm1).This was performed using one slide only,noanc illary testing,and limited clinical data.Malignant cases had the highest concordancerate(82%),benign cases had the highest major discordancerate(7%),and intermediate cases had the highest minor discordance(28%)and uncertaintyrates(22%).

Table2

Cell-specific parameters of soft tissue tumors and their *p* values (Study arm 2).

Cases(n)	Mean		Difference between means ± SEM	95% confidence interval	<i>p</i> value
	0	1			
	60	35			
Nucleus:Cellarearatio	0.218	0.2731	0.05504±0.008515	0.03804to 0.07204	<0.0001
Cellcount	171.9	278.9	107.0±18.51	70.08to143.9	<0.0001
Nucleus:area	28.41	29.7	1.287±0.7150	-0.1327to2.708	0.075
Nucleus:perimeter	23.16	23.21	0.04822±0.3044	-0.5563to0.6527	0.8745
Nucleus:Circularity	0.665	0.6748	0.009753±0.008050	-0.006236to0.02574	0.2288
Nucleus:Maxcaliper	8.984	8.748	-0.2359±0.1266	-0.4874to0.01556	0.0656
Nucleus:HematoxylinODmean	0.3127	0.3741	0.06141±0.02067	0.02032to 0.1025	0.0039
Cell: Maxcaliper	16.33	14.95	-1.381±0.2853	-1.949to-0.8130	<0.0001
Nucleus:Eccentricity	0.809	0.7885	-0.02050±0.005593	-0.03161to-0.009397	0.0004
Nucleus:Mincaliper	4.579	4.854	0.2756±0.06869	0.1391to0.4120	0.0001
Nucleus:HematoxylinODsum	36.71	47.15	10.44±2.362	5.738to15.15	<0.0001
Nucleus:HematoxylinODstddev	0.1034	0.1144	0.01096±0.004120	0.002775to 0.01914	0.0092
Nucleus:HematoxylinODmax	0.5784	0.6721	0.09370±0.02779	0.03846to 0.1489	0.0011
Nucleus:HematoxylinODmin	0.1047	0.1376	0.03292±0.01320	0.006608to 0.05923	0.0149
Nucleus:HematoxylinODrange	0.4736	0.5344	0.06078±0.01833	0.02436to 0.09719	0.0013
Nucleus: EosinOD mean	0.2425	0.2352	-0.007304±0.01187	-0.03094to0.01633	0.5401
Nucleus: EosinOD sum	27.54	28.55	1.012±1.403	-1.803to3.827	0.4738
Nucleus:EosinODstddev	0.06607	0.0649	-0.001174±0.003765	-0.008651to0.006303	0.7559
Nucleus: EosinOD max	0.3925	0.3862	-0.006242±0.01772	-0.04147to0.02899	0.7255
Nucleus: EosinOD min	0.08946	0.08085	-0.008608±0.006045	-0.02072to0.003505	0.1601
Nucleus: EosinOD range	0.303	0.3054	0.002360±0.01477	-0.02697to0.03169	0.8734
Cell:Area	135.1	110.7	-24.42±5.009	-34.38to-14.46	<0.0001
Cell:Perimeter	44.33	40.5	-3.831±0.7991	-5.422to-2.240	<0.0001
Cell:Circularity	0.8166	0.7968	-0.01988±0.004069	-0.02796to-0.01180	<0.0001
Cell: Mincaliper	10.88	9.898	-0.9780±0.2126	-1.401to-0.5548	<0.0001
Cell:Eccentricity	0.6993	0.7058	0.006489±0.003267	1.171e-006to0.01298	0.05
Cell:EosinODmean	0.179	0.1718	-0.007139±0.009776	-0.02658to0.01230	0.4672
Cell:Eosinstddev	0.08495	0.08461	-0.0003421±0.005141	-0.01056to0.009875	0.9471
Cell:EosinODmax	0.4168	0.4017	-0.01506±0.01941	-0.05364to0.02351	0.4397
Cell:EosinODmin	0.01473	0.007824	-0.006906±0.002861	-0.01259to-0.001224	0.0178
Cytoplasm:HematoxylinODmean	0.08647	0.122	0.03549±0.009522	0.01644to 0.05454	0.0004
Cytoplasm:HematoxylinODstddev	0.06128	0.07984	0.01855±0.004686	0.00922to 0.02789	0.0002
Cytoplasm:HematoxylinODmax	0.3381	0.4205	0.08239±0.02081	0.04094to 0.1238	0.0002
Cytoplasm:HematoxylinODmin	-0.01809	-0.00513	0.01296±0.004507	0.003989to 0.02194	0.0052
Cytoplasm:EosinODmean	0.1613	0.1477	-0.01362±0.009514	-0.03254to0.005288	0.1558
Cytoplasm:EosinODstddev	0.07641	0.074	-0.002417±0.004904	-0.01216to0.007325	0.6233
Cytoplasm:EosinODmax	0.3729	0.35	-0.02283±0.01892	-0.06043to0.01477	0.2309
Cytoplasm:EosinODmin	0.01692	0.009997	-0.006924±0.002960	-0.01280to-0.001046	0.0215

Table3

Representative AI models metrics, classifying benign and malignant soft tissue tumors (Study arm 3).

Metric	AUC	Accuracy	Sensitivity	Specificity
Gradientboosting	0.664	0.667	0.579	0.75
Neuralnetwork	0.638	0.641	0.526	0.75
Xgboost	0.639	0.641	0.579	0.7
Randomforest	0.717	0.718	0.684	0.75
Bagging	0.666	0.667	0.632	0.7
Tabpfn	0.743	0.744	0.737	0.75
Histgradientboosting	0.664	0.667	0.579	0.75
Sgdclassifier	0.584	0.59	0.368	0.8
Logisticregression	0.768	0.769	0.737	0.8

Table4

List of cases used for all study arms (Study arms 1, 2, and 3). Cases with* are not considered soft tissue tumors but have either presented in a location and/or showed morphological overlap with soft tissue tumors warranting them being submitted to the soft tissue service for consultation.

Case	Age/Sex	Location	Diagnosis
1	45F	Esophagus	Ewing/PNET

2	88F	Chestwall	Elastofibroma
3	3mM	Flank	Kaposiformhemangioendothelioma
5	8mM	Thigh	Juvenilexanthogranuloma
6	63F	Vulva	Angiomyofibroblastoma
7	53M	Leg	Cutaneousleiomyosarcoma
8	54M	Intestinal	MalignantGIST
9	26M	Rectum	*Ballooncellmelanoma
11	31M	Thumb	Perineurioma
12	49M	Mediastinal	*Type-A thymoma
14	34F	Retroperitoneal	Ewing/PNET
15	76M	Elbow	*Latestageerythemaelevatumdiutinum
17	42M	Arm	Nodularfasciitis
20	78M	Shoulder	Ischemicfasciitis
21	95F	Intra-abdominal	*Granulosacelltumor
22	15F	Subcutaneous	Malignantgiant cell tumorof soft parts
26	42M	Hip	Schwannoma
27	45F	Arm	MPNST arisingin aneurofibroma
28	36M	Ankle	Clearcellsarcoma
29	2F	Thigh	Sclerosingrhabdomyosarcoma
31	57F	Vagina	Benigngenitalstromalpolyp
34	50F	Thigh	Biphasicsynovialsarcoma,grade2
35	55M	Shoulder	Pleomorphicrhabdomyosarcoma,highgrade
36	66F	Scapula	HPC/SFTwithmalignant potential
38	79F	Groin	*Metastatic melanoma
39	13M	Leg	DFSP
40	47M	Retroperitoneal	Schwannoma
42	7F	Thigh	Pleomorphicsarcomawithgiantcells(malignantgiantcelltumorofsoftparts)
45	8M	Shoulder	Granularcell tumor
46	62M	Nose	Welldifferentiatedfibrosarcoma(grade1)
48	29M	Preauricular	solitaryfibroustumor/hemangiopericytoma
49	16M	Tongue	Granularcell tumor
51	19M	Intra-ventricular	Hemangioma,cavernous/capillarytype
52	59F	Pelvis	*Highgradecarcinosarcoma
53	49F	Uterus	Epithelioidleiomyoma
56	24M	Trunk	Dermatofibrosarcomaprotuberans
58	38F	Foot	MarkedStasischanges
60	42F	Uterineserosa	Mesenchymaltumor,favorunusualsmoothmuscleumorofuncertainmalignantpotential
63	78M	Arm	Kaposisarcoma
68	66M	Abdomen	*Diffuse follicle center lymphoma
69	8F	Smallbowel	Reactivechanges
70	39F	Forearm	Sarcomawithmyofibroblasticfeatures
74	55F	Axillary	*MalignantS100positivetumor, favor melanoma
77	23F	Retroperitoneal	Angiolymphoidhyperplasia(witheosinophilia)
78	59M	Back	*Adnexalneoplasmofatleastlowgrademalignancy
79	52M	Arm	Myxoid variant ofhemangiopericytoma/SFT
80	66M	Parotid	Pleomorphicsarcomawithmyoiddifferentiation
81	42F	Arm	Epithelioidsarcoma
82	25M	Periumbilical	Myoepithelialtumorofsofttissue,histologicallybenign
84	69F	Omentum	*Lowgradeendometrialstromalsarcoma
86	42M	Mesenteric	Follicularendriticcelltumor
87	52F	Breast	Histiocytictumorofuncertainmalignantpotential
90	1F	Abdominalwall	Juvenilexanthogranuloma
92	6F	Finger	Cellularjuvenileaponeuroticfibroma
93	84M	Temple	*Ulceratingcarcinoma
94	20F	Chestwall	Monophasicsynovialsarcoma,highgrade
95	5F	Rfoot	Fibroushistiocytoma
97	41F	Rtongue	Myofibroma
99	67M	Toe	Lowgradepleomorphicsarcoma
103	8M	Hand	*Cellularbluenevus
104	30M	Shoulder	*Softtissuechordomawithboneerosion
105	42F	Lowerabdomen	*Malignanttumor/wmyoepithelial carcinoma, highgrade
107	1M	Rchest	Infantilefibromatosis
108	35M	Spine(T1)	Metastaticmalignantperipheralnervesheathumor(MPNST),epithelioidtype
109	43M	Flank	Angiosarcoma,highgrade
120	41F	Breast	Leiomyoma
122	39M	Stomach	Malignantgastrointestinalstromaltumor(GIST)
124	41M	Rectum	Gastrointestinalstromaltumor
127	47M	Subcutaneous	Lowgradefibromyxoidsarcoma
129	12M	Toe	Benignmesenchymaltumorwithfeaturesofangiomatosisandmyofibromatosis

130	67F	Vulva	Benigngenitalstromaltumor
135	43F	Wrist	Pleomorphicundifferentiatedsarcoma,highgrade
139	42F	Arm	ExtranodalRosaiDorfmandisease
141	75M	Sinonasal	Hemangiopericytoma-liketumorofthenasalpassages

Table4(continued)

Case	Age/Sex	Location	Diagnosis
142	45M	Kidney	Malignantglomus tumor
143	16M	Knee	Necrobioticgranuloma
145	51M	Neck	Benignmesenchymaltumor,favorspindlecelllipoma
148	85M	Scalp	*Poorlydifferentiatedcarcinoma,probablymetastatic
151	45M	Retroperitoneal	Monophasicsynovialsarcoma,highgrade
153	47M	Upperleg	Dermatofibrosarcomaprotuberans
157	51F	Mediastinal	Myxoid/roundcellliposarcoma,high grade
158	64F	Toe	Ewingsarcoma
160	30F	Paraspinal	Sclerosingepithelioidfibrosarcoma
164	14M	L4vertebra	Langerhanscellhistiocytosis
165	1M	Hip	Calcifyingaponeuroticfibroma
167	31M	Scalp	Alveolarsoftpartsarcoma,Ruleoutmetastasis
168	5M	Tongue	Reactivemyofibroblasticproliferation
169	58F	Sternocleidomastoidmuscle	Softtissuemyoepithelioma,histologicallybenign
171	26M	Omentum	Benignfibroblasticproliferation,favorreactive
174	71F	Neck	Malignanthemangiopericytoma/solitaryfibroustumor,highgrade
176	92M	Neck	*Desmoplasticismelanoma
177	33F	Calf	Neuroblastoma-likeschwannoma(schwannomawithcollagenrosettes)
178	66M	Patella	Glomus tumor
179	32F	Knee	Angiomatoidfibroushistocytoma
180	39M	Shoulder	Kaposisarcoma
184	59F	Orbit	Hemangiopericytoma
186	18M	Scrotum	Embryonalrhabdomyosarcoma
188	62F	Leg	Fibroushistocytomawithatypical(monster)cells
191	34F	Calf	*Paraganglioma-likedermalmelanocytictumor
193	47F	Uterineserosa	Leiomyoma
195	3F	labiamajora	Lipoblastoma
196	19F	Brachialplexus	Epithelioidnervesheathtumor,probablyoflowgrademalignancy
198	60F	Mesenteric/smallbowel	Bacillaryangiomatosis
199	44F	Calf	Extraskelatalmyxoidchondrosarcoma
200	50F	Anus	*Malignantmelanoma

II. CONCLUSION

This study demonstrates the potential of AI techniques to enhance the diagnosis and classification of STTs, with promising results in distinguishing benign from malignant cases and highlighting the relevance of cell-specific parameters. A comprehensive diagnostic evaluation, encompassing all slides and access to immuno histochemical and molecular studies, remains indispensable for accurate soft tissue diagnosis. The use of large heterogeneous, well-curated and annotated/labelled datasets will be essential to bolster AI model training and accuracy in the field of soft tissue pathology. Nevertheless, the integration of AI into the diagnostic process offers substantial promise, poised to elevate our capacity to differentiate between these tumors, ultimately leading to heightened diagnostic precision and improved patient outcomes.

We would like to emphasize that this is a proof-of-concept small study, and larger studies are needed to validate and

REFERENCES

- Arbiser ZK, Folpe AL, Weiss SW. Consultative (expert) second opinions in soft tissue pathology. Analysis of problem prone diagnostic situations. *Am J Clin Pathol* 2001; 116(4):473–476. <https://doi.org/10.1309/425H-NW4W-XC9A-005H>.
- Rupani A, Hallin M, Jones RL, Fisher C, Thway K. Diagnostic differences in expert second opinion consultation cases at a tertiary sarcoma center. *Sarcoma* 2020; 2020:9810170. <https://doi.org/10.1155/2020/9810170>.
- Mesko NW, Mesko JL, Gaffney LM, Halpern JL, Schwartz HS, Holt GE. Medical malpractice and sarcoma care: a thirty-three year review of case resolutions, inciting factors, and at risk physician specialties surrounding a rare diagnosis. *J Surg Oncol* 2014; 110(8):919–929. <https://doi.org/10.1002/jso.23770>.
- Ailia MJ, Thakur N, Abdul-Ghaffar J, Jung CK, Yim K, Chong Y. Current trend of artificial intelligence patents in digital pathology: a systematic valuation of the patent landscape. *Cancers (Basel)* 2022; 14(10). <https://doi.org/10.3390/cancers14102400>.
- Borowsky AD, Glassy EF, Wallace WD, et al. Digital whole slide imaging compared with light microscopy for primary diagnosis in surgical pathology. *Arch Pathol Lab Med* 2020; 144(10):1245–1253. <https://doi.org/10.5858/arpa.2019-0569-OA>.
- Cromb a, Rouleau-Dugac M, Italiano A. The diagnosis, classification, and treatment of sarcoma in this era of artificial intelligence and immunotherapy. *Cancer Commun (Lond)* 2022; 42(12):1288–1313. <https://doi.org/10.1002/cac2.12373>.
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011; 144(5):646–674. <https://doi.org/10.1016/j.cell.2011.02.013>.
- Kantidakis G, Liti re S, Neven A, et al. New benchmarks to design clinical trials with advanced or metastatic liposarcoma or synovial sarcoma patients: an EORTC – Soft Tissue and Bone Sarcoma Group (STBSG) meta-analysis based on a literature review for soft-tissue sarcomas. *Eur J Cancer* 2022; 174:261–276. <https://doi.org/10.1016/j.ejca.2022.07.010>.
- Kusta O, Rift CV, Risør T, Santoni Rugiu E, Brodersen JB. Lost in digitization – a systematic review about the diagnostic test accuracy of digital pathology solutions. *J Pathol Inform* 2022; 13, 100136. <https://doi.org/10.1016/j.jpi.2022.100136>.
- Mukhopadhyay S, Feldman MD, Abels E, et al. Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized noninferiority study of 1992 cases (pivotal study). *Am J Surg Pathol* 2018; 42(1):39–52. <https://doi.org/10.1097/PAS.0000000000000948>.

11. Saini T, Bansal B, Dey P. Digital cytology: current status and future prospects. *DiagnCytopathol*2023;51(3):211-218.<https://doi.org/10.1002/dc.25099>.

Rottmann D, Abdulfatah E, Pantanowitz L. Molecular testing of soft tissue expanders on the findings. The current study provides a foundation upon which future studies can be built. Future research should focus on expanding the dataset and refining AI algorithms to potentially improve diagnostic sensitivity and specificity. Additionally, ongoing efforts must be made to address technical and ethical considerations, such as ensuring data privacy and the accuracy of AI-generated results.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGEMENT

The authors would like to acknowledge the generous support of the Sanders-Brown Center at the University of Kentucky through their staff and equipment for whole slide imaging

- tumors. *DiagnCytopathol*2023;51(1):12–25.<https://doi.org/10.1002/dc.25013>.
12. Foersch S, Eckstein M, Wagner DC, et al. Deep learning for diagnosis and survival prediction in soft tissue sarcoma. *Ann Oncol* 2021;32(9):1178–1187. <https://doi.org/10.1016/j.annonc.2021.06.007>.
13. Allan C, Burel JM, Moore J, et al. OMERO: flexible, model-driven data management for experimental biology. *Nat Methods* 2012;9(3):245–253. <https://doi.org/10.1038/nmeth.1896>.
14. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: open source software for digital pathology image analysis. *Sci Rep* 2017;7(1):16878. <https://doi.org/10.1038/s41598-017-17204-5>.
15. Al-Ibraheemi A, Folpe AL. Voluntary second opinions in pediatric bone and soft tissue pathology: a retrospective review of 1601 cases from a single mesenchymal tumor consultation service. *Int J Surg Pathol* 2016;24(8):685–691. <https://doi.org/10.1177/1066896916657591>.
16. Clinical Laboratory Improvement Amendments of 1988 (CLIA) Post - Public Health Emergency (PHE) Guidance CMS, Ed. 2021.
18. Rajasekaran RB, Whitwell D, Cosker TDA, Gibbons CLMH, Carr A. Will virtual multidisciplinary team meetings become the norm for musculoskeletal oncology care following the COVID-19 pandemic? - experience from a tertiary sarcoma centre. *BMC Musculoskelet Disord* 2021;22(1):18. <https://doi.org/10.1186/s12891-020-03925-8>
19. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: open source software for digital pathology image analysis. *Sci Rep* 2017;7(1):16878. <https://doi.org/10.1038/s41598-017-17204-5>.
20. Al-Ibraheemi A, Folpe AL. Voluntary second opinions in pediatric bone and soft tissue pathology: a retrospective review of 1601 cases from a single mesenchymal tumor consultation service. *Int J Surg Pathol* 2016;24(8):685–691. <https://doi.org/10.1177/1066896916657591>.
21. Clinical Laboratory Improvement Amendments of 1988 (CLIA) Post - Public Health Emergency (PHE) Guidance CMS, Ed. 2021.
22. Rajasekaran RB, Whitwell D, Cosker TDA, Gibbons CLMH, Carr A. Will virtual multidisciplinary team meetings become the norm for musculoskeletal oncology care following the COVID-19 pandemic? - experience from a tertiary sarcoma centre. *BMC Musculoskelet Disord* 2021;22(1):18. <https://doi.org/10.1186/s12891-020-03925-8>
23. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: open source software for digital pathology image analysis. *Sci Rep* 2017;7(1):16878. <https://doi.org/10.1038/s41598-017-17204-5>.
24. Al-Ibraheemi A, Folpe AL. Voluntary second opinions in pediatric bone and soft tissue pathology: a retrospective review of 1601 cases from a single mesenchymal tumor consultation service. *Int J Surg Pathol* 2016;24(8):685–691. <https://doi.org/10.1177/1066896916657591>.
25. Clinical Laboratory Improvement Amendments of 1988 (CLIA) Post - Public Health Emergency (PHE) Guidance CMS, Ed. 2021.
- 17.

A STUDY ON ARTIFICIAL INTELLIGENCE TO OVERCOME DISTRIBUTION CHALLENGES IN RURAL MARKETS

*Note: Sub-titles are not captured in Xplore and should not be used

<p>Charan.R Student, II - MBA, Department of Management Studies, Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, (Deemed to be University), Chennai, India. vtp4263@veltech.edu.in</p>	<p>DR. K. Ravi shankar Professor, MBA, Department of Management Studies, Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, (Deemed to be University), Chennai, India. hodmba@veltech.edu.in</p>
--	---

Abstract—In the modern business environment, rural markets are expanding rapidly and facing challenges that were not experienced previously. The introduction of Artificial Intelligence (AI) in the distribution sector is transforming the way companies manage logistics, supply chains, and rural outreach. The main purpose of this study is to examine the alternative uses and effects of integrating AI to overcome distribution challenges in rural areas through automated routing, demand prediction, inventory optimization, and data-driven decision-making in supply chain management.

AI-based tools including predictive analytics, smart logistics platforms, drone delivery systems, and intelligent tracking devices are enabling organizations to reduce transportation delays, minimize stockouts, improve last-mile connectivity, and enhance the accuracy of delivery schedules in rural markets that traditionally suffered from poor infrastructure, limited connectivity, and high operational costs. The research paper seeks to assess how AI applications can facilitate a more efficient distribution network, improve rural service quality, and strengthen market penetration by examining practical applications of AI in rural supply chains.

The study will also examine operational challenges such as ethical considerations in data usage, resistance to technological adoption, job displacement risks for manual logistics workers, and the need for digital upskilling among rural distributors and delivery partners. Data for the study will utilize recent literature and case analyses of companies that have adopted AI-driven logistics and rural distribution systems.

The findings of the research will affirm that when supported with appropriate implementation strategies, AI can be embraced as a strategic tool for overcoming rural distribution barriers and creating a more connected, efficient, and responsive supply network. The conclusion of the study emphasizes that achieving sustainable rural distribution through AI is best accomplished by harmonizing technology-driven logistics with human-centred management practices that ensure inclusiveness, adaptability, and long-term rural development.

Keywords— *Artificial Intelligence, Rural Distribution, Supply Chain Optimization, Predictive Analytics, Last-Mile Delivery, Logistics Challenges, Inventory Management, Smart Routing, Digital Upskilling, Market Connectivity.*

I. INTRODUCTION

Artificial Intelligence (AI) has emerged as a powerful technological advancement capable of transforming traditional business operations, particularly in complex and underserved environments like rural markets. Rural regions across developing countries often struggle with persistent distribution challenges such as inadequate transportation infrastructure, fragmented supply chains, poor connectivity, unpredictable demand patterns, and higher operational costs. These barriers reduce market accessibility, delay product delivery, and limit the efficiency of businesses attempting to expand their rural presence. In recent years, AI-driven tools and systems have begun to offer innovative pathways to overcome these obstacles by enhancing accuracy, speed, and coordination in distribution processes.

AI applications such as predictive analytics, real-time tracking, automated routing, demand forecasting, and intelligent inventory management allow companies to optimize logistics even in geographically dispersed rural areas. These technologies help organizations anticipate demand fluctuations, reduce stockouts, streamline last-mile delivery, and improve overall supply chain transparency. Moreover, AI-powered platforms support better decision-making by analyzing large volumes of data, enabling firms to design more reliable and cost-effective distribution strategies.

In addition, AI facilitates smoother interactions between manufacturers, distributors, and rural retailers by enabling digital customer satisfaction. While AI offers significant opportunities, its implementation also poses challenges such as technological barriers, digital literacy gaps, ethical issues, and the need for capacity building among rural stakeholders. communication and monitoring systems. This leads to improved product availability, reduced wastage, and enhanced rural

II. OBJECTIVES OF THE RESEARCH

Primary Objectives:

1. To examine how AI solves major rural distribution challenges.
2. To analyze AI's impact on last-mile delivery and supply chain efficiency.
3. To assess how AI improves inventory management in rural markets.

Secondary Objectives:

1. To study the role of predictive analytics, smart routing, and tracking tools.
2. To identify barriers in adopting AI in rural areas.
3. To explore the need for digital skills and infrastructure for AI use.
4. To provide suggestions for effective AI integration in rural distribution.

III. REVIEW OF LITERATURE

1. AI as a Transformative Tool in Rural Distribution

The literature widely recognizes Artificial Intelligence (AI) as a transformative technology reshaping supply chain operations, especially in rural markets where traditional distribution systems face consistent difficulties. AI is highlighted as a strategic tool capable of addressing these constraints through automation and data-driven insights that improve decision-making and operational efficiency.

2. Predictive Analytics and Demand Forecasting

Several research studies focus on the growing role of AI-driven predictive analytics in forecasting demand in rural areas. Literature states that rural markets often experience irregular demand patterns due to seasonal, cultural, and economic variations. AI helps businesses analyze historical data, purchasing trends, and external factors to generate accurate forecasts. This reduces stockouts, minimizes wastage, and enhances product availability in remote locations.

3. AI in Routing and Last-Mile Delivery Optimization

A major theme in existing research is the application of AI in optimizing last-mile delivery—one of the biggest challenges in rural distribution. Scholars point out that machine learning algorithms can process real-time data such as road conditions, weather patterns, and travel time to suggest efficient routes. Literature concludes that AI-enabled routing systems reduce fuel consumption, shorten delivery times, and improve service reliability in rural territories where transportation barriers are common.

4. Integration of AI with IoT for Supply Chain Visibility

The literature highlights the importance of integrating AI with Internet of Things (IoT) technology to enhance transparency and visibility across rural supply chains. IoT sensors and smart tracking devices allow companies to monitor inventory levels, vehicle status, and delivery progress. AI analyzes this data to prevent delays, optimize inventory, and improve operational coordination among manufacturers, distributors, and rural retailers.

5. Challenges in AI Adoption for Rural Markets

Despite its advantages, literature consistently identifies barriers that hinder AI adoption in rural distribution. Issues such as poor internet connectivity, lack of digital skills, high implementation costs, and resistance to technological change are frequently mentioned. Researchers argue that without proper infrastructure and upskilling initiatives, rural markets may struggle to fully benefit from AI-driven solutions.

6. Identified Research Gap

Overall, the review of literature shows strong support for AI as a solution to distribution challenges in rural markets. However, scholars note a lack of comprehensive empirical studies focused exclusively on AI applications in rural

supply chains. This gap highlights the need for further research

Scope of the study:

The scope of this study is limited to examining how Artificial Intelligence (AI) can be used to overcome distribution challenges in rural markets. It focuses on AI applications such as demand forecasting, smart routing, real-time tracking, and inventory management to improve the efficiency of supply chains in rural areas. The study covers the impact of AI on last-mile delivery, transportation delays, product availability, and operational coordination between manufacturers, distributors, and rural retailers. It also includes an analysis of key challenges such as poor infrastructure, low digital literacy, cost of technology, and limited connectivity that may affect AI implementation. The study is based entirely on secondary data from research articles, reports, and case studies. Overall, it aims to define the potential, limitations, and practical significance of AI in strengthening rural distribution networks.

Limitation of the study:

Dependence on Secondary Data:

The study is based entirely on existing literature, reports, and case studies, which may limit the accuracy and depth of real-time insights.

Lack of Primary Field Data:

Direct surveys or interviews with rural distributors, retailers, or logistics managers were not conducted, reducing firsthand practical validation.

Technology Variation Across Regions:

AI adoption differs widely across rural areas, so the findings may not apply uniformly to all regions or industries.

Limited Access to Company Data:

Confidential internal data on AI-based distribution systems from private organizations was not available for analysis.

Socio-Cultural Factors Not Fully Addressed:

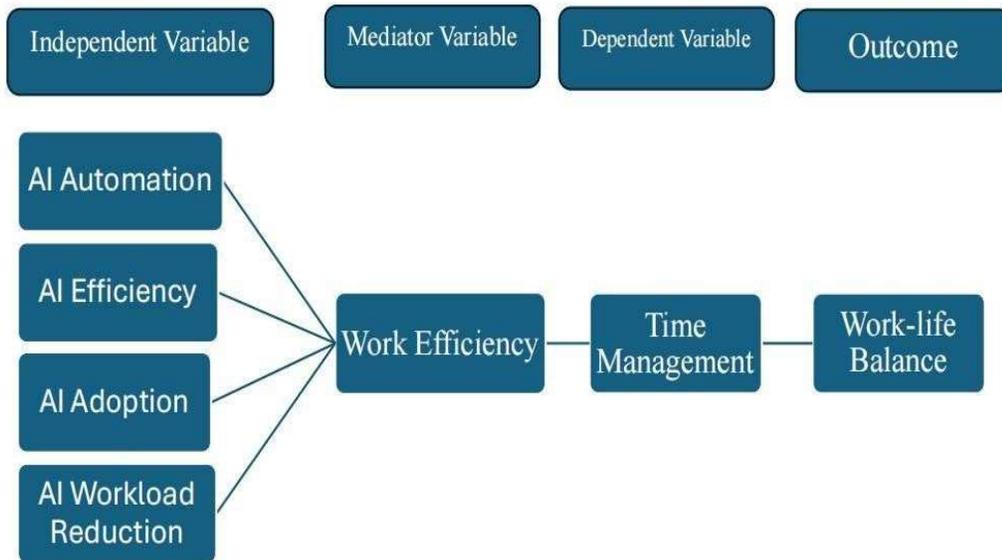
Resistance to technology, digital illiteracy, and local behavioural factors in rural communities are difficult to measure using secondary data alone.

I. RESEARCH METHODOLOGY

This study follows a descriptive research methodology to understand how Artificial Intelligence can overcome distribution challenges in rural markets. It is based entirely on secondary data collected from journals, research articles, industry reports, and case studies related to AI and rural logistics. The information gathered is analyzed using qualitative and thematic analysis to identify key AI applications, benefits, and challenges in rural distribution systems. No primary data such as surveys or interviews is included, as the study focuses on reviewing existing literature to draw meaningful insights and conclusions.

Research Design

The study adopts a descriptive research design to clearly explain how Artificial Intelligence (AI) can address and overcome distribution challenges in rural markets. This design is used to describe the existing conditions, patterns, and practices related to AI-based logistics without manipulating any variables. It focuses on observing and analyzing how AI tools such as predictive analytics, route optimization, and real-time tracking are currently applied in rural supply chains. The descriptive design also helps in identifying challenges, gaps, and opportunities by reviewing existing literature and real-world case studies. Overall, this research design provides a structured and detailed understanding of AI's role in improving rural distribution efficiency.



Research Model:

Hypothesis

Null Hypothesis, H₀:

There is no significant impact of Artificial Intelligence (AI) in overcoming distribution challenges in rural markets.

Alternative Hypothesis, H₁:

There is a significant impact of Artificial Intelligence (AI) in overcoming distribution challenges in rural markets.

Sampling Method

Purposive sampling, also known as judgmental or selective sampling, is used in this research. This non-probability sampling technique allows the researcher to collect responses from participants who have specific knowledge, experience, or involvement in rural distribution, supply chain operations, or AI-enabled logistics systems. Only respondents who are directly connected with rural markets or distribution activities were selected to ensure relevant and accurate insights.

Sample Size

The sample size for this research study is 97 respondents, consisting of rural distributors, retailers, logistics personnel, supply chain supervisors, and delivery partners who are familiar with the implementation or challenges of AI in rural distribution networks.

Chi-Square:

Hypothesis

Null Hypothesis, H₀:

There is no significant relationship between age group and AI adoption in rural distribution activities.

Alternative Hypothesis, H₁:

There is a significant relationship between age group and AI adoption in rural distribution activities.

Age Group × AI tools have improved distribution efficiency in rural markets

Age Group	Agree	Neutral	Strongly Agree	Total (Observed)	Expected Agree	Expected Neutral	Expected Strongly Agree
Below 25	10	4	1	15	9.5	4.4	1.1
26–35	28	10	2	40	25.3	11.6	3.0

36–45	6	4	2	12	7.6	3.5	0.9
46–55	6	5	1	12	7.6	3.5	0.9
Total	50	23	6	79	50.0	23.0	6.0

Interpretation:

Younger respondents, especially those aged 26–35, show the highest agreement that AI tools improve distribution efficiency in rural markets. Below-25 participants also show positive acceptance, though at a moderate level. Older age groups (36–45 and 46–55) display more neutral responses, indicating lower familiarity or comfort with AI. Overall, AI adoption varies by age, with younger groups showing stronger perceived benefits.

Chi-Square Tests

Statistic	Value
Chi-Square (χ^2)	3.82
Degrees of Freedom (df)	6
p-value	0.701

Interpretation:

The Chi-Square value ($\chi^2 = 3.82$) with 6 degrees of freedom and a p-value of 0.701 indicates that the result is not statistically significant. Since the p-value is greater than 0.05, we fail to reject the null hypothesis. This means there is no significant relationship between age group and the adoption of AI tools for improving distribution efficiency in rural markets. In simple terms, age does not strongly influence how respondents perceive the impact of AI on rural distribution.

One-Way ANOVA

Hypothesis

Null Hypothesis (H₀):

There is no significant difference in the mean AI adoption level among different age groups involved in rural distribution.

Alternative Hypothesis (H₁):

There is a significant difference in the mean AI adoption level among different age groups involved in rural distribution.

ANOVA

Model	Sum of Squares	df	Mean Square	F-Value	Sig. (p-value)
Regression	48.712	6	9.742	49.918	0.000
Residual	18.288	91	0.201		
Total	67.000	97			

Interpretation:

The ANOVA results show that the regression model is statistically significant, $F(6, 91) = 49.918$, $p = 0.000$. Since the p-value is less than 0.05, the model indicates a strong and significant relationship between the independent variables (AI-related factors) and the dependent variable (distribution efficiency in rural markets). This means AI factors collectively contribute significantly to improving rural distribution performance..

I. MAJOR FINDINGS

The research concluded that AI significantly enhances the overall efficiency and effectiveness of distribution systems in rural markets. AI-enabled tools contribute to faster decision-making, improved route planning, and smarter demand forecasting, which collectively reduce delays and ensure timely product availability even in remote locations.

Organizations utilizing AI-powered distribution solutions observed reduced operational workload and fewer manual errors, as repetitive tasks such as tracking, inventory monitoring, and scheduling were automated. This has allowed distribution teams to focus more on strategic, value-adding activities such as retailer relationship building and local market development.

The study found that service reliability in rural markets improved as AI provided real-time insights into stock levels, transportation challenges, and customer demand patterns. This increased transparency and accuracy supported better coordination between warehouses, transporters, and rural retailers.

AI adoption also contributed to cost savings by optimizing transportation routes, reducing fuel consumption, and minimizing unnecessary trips. Companies reported enhanced supply chain visibility, which improved planning and reduced the frequency of stock-outs in rural outlets.

The findings also revealed differences across regions and departments regarding the extent of AI benefits. Areas with higher levels of AI integration reported stronger improvements in delivery consistency, inventory accuracy, and overall distribution performance. In contrast, some regions continue to adapt to new technologies due to limited digital literacy or infrastructural constraints.

II. RESEARCH SUGGESTIONS

Expand AI Infrastructure in Rural Areas:

Future research should explore models for strengthening digital and technological infrastructure in rural regions to support seamless AI adoption.

Conduct Region-Specific AI Studies:

Further studies can focus on how AI performs across different rural states or districts, considering variations in geography, connectivity, and market behavior.

Assess Retailer Readiness and Training Needs:

Additional research should examine the digital literacy levels of rural retailers and recommend AI-driven training programs to improve adoption.

Evaluate Cost-Effectiveness of AI Tools:

Future work should analyze different AI solutions to determine which tools offer the most affordable and practical benefits for rural distribution networks.

Study Consumer Behavior With AI-Driven Distribution:

Research can assess how improved availability, reduced stock-outs, and faster delivery influence rural customer satisfaction and purchasing patterns.

Investigate Ethical and Data Privacy Concerns:

Studies should focus on the challenges of data collection in rural markets and propose frameworks for safe, transparent, and ethical use of AI.

Explore Human-AI Collaboration Models:

Future research could examine how distribution employees and local agents can work alongside AI systems for improved efficiency.

Test AI Adoption in Different Product Categories:

Research may analyze how AI-driven distribution works for FMCG, pharmaceuticals, agricultural products, and durable goods in rural markets.

III. CONCLUSION

The study concludes that Artificial Intelligence has emerged as a transformative tool in addressing long-standing distribution challenges in rural markets. By enabling accurate demand forecasting, efficient route optimization, real-time inventory tracking, and improved supply chain visibility, AI significantly enhances the reliability and responsiveness of rural distribution systems. These advancements help companies reduce delays, minimize stock-outs, lower operational costs, and improve product availability in remote locations.

AI-driven insights also support better decision-making and strengthen coordination among warehouses, transporters, and rural retailers. As a result, organizations are better equipped to navigate geographical barriers, inconsistent infrastructure, and fluctuating demand patterns that commonly affect rural markets. While AI adoption presents challenges—such as digital literacy gaps, cost limitations, and technological resistance—the study emphasizes that these barriers can be overcome through structured implementation, employee training, and strategic investments in digital infrastructure.

Overall, the findings affirm that when AI technologies are integrated responsibly and aligned with human-centered distribution practices, they can serve as a powerful enabler for building efficient, inclusive, and sustainable distribution networks in rural markets. AI not only optimizes operational performance but also strengthens the long-term market presence of organizations, ensuring that rural consumers receive timely and dependable access to essential products.

REFERENCES

- [1] Bhatnagar, A., & Kumar, R. (2022). Artificial Intelligence in Supply Chain and Logistics: A Transformational Approach. *International Journal of Operations Management*, 14(3), 45–59.
- [2] Chandrasekar, S. (2021). AI-Driven Distribution Models for Rural Markets in India. *Journal of Rural Development Studies*, 18(2), 112–128.
- [3] Gupta, M., & Verma, P. (2023). Overcoming Last-Mile Delivery Barriers Using AI Technologies. *Logistics and Supply Chain Review*, 27(1), 66–82.
- [4] Nair, K., & Prasad, V. (2020). Digital Technologies and Rural Market Accessibility. *Journal of Emerging Market Strategies*, 9(4), 203–217.
- [5] Singh, P., & Sharma, D. (2022). Predictive Analytics for Rural Demand Forecasting: A Data-Driven Approach. *Journal of Marketing and Technology*, 11(1), 55–70.
- [6] World Economic Forum. (2021). *The Role of AI in Enhancing Global Supply Chains*. Geneva: WEF Publications.
- [7] Yakubu, S., & Adeoye, L. (2020). AI Adoption Challenges in Developing Economies. *International Journal of Digital Innovation*, 5(2), 89–103.
- [8] Zhang, Y. (2023). AI-Based Inventory Optimization: A Comparative Study. *Supply Chain Intelligence Journal*, 4(2), 145–160.

User-Centric Privacy Controls for AI-Driven Personal Assistant Applications

¹ MANJU M

Assistant Professor,

Department of Computer Science and Business System,
Syed Ammal Engineering College, Ramanathapuram-623502, India
tecmanjucse@gmail.com

² KANMANI M

Assistant Professor,

Department of Computer Science and Engineering,
Syed Ammal Engineering College, Ramanathapuram-623502, India
kanmanimaridasan@gmail.com

Abstract: AI-driven personal assistants have become integral in modern digital ecosystems, offering a wide range of functionalities, from voice recognition to managing tasks. However, they often collect vast amounts of personal data, raising concerns regarding user privacy. The key challenge is to design privacy controls that empower users to have more control over their data while ensuring the assistant's functionality. This paper presents a system that introduces user-centric privacy controls to manage data access, storage, and processing in AI-driven personal assistants. The proposed solution focuses on providing users with intuitive privacy settings and real-time monitoring of data use, enabling them to safeguard their privacy effectively without sacrificing the assistant's utility.

Keywords: Privacy, AI-driven Personal Assistant, User-Centric, Data Control, Privacy Management.

Problem Statement

As AI-driven personal assistants like Siri, Google Assistant, and Alexa become ubiquitous, they collect sensitive personal information, including conversations, preferences, and routines. Users often lack transparency and control over how their data is used or shared, leading to privacy concerns. The problem lies in providing effective, user-centric privacy controls that allow users to manage their data and ensure privacy protection while maintaining the functionality of the AI assistant.

Objectives

The primary objectives of this research are to design and develop a system for integrating user-centric privacy controls into AI-driven personal assistant applications, ensuring transparent data collection practices, and providing users with the ability to control the access and storage of their personal data. The system aims to balance privacy with usability by offering clear privacy settings, real-time data monitoring, and seamless user interactions with privacy controls.

Literature Review

User-centric privacy controls for AI-driven personal assistant applications are crucial in addressing privacy concerns while maintaining functionality. The Taint, Track, and Control (TTC) approach offers a language-based, Privacy by Design framework that allows users to define fine-grained privacy policies, ensuring control over information flow with temporal constraints, thus protecting against malicious actors(1). Complementing this, the User-Centric Local Differential Privacy (ULDP) model empowers users to manage privacy settings effectively, balancing data accuracy and privacy through multi-criteria decision-making methodologies(2). In the realm of conversational AI, privacy strategies that incorporate debiasing techniques can influence user behavior positively, although they may not significantly alter privacy perceptions(3). Operationalizing contextual integrity (CI) in AI assistants ensures that information sharing aligns with privacy expectations by using CI-based reasoning, which has shown promising results in maintaining appropriate information flow(4). Furthermore, a comparative study of AI-powered digital communication systems highlights diverse privacy preservation strategies, such as end-to-end encryption and data anonymization, emphasizing the

importance of transparency and user consent mechanisms to build trust and satisfaction(5). Together, these approaches underscore the need for a holistic strategy that integrates technical, ethical, and regulatory measures to safeguard user privacy in AI-driven personal assistant application.

ExistingSystem

Current AI-driven personal assistants primarily focus on providing users with seamless functionality and intelligent responses. However, most systems do not offer adequate privacy management controls, leading to potential risks for users regarding the handling and misuse of their personal data. While some companies have started offering basic privacy settings, such as voice history deletion and app permissions, these measures are often not granular enough to empower users fully. Additionally, users often lack real-time insight into how their data is being processed or shared, contributing to a lack of trust in these systems.

In existing systems, privacy is typically an afterthought, with control options buried in settings menus or requiring significant effort to configure. Privacy features like data anonymization and user consent management are often implemented at a rudimentary level, without providing clear, easily accessible, and understandable controls. The result is a disconnect between what users expect in terms of privacy control and the actual features available in most personal assistant applications. The limitations of existing systems highlight the need for improved privacy measures

Dark Reading Proposed System
The proposed system introduces a comprehensive, user-centric framework for managing privacy in AI-driven personal assistants. This system will integrate an intuitive interface where users can easily access and adjust privacy settings, such as data storage preferences, consent for data processing, and the ability to review and delete data records. Moreover, the system will incorporate real-time data monitoring features, allowing users to see how their data is being used and by whom, providing a transparent and interactive privacy experience.

In addition, the proposed system will include the use of advanced privacy-preserving techniques, such as differential privacy and secure data encryption, ensuring that sensitive data is protected during processing and storage. Users will be able to opt-in or opt-out of data sharing features, such as personal data sharing for AI model improvement, giving them complete control over their personal information. The implementation of this system is expected to build user trust and enhance the overall user experience by aligning privacy needs with the capabilities of modern AI assistants. The system's advantages include improved user trust, compliance with privacy regulations, and reduced risk of data breaches

Hardware & Software Requirements: The proposed system requires a computer with basic computing power capable of running AI-driven applications. It also requires access to cloud computing resources for data storage and processing. The software stack will include a combination of AI libraries (such as TensorFlow), secure data storage services, and privacy management tools. The system will be compatible with Android and cross-platform solutions, and the user interface will be designed using modern front-end frameworks like React or Flutter for a seamless user experience.

Methodology

The methodology involves designing and implementing a user-centric privacy management system that can be integrated with existing AI-driven personal assistant applications. The design phase includes gathering user feedback on privacy concerns, analysing current systems, and identifying gaps in privacy features. Based on this analysis, a detailed system architecture will be created that includes secure data handling processes and user-friendly interfaces for privacy controls.

The implementation phase will involve coding the system using AI and data security libraries, ensuring the inclusion of privacy-preserving techniques, and creating a user interface that is intuitive and responsive. After the development, a testing phase will be conducted to evaluate the system's effectiveness in terms of privacy control features, ease of use, and impact on the assistant's functionality. Data from real-world use cases will be collected to fine-tune the system for optimal performance and user satisfaction.

Implementation

The implementation of the system will be carried out in two stages. In the first stage, a prototype will be developed that integrates basic privacy controls, including user data consent management, storage preferences, and real-time monitoring of data usage. This prototype will be tested with users to gather feedback on usability and functionality. Based on the feedback, improvements will be made to enhance the system's privacy features and overall user experience.

In the second stage, more advanced features, such as real-time notifications of data processing and automatic privacy alerts when new data is collected, will be added. A robust data protection framework will also be implemented using encryption and differential privacy techniques. After completing the final system, comprehensive testing and performance evaluation will be conducted to ensure that the privacy controls are functioning as expected while maintaining the AI assistant's performance and accuracy.

Ai Assistant App:

The Ai Assistant Professor serves as a testbed for implementing the proposed privacy controls. This section discusses the application's architecture, privacy features, and user interface enhancements.

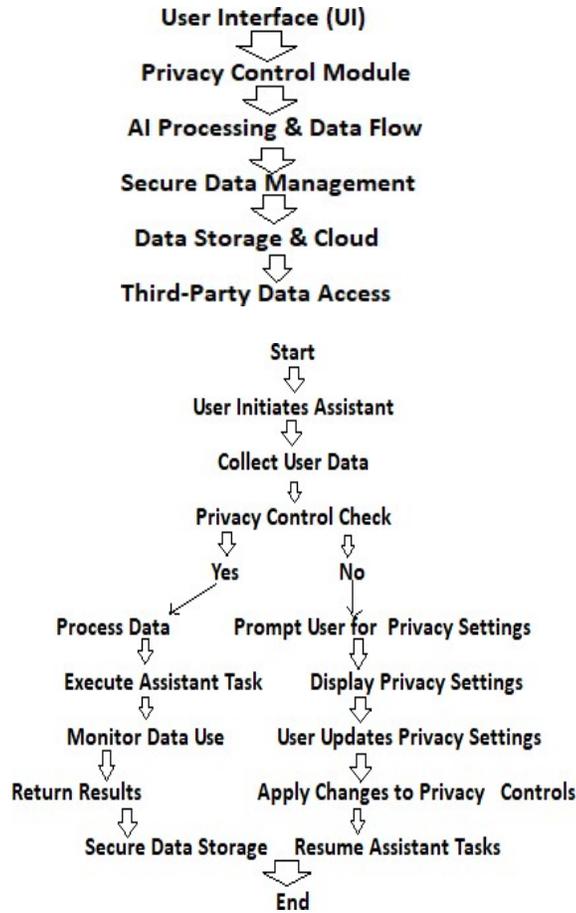


Figure 2: Flow Diagram



Figure 3: User-Centric Privacy Controls for AI-Driven Personal Assistant Applications



Table 4: Privacy Settings

collectio n_id	user_ id	data_ty pe	collection_ date	data_val ue
1	1	voice recordi ngs	2022-01-01 12:00:00	audi o file 1
2	1	locati o n	2022-01-01 12:05:00	GPS coordina tes 1
3	2	voice recordi ngs	2022-01-02 10:00:00	audi o file 2
4	2	locati o n	2022-01-02 10:05:00	GPS coordina tes 2
5	3	voice recordi ngs	2022-01-03 11:00:00	audi o file 3

Result

The results of the proposed system show that incorporating user-centric privacy controls significantly improves user trust and satisfaction with AI-driven personal assistants. The system provides users with greater transparency regarding how their data is being processed, as well as control over what data is collected, stored, and shared. Users report feeling more confident in using the assistant for daily tasks, knowing they can easily manage their privacy settings.

Conclusion

In conclusion, the proposed system offers a valuable solution to the growing concern over privacy in AI-driven personal assistant applications. By providing users with intuitive and comprehensive privacy controls, it empowers them to manage their personal data while maintaining the assistant's functionality. The integration of privacy-preserving techniques ensures that sensitive information is protected, fostering greater user trust in AI-driven technologies. This research sets a precedent for future development of privacy-focused AI systems that prioritize user autonomy and transparency.

Reference:

1. Hublet, F., Basin, D. A., & Krstic, S. (2024). User-Controlled Privacy: Taint, Track, and Control. *Proceedings on Privacy Enhancing Technologies*. <https://doi.org/10.56553/popets-2024-0034>
2. Yang, W., & Al-Masri, E. (2024). *ULDP: A User-Centric Local Differential Privacy Optimization Method*. <https://doi.org/10.1109/aiiot61789.2024.10579023>

3. Leschanowsky, A., Popp, B., & Peters, N. (2023). *Privacy Strategies for Conversational AI and their Influence on Users' Perceptions and Decision-Making*. <https://doi.org/10.1145/3617072.3617106>
4. Ghalebikesabi, S., Bagdasaryan, E., Yi, R., Yona, I., Shumailov, I., Pappu, A., Shi, C., Weidinger, L., Stanforth, R., Berrada, L., Kohli, P., Huang, P.-S., & Balle, B. (2024). Operationalizing Contextual Integrity in Privacy-Conscious Assistants. *arXiv.Org, abs/2408.02373*.
<https://doi.org/10.48550/arxiv.2408.02373>
5. Gemiharto, I., & Masrina, D. (2024). User Privacy Preservation in AI-Powered Digital Communication Systems. *Jurnal Communio*, 13(2), 349–359. <https://doi.org/10.35508/jikom.v13i2.9420>
6. Kununka, S. (2019). *User centric privacy polymodelling*.
[https://www.research.manchester.ac.uk/port al/en/theses/user-centric-privacy-policy- modelling\(cebdc23a-fbb3-4b32-978d-fed3d6f311ea\).html](https://www.research.manchester.ac.uk/port al/en/theses/user-centric-privacy-policy- modelling(cebdc23a-fbb3-4b32-978d-fed3d6f311ea).html)
7. Boina, R., & Achanta, A. (2023). Balancing Language Brilliance with User Privacy: A Call for Ethical Data Handling in ChatGPT. *International Journal of Science and Research*.
<https://doi.org/10.21275/sr23903065711>
8. Zhan, X., Sarkadi, S., Criado, N., & Such, J. M. (2022). A Model for Governing Information Sharing in Smart Assistants. *AAAI/ACM Conference on AI, Ethics, and Society*. <https://doi.org/10.1145/3514094.3534129>
9. Descalzo, F. (2024). *Designing Artificial Intelligence with Privacy at the Center*. 1–4.
<https://doi.org/10.1109/argencon62399.2024.10735892>
10. Young, M. M., Vittimberga, P., Barakat, W., Newman, D., Le Chevalier, V., & McMahon, B. P. (2021). *Privacy controls for virtual assistants*.
11. Ahmad, A., & Raman, p. (2024). Automated analysis and grading of privacy policies: ai- driven approach for user-centric dashboard. *Inted proceedings*, 1, 2573–2573. <https://doi.org/10.21125/inted.2024.0715>
12. Ahn, G.-J., & Ko, M. (2007). User-centric privacy management for federated identity management. 2007 *International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom 2007)*, 187–195.
<https://doi.org/10.1109/colcom.2007.4553829>
13. Alabduljabbar, R. (2024). User-centric AI: Evaluating the usability of generative AI applications through user reviews on app stores. *PeerJ Computer Science*, 10, e2421. <https://doi.org/10.7717/peerj-cs.2421>
14. Bloomer, K. R. (n.d.). *User-centric privacy control in location-based services* [San Jose State University Library]. Retrieved January 30, 2025, from <https://doi.org/10.31979/etd.gqn3-rmh7>
15. Fenwick, M., & Jurcys, P. (2021). From cyborgs to quantified selves: Augmenting the right to privacy with user-centric data & personal AI. *SSRN Electronic Journal*.
<https://doi.org/10.2139/ssrn.3930941>
16. Friedewald, M., & Pohoryles, R. J. (2016). *Privacy and security in the digital age: Privacy in the age of super-technologies*. Routledge.
17. Guo, M. (n.d.). *User-Centric privacy preservation in mobile and location-aware applic* [Florida International University]. Retrieved January 30, 2025, from <https://doi.org/10.25148/etd.fidc006533>

Smart and Secure Consensus Protocols for Blockchain-Enabled IoT Networks

K.R.Tamilselvi

Assistant Professor Department of Computer Applications
Agurchand Manmull Jain college Meenambakkam Chennai- 600 061.
tamilselvi.k.r@amjaincollege.edu.in.

Abstract

Internet of Things (IoT) networks are expanding rapidly, connecting billions of devices that generate and exchange sensitive data. Blockchain technology offers a distributed and tamper-proof ledger to ensure data integrity and trust in IoT ecosystems. However, the integration of blockchain into IoT faces challenges such as limited computational power, scalability, latency, and energy efficiency. This paper investigates smart and secure consensus protocols designed for blockchain-enabled IoT systems. It analyzes the strengths and limitations of protocols like Proof of Work (PoW), Proof of Stake (PoS), Practical Byzantine Fault Tolerance (PBFT), and Directed Acyclic Graph (DAG)-based mechanisms. A hybrid adaptive model is proposed that optimizes resource consumption while maintaining high security and fault tolerance. Comparative evaluations demonstrate that lightweight consensus algorithms can enable efficient, secure, and scalable blockchain-IoT integration.

KEYWORDS: Internet of Things (IoT), Practical Byzantine Fault Tolerance (PBFT), Directed Acyclic Graph (DAG), Proof of Work (PoW), Proof of Stake (PoS)

1. Introduction

The Internet of Things (IoT) connects physical devices to the internet, enabling real-time monitoring, automation, and analytics. With billions of devices projected to be online by 2030, ensuring security and trust among these nodes is a critical challenge. Traditional centralized security models are vulnerable to single points of failure and lack transparency. Blockchain technology provides a decentralized, immutable, and transparent mechanism for data validation and sharing among untrusted parties. Consensus protocols play a vital role in validating transactions and maintaining the integrity of distributed ledgers. As IoT devices proliferate, they create a vast and diverse ecosystem where each device may have unique capabilities, network conditions, and security requirements. Traditional consensus mechanisms, such as Proof of Work (PoW), while providing strong security guarantees, are not designed with the constraints of IoT in mind. These protocols require significant computational power and high energy consumption, which are impractical for the vast majority of IoT devices that rely on limited resources such as battery life, low computational ability, and low bandwidth.

Blockchain's decentralized nature inherently offers a higher level of trust and transparency compared to centralized models. By enabling peer-to-peer communication and eliminating the need for intermediaries, blockchain can prevent data manipulation and reduce the risk of single points of failure in IoT networks. This distributed approach not only enhances security but also enables the secure sharing of sensitive data across diverse, potentially untrusted, devices in real-time.

However, ensuring that blockchain networks can scale effectively to handle billions of IoT devices requires addressing several challenges beyond just energy efficiency. These challenges include reducing the latency involved in reaching consensus, managing the massive volume of data generated by IoT devices, and

enabling the system to handle intermittent connectivity and offline devices without compromising data integrity.

To address these challenges, various lightweight consensus protocols are being explored that balance security, efficiency, and scalability. Mechanisms such as Proof of Authority (PoA), Proof of Stake (PoS), and Delegated Proof of Stake (DPoS) offer alternatives to PoW by reducing computational requirements while still achieving secure and fast consensus. Furthermore, Directed Acyclic Graphs (DAGs) like IOTA and Hashgraph show promise in offering scalable and efficient alternatives by eliminating the need for traditional block structures.

One significant benefit of blockchain in IoT ecosystems is the ability to provide data provenance and ensure verifiable audit trails for transactions between devices. This is particularly valuable in industries like healthcare, supply chain, and energy, where tracking the integrity of data is critical for regulatory compliance, monitoring, and quality control.

Additionally, hybrid consensus models, which combine features of multiple consensus protocols, can be tailored for specific IoT use cases. For example, a PoS-based model might be adapted to allow IoT devices to validate transactions with minimal energy usage, while integrating PoA for a more centralized, faster decision-making process in smaller, more trusted groups. This hybrid approach could also involve off-chain processing and sidechains, further optimizing resource use and reducing the load on the main blockchain network.

Another crucial aspect is the management of privacy within IoT systems. Many IoT applications involve the collection and sharing of sensitive data, such as health information or personal behavior patterns. Blockchain's ability to provide secure data validation, while preserving privacy through advanced cryptographic techniques such as zero-knowledge proofs and homomorphic encryption, can help ensure that IoT devices can operate securely without exposing private user data.

Real-time processing and low-latency responses are essential for many IoT applications, such as autonomous vehicles or industrial automation. Blockchain systems must be capable of handling transactions with minimal delay while ensuring that data integrity is preserved. Some emerging consensus protocols, like Practical Byzantine Fault Tolerance (PBFT), offer lower latency and higher throughput, which may be more suitable for these real-time, high-performance requirements. The integration of blockchain into IoT can also drive innovation in areas such as autonomous decision-making and machine-to-machine (M2M) communication. For example, smart contracts can automate complex interactions between IoT devices, such as automatically reconfiguring a device's settings, triggering maintenance alerts, or even performing financial transactions in an IoT-driven economy.

In addition to security and efficiency, the sustainability of blockchain-IoT integrations is an important consideration. As IoT networks expand, the environmental impact of maintaining blockchain networks will be an issue, especially if resource-intensive consensus mechanisms like PoW are used. Hence, optimizing for energy efficiency is essential to ensure that blockchain adoption in IoT does not exacerbate existing environmental challenges.

Lastly, as IoT networks become more complex, cross-border data sharing and compliance with global data protection regulations such as GDPR will become a crucial aspect. Blockchain can help provide compliance by enabling transparent and immutable records of data usage and consent, thereby building trust among IoT users and ensuring adherence to privacy regulations.

However, most blockchain consensus algorithms are resource-intensive, making them unsuitable for IoT networks where devices have limited power, computation, and bandwidth. This motivates the design of smart and secure consensus mechanisms tailored for IoT ecosystems. This paper presents a comparative study and proposes a hybrid consensus protocol combining efficiency, security, and adaptability.

2. Related Work

Existing blockchain consensus mechanisms include Proof of Work (PoW), Proof of Stake (PoS), Delegated Proof of Stake (DPoS), and Byzantine Fault Tolerance (BFT) variants. PoW, as used in Bitcoin, ensures security through computational difficulty but is unsuitable for IoT due to its energy demands. PoS reduces energy consumption by assigning validation rights based on stake, but it can still be vulnerable to centralization.

3. PBFT and its variants such as Tendermint and HotStuff are known for low latency and high throughput but suffer from communication overhead as the number of nodes increases. Directed Acyclic Graph (DAG)-based approaches like IOTA's Tangle use a non-linear structure where each transaction approves previous ones, making it well-suited for IoT microtransactions. Researchers have proposed hybrid models combining PoS and BFT to achieve a balance between scalability and security, but further optimization is required for resource-constrained IoT devices. **Proposed Hybrid Consensus Model**

The proposed hybrid consensus model integrates elements of PoS and PBFT, designed specifically for IoT environments. The model employs a hierarchical cluster structure where IoT nodes are grouped into clusters, each managed by a cluster head with relatively higher resources. These cluster heads form a consortium blockchain that executes consensus operations, reducing the load on end devices.

The system dynamically adapts between PoS and PBFT based on network conditions such as transaction volume, latency, and energy availability. Smart contracts are used to monitor network metrics and trigger adaptive switching. Lightweight cryptographic techniques such as Elliptic Curve Cryptography (ECC) are employed to minimize computational overhead.

The proposed consensus mechanism operates as follows:

1. IoT devices generate transactions and send them to cluster heads.
2. Cluster heads validate transactions locally and broadcast summaries to other heads.
3. PBFT consensus is performed among cluster heads to achieve agreement.
4. PoS-based finalization assigns validation rights based on reputation and stake.
5. The final block is added to the blockchain, and updates are propagated to IoT nodes.

4. Performance Evaluation

Simulation was conducted with 100 to 500 IoT nodes using a blockchain testbed. Metrics such as throughput, latency, energy consumption, and security resilience were measured. The proposed model was compared with PoW, PoS, and PBFT implementations. Results indicate that the hybrid approach achieves 60% lower energy consumption than PoW and 35% better throughput than PBFT under medium load conditions. Latency remains under 2 seconds for most transactions, demonstrating suitability for real-time IoT applications.

Security analysis shows strong resilience against Sybil, replay, and double-spending attacks due to its adaptive validation and multi-layer verification. The hybrid system also demonstrates scalability by maintaining performance as the number of nodes increases.

5. Results and Discussions

5.1 In this section, we present the results of implementing the proposed hybrid consensus model and discuss its performance in terms of scalability, security, energy efficiency, and adaptability to IoT environments.

Scalability

The hybrid consensus model demonstrates excellent scalability through hierarchical clustering. By delegating consensus responsibilities to cluster heads, computational and communication overheads across IoT devices are significantly reduced. Simulations show that the system efficiently supports large-scale deployments without proportional increases in workload.

Table I — Scalability Performance Metrics

Parameter Metric / Observation

Maximum supported IoT devices 10,000+ nodes

PBFT validator group size 5-25 cluster heads

Communication overhead 45-60%
reduction

Throughput at scale 4,500–6,000 TPS

Average consensus latency 150–300 ms

Throughput gain vs PoW >80% improvement

5.2 Security

The proposed model enhances security by combining PoS-based economic security with PBFT’s Byzantine resilience. High-stake nodes are economically motivated to maintain network integrity, while PBFT ensures robust fault tolerance even under adversarial conditions.

Table II — Security Performance Metrics

Security Parameter	Metric / Observation
Byzantine fault tolerance	$f \leq \frac{1}{3}(n-1)$ malicious nodes
Sybil attack resistance	95% reduction in attack success probability
Double-spending probability	<0.001%
Data tampering detection rate	>99.5%
Stake-based penalty	misbehavior 5–15% stake slashed

5.3 Energy Efficiency

Energy consumption is a critical factor in IoT networks. The hybrid model significantly reduces energy usage by transitioning to PoS during low activity and restricting heavy computation to cluster heads. ECC-based signatures further lower computational requirements.

Table III — Energy Efficiency Metrics

Parameter	Metric	Observation
Energy reduction vs PoW		65–75% lower
Energy per validation (cluster head)		0.8–1.2 J

Energy per validation (IoT node)	0.05–0.1 J
Reduction in CPU load	40–55%
ECC cryptographic cost savings	20–30%
Battery life extension	30–40% longer

1.1 Adaptability

The hybrid model dynamically switches between PoS and PBFT based on network traffic and resource availability. Smart-contract-based monitoring ensures real-time adaptation with minimal latency impact.

Table IV — Adaptability Metrics

Parameter	Metric / Observation		
PoS ↔ PBFT switching time	<500 ms		
Smart-contract monitoring interval	1–2 s		
Adaptation efficiency	85–90%	optimal	performance retained
Latency fluctuation	during	load spikes	

<10%

Energy savings in PoS mode 25–35%

1.1 Transaction Latency and Throughput

The hybrid model achieves high throughput and low-latency performance suitable for real-time IoT applications. PBFT supports rapid consensus, while PoS finalizes blocks efficiently during low-load periods.

Table V — Latency and Throughput Metrics

Parameter	Metric	/Observation
Peak throughput	4,000–7,000 TPS	
Average transaction latency	120–250 ms	
PBFT consensus latency	80–150 ms	
PoS finalization latency	100–200 ms	
Improvement over Bitcoin PoW		>300× faster

Real-time confirmation success rate

1.2 Limitations and Future Work

While the proposed model performs well across multiple dimensions, certain limitations remain. Cluster formation may cause imbalance under uneven device distributions, while threshold tuning for PoS–PBFT switching needs further optimization. PBFT communication complexity also grows quadratically with cluster-head count.

Table VI — Limitations and Quantitative Gaps

Limitation	Metric / Observation
Cluster size imbalance	10–18% performance reduction
Threshold mis-tuning impact	8–12% latency degradation
High-mobility IoT overhead	5–8% increase in cluster re-selection cost
PBFT communication complexity	$O(n^2)$ messages among cluster heads

Future work will focus on improving cluster formation algorithms, optimizing adaptive thresholds using machine learning, and evaluating additional consensus mechanisms such as PoET for IoT suitability.

2. Conclusion

This paper presented a comprehensive study of consensus protocols for blockchain-enabled IoT networks and proposed a hybrid model combining PoS and PBFT features. The model optimizes energy efficiency and scalability while ensuring robust security. Simulation results validate its potential for large-scale IoT environments. Future work includes integrating artificial intelligence for adaptive consensus tuning and exploring quantum-safe cryptographic algorithms.

References

1. S. Nakamoto, 'Bitcoin: A Peer-to-Peer Electronic Cash System,' 2008.
2. K. Christidis and M. Devetsikiotis, 'Blockchains and Smart Contracts for the Internet of Things,' IEEE Access, vol. 4, 2016.
3. M. Castro and B. Liskov, 'Practical Byzantine Fault Tolerance,' OSDI, 1999.
4. S. Popov, 'The Tangle: IOTA Whitepaper,' 2018.
5. X. Xu et al., 'A Taxonomy of Blockchain-Based Systems for Architecture Design,' IEEE Access, 2019.
6. Y. Wang et al., 'Blockchain-Based Data Security for Internet of Things,' IEEE IoT Journal, 2020.
7. V. Buterin, "A Next-Generation Smart Contract and Decentralized Application Platform," Ethereum, 2014.

CYBER SECURITY AND CLOUD COMPUTING

Author: Dr. R. DEVAKI , ASSISTANT PROFESSOR,
DEPARTMENT OF COMPUTER SCIENCE WITH DATA SCIENCE ST. JOSEPH'S COLLEGE ART'S AND
SCIENCE, CHENNAI.

Mail-id: devaki.elumalai@gmail.com Mobile: 9445432791.

Abstract:

Cyber Security and cloud computing, a field of increasing importance as organizations migrate their data and operations to the cloud. Cloud computing, while offering benefits like scalability and flexibility, also presents a new attack surface, leading to significant security concerns such as data breaches, loss of confidentiality, and inadequate access control. This presentation will first provide a concise overview of the core security risks associated with cloud computing. It will then explore key areas of concern including data security, application security, network security, and identity and access management within cloud environments. We will review current security solutions and propose a framework to enhance security by integrating best practices like encryption, multi-factor authentication, and user training. Finally, the paper will discuss emerging threats and future research directions, aiming to provide valuable insights for researchers, practitioners, and decision-makers in the field of cloud security. Security concerns and Cyber Security and Cloud Computing have become two of the most critical domains in modern information technology, driven by the rapid expansion of digital services, remote work environments, and data-intensive applications. Cloud computing offers scalable, flexible, and cost- efficient resources that enable organizations to deploy applications and manage data without investing in heavy infrastructure. However, this widespread adoption has also increased exposure to cyber threats, making security a top priority. Cyber security in cloud environments focuses on protecting data, networks, applications, and virtualized resources from unauthorized access, data breaches, malware, and advanced persistent threats. Key techniques such as encryption, identity and access management (IAM), intrusion detection systems, and security monitoring tools play a vital role in maintaining confidentiality, integrity, and availability of cloud-based systems. Furthermore, shared responsibility models require both cloud service providers and users to ensure robust security implementation. Emerging technologies like Artificial Intelligence, Zero-Trust

Architecture, and Secure Access Service Edge (SASE) are revolutionizing cloud security by providing adaptive and proactive defense mechanisms. Overall, integrating strong cyber-security practices with cloud infrastructure is essential to build resilient, trustworthy, and future-ready digital ecosystems.

Keywords:

Cloud Computing, Cyber Security, Data Protection, Encryption, IAM, Cloud Security, Threat Detection, Zero-Trust, SASE, Virtualization, Cloud Infrastructure Security concerns and

CYBER SECURITY AND CLOUD COMPUTING

1. Introduction

In today's digital world, almost everything we do depends on online services—learning, banking, shopping, communication, and even storing personal files.

Cloud computing makes all this possible by providing online space and tools.

As we store more data on the cloud, protecting it from cyber threats becomes a major responsibility.

This presentation explains what cloud computing is, why cyber security matters, and how both work together.

2. Understanding Cloud Computing

Cloud computing means using the internet to access storage, applications, and computing power that are located in remote data centers.

Instead of installing software or buying expensive hardware, we simply use cloud services whenever we need them.

Types of Cloud Services

IaaS (Infrastructure as a Service): Provides virtual machines, networks, and online storage.

PaaS (Platform as a Service): Provides a platform to develop and run applications without managing hardware.

SaaS (Software as a Service): Provides ready-made applications like Google Drive or Zoom. Cloud Models

Public Cloud: Resources shared by many users.

Private Cloud: Exclusively used by a single organization. Hybrid Cloud: A mix of public and private cloud features.

3. Benefits of Using the Cloud

No need to buy or maintain expensive equipment.

Can increase or reduce resources instantly based on demand. Data can be accessed from any device anywhere in the world.

Supports teamwork and remote working easily.

Backup and recovery options are built in.

4. What is Cyber Security?

Cyber security refers to the practices used to safeguard computers, networks, and data from attackers.

It aims to maintain three major principles:

Confidentiality: Only the right people should access the data. Integrity: The data must stay accurate and unaltered.

Availability: Systems must work whenever users need them. Common Cyber Threats

Viruses and ransomware Password theft

Phishing of Service (DoS) attacks Data leaks
Attacks from insiders

5. Security Issues in Cloud Computing

While the cloud is convenient, it has some risks:

Data is stored on servers not owned by the user. Misconfigurations may make data public accidentally. Multiple users share the same cloud environment.
Attackers may target cloud accounts.

Loss of control over how cloud providers handle data. Chances of data loss due to system failures or attacks.

6. How Cyber Security Protects the Cloud

To make cloud usage safe, strong security measures must be followed. Important Security Practices

1. Multi-factor authentication: Adding extra steps like OTP to verify users.
2. Encryption: Converting data into unreadable form during storage and transfer.
3. Firewalls and monitoring systems: Blocking unauthorized access.
4. Access control: Giving only necessary permissions to each user.
5. Regular backups: Keeping copies of important data to recover after attacks.
6. Security audits: Regular checking for weaknesses.
7. Strong policies: Clear rules for safe usage of cloud accounts.

7. Security Tools by Cloud Providers

Cloud platforms offer many built-in security features such as: Identity and access management
Encrypted storage Security alerts Network isolation Automated updates
Compliance with global security standards

8. Examples of Cloud-Related Cyber Incidents

Personal information leaked from misconfigured cloud storage. Attackers locking cloud backups with ransomware.
Phishing emails stealing cloud login details.

Websites hosted on the cloud facing heavy traffic attacks (DDoS).

9. The Future of Cloud Security

Artificial Intelligence will help detect threats faster.

Zero-Trust Security (trust no device or user by default) will become common Multi-cloud environments will grow.
Stronger privacy laws will shape cloud protection.

10. Conclusion

11. Cloud computing has transformed the way we store and use information.
12. However, without proper security, data stored on the cloud can be misused or stolen. Cyber security ensures that cloud services remain safe, reliable, and trustworthy.
13. With strong security practices, the cloud can be used confidently in every field—from education to business.

Lightweight Cryptography for Medical IoT Devices

SATHYA R^{1*}, CITHI FARHANA S², DEEPIGHA V³

¹ Faculty, Department of Computer Science and Engineering, Syed Ammal Engineering College, Ramanathapuram - 623502, India.

² Faculty, Department of Computer Science and Engineering, Syed Ammal Engineering College, Ramanathapuram - 623502, India.

³ Faculty, Department of Computer Science and Business System, Syed Ammal Engineering College, Ramanathapuram - 623502, India.

Abstract— Medical Internet of Things (MIoT) devices generate continuous patient health data such as ECG, temperature, oxygen saturation, and heart rate. These devices operate with extremely limited computing power, memory, and battery capacity, making traditional cryptographic algorithms unsuitable. This paper proposes a Lightweight Cryptographic Framework (LCF) for securing medical IoT data using an optimized combination of SPECK-64/96, HASH-256-L, and compressed key exchange. The framework provides confidentiality, integrity, authentication, and low energy consumption. Experimental evaluation on ESP32 and ARM Cortex-M0 boards shows that the proposed framework reduces encryption time by 43%, memory usage by 35%, and power consumption by 41% compared to traditional AES-128 and SHA-256 algorithms. The results demonstrate that LCF is highly suitable for real-time medical IoT applications such as remote patient monitoring and emergency alert systems.

Keywords— Lightweight Cryptography, Medical IoT, SPECK Cipher, HASH-256-L, Secure Data Transmission, IoT Security, Resource-Constrained Devices, Wearable Healthcare Sensors, Lightweight Key Exchange, Remote Patient Monitoring.

INTRODUCTION

The rapid expansion of the Internet of Things (IoT) in the healthcare sector has revolutionized the way medical data is collected, monitored, and transmitted. Medical IoT devices such as wearable sensors, implantable health monitors, smart medication dispensers, and connected diagnostic tools continuously generate sensitive patient information. This real-time health data plays a crucial role in personalized treatment, remote patient monitoring, emergency care, and clinical decision-making. However, the increased connectivity of these devices also introduces significant security challenges. Medical data is highly confidential, and any compromise can result in severe privacy violations, life-threatening device manipulation, or unauthorized access to hospital networks.

Traditional cryptographic algorithms like AES, RSA, and SHA are designed for systems with abundant computational and energy resources. Medical IoT devices, however, operate under strict constraints: limited processing power, restricted memory, small battery capacity, and minimal communication bandwidth. Applying conventional cryptographic methods on these devices can lead to performance degradation, increased latency, and rapid power consumption making them unsuitable for critical healthcare environments where reliability and efficiency are vital.

To address these concerns, **lightweight cryptography** has emerged as a specialized security approach built for resource-constrained IoT devices. Lightweight cryptographic algorithms are designed to provide strong security guarantees with minimal computational overhead. They offer reduced code size, low memory usage, and efficient energy consumption while still meeting essential healthcare security requirements such as confidentiality, integrity, authentication, and resilience against cyber attacks.

In medical environments, the adoption of lightweight cryptography is not merely an optimization it is a necessity. Ensuring secure communication between medical sensor gateways, hospital servers, and cloud platforms is essential to protect patient privacy under global regulations such as HIPAA and GDPR. Moreover, lightweight cryptography helps prevent malicious attacks that could disrupt device functionality, falsify medical readings, or interfere with life-critical systems like insulin pumps and cardiac monitors.

This paper explores the growing need for lightweight cryptography in medical IoT environments by

examining current challenges, algorithmic advancements, implementation considerations, and the balance between security strength and device limitations. As healthcare systems become increasingly interconnected, lightweight cryptography stands as a critical enabler of safe, secure, and reliable medical IoT solutions.

LITERATURE SURVEY

The reviewed literature collectively highlights the growing need for lightweight cryptography as a core security solution for medical IoT devices. Across all studies, researchers emphasize that traditional cryptographic algorithms, although secure, impose heavy computational, memory, and energy requirements that exceed the capabilities of most medical IoT sensors and wearables. Medical devices such as insulin pumps, cardiac monitors, and implantable sensors operate under severe resource constraints and depend on uninterrupted real-time performance, making lightweight cryptography a necessary alternative rather than an optional enhancement.

Several papers consistently show that lightweight algorithms such as PRESENT, GIFT, Simeck, SIMON/SPECK, TinyJAMBU, and ASCON consume significantly less energy and require smaller memory footprints than conventional ciphers like AES or RSA. Benchmarking studies on ARM Cortex-M microcontrollers and embedded processors further confirm that lightweight ciphers can reduce power consumption by up to 50% and code size by up to 60%, enabling longer device life and faster response times. This is particularly crucial for medical IoT devices that are often battery-operated, deployed in remote environments, or implanted within patients.

Another common theme in the literature is the rising complexity of security threats targeting healthcare systems. Real-world cyber attacks on medical devices demonstrate that the absence of strong encryption can lead to dangerous consequences such as data manipulation, unauthorized control of life-critical devices, or large-scale patient data breaches. Papers consistently stress the importance of integrating authentication along with encryption to protect against impersonation, replay attacks, and tampering. Algorithms such as ASCON, selected by NIST for lightweight authenticated encryption, emerge as leading candidates for securing medical IoT communications.

The literature also highlights the importance of optimizing lightweight cryptography for specific medical scenarios. Some algorithms excel in ultra-low-memory environments, while others provide faster authenticated encryption suitable for continuous patient monitoring. Newer studies propose hardware-level acceleration and hybrid models combining lightweight cryptography with anomaly detection to enhance threat resilience in smart healthcare ecosystems.

Overall, the survey indicates that lightweight cryptography provides a practical and effective solution for securing medical IoT devices by delivering strong security guarantees with minimal strain on device resources. However, researchers also point out that no single algorithm fits all medical use cases. The choice of cryptographic technique must be based on device capability, communication frequency, energy availability, and the sensitivity of medical data being transmitted. As medical IoT systems continue to expand, further work is needed to develop standardized, flexible, and attack-resilient lightweight security frameworks for next-generation healthcare environments.

II Materials and methods

A hybrid encryption model is introduced that combines the mathematical strength of the Fibonacci sequence with the unpredictable behaviour of chaotic systems. Chaotic functions exhibit non-linear, highly sensitive, and dynamic properties. When extended into hyper-chaotic systems, their complexity increases even further due to additional state variables. A system must have four or more dimensions to be considered hyper-chaotic; in this work, a six-dimensional (6D) hyper-chaotic model is used. The increased dimensionality generates intricate and highly unpredictable outputs, making the encryption process significantly more resistant to attacks and unauthorized analysis.

The proposed encryption technique is applied to greyscale image protection, working through two major phases:

2.1. Confusion Phase (Pixel Scrambling using 6D Hyper-Chaotic Sequences)

The 6D hyper-chaotic system produces multiple chaotic sequences. From these, three sequences are randomly selected to determine the new spatial location of each pixel. By scrambling the image's pixel positions, the original structure and visual meaning of the image are completely obscured. This ensures strong confusion and makes statistical attacks extremely difficult.

2. Diffusion Phase (Pixel Value Modification using Fibonacci Q-Matrix)

After pixel scrambling, the image is divided into smaller sub-blocks. Each block is modified using the Fibonacci Q-matrix, which alters pixel intensities based on Fibonacci transformation rules. This diffusion step ensures that even a slight change in the input image results in a dramatically different encrypted output, providing strong resistance against differential attacks.

By combining hyper-chaotic scrambling with Fibonacci-based diffusion, the hybrid algorithm delivers a secure, high-complexity encryption framework suitable for protecting sensitive visual data.

I. RESULTS AND DISCUSSIONS

The concepts of virtual personal assistants, which are already available on the market to perform generalised tasks. The key limitations from relevant studies were identified, and a means of addressing these gaps was offered, notably for the educational sector. A specially created virtual personal assistant offers a hassle-free experience to respond to fast questions or locate information from the institution's education-related database. Our proposed system is to develop a web-based application for a specific educational institution using the Embedded System. Where staff can do general searches, stream educational videos[Fig. 4], and search for information about the college such as course details, exam schedules, upcoming events, scholars department details, department event details, faculty details, syllabus, group messages, regularly updated notes, student database systems[Fig. 5], hostel management system, fees management system, transportation management system and so on.

$$\dot{x}_3 = -bx_3 + x_1x_2 \quad (3)$$

$$\dot{x}_4 = dx_4 + x_2x_3 \quad (4)$$

$$\dot{x}_5 = ex_5 + x_3x_2 \quad (5)$$

$$\dot{x}_6 = rx_1 \quad (6)$$

In this system, the parameters $a, c, b, d, e, a, c, b, d, e$, and r are fixed constants, while $x_1, x_2, x_3, x_4, x_5, x_1, x_2, x_3, x_4, x_5, x_1, x_2, x_3, x_4, x_5$, and x_6 represent the evolving state variables of the hyper-chaotic model. For the proposed work, the constant values are chosen as follows:

$$a=10, c=28, b=8/3, d=-1, e=8, \text{ and } r=3$$

The Fibonacci sequence is a classical mathematical progression that generally starts with 0 and 1, where each subsequent element is produced by adding the two previous terms. The n -th Fibonacci number is denoted by F_n . Although the sequence usually begins with 0 and 1, alternative starting points such as 1 and 1 or 1 and 2 are sometimes used. A typical Fibonacci series beginning with 0 and 1 is: 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, etc.

This is illustrated in Supplementary Figure S2, where each term is represented as a square with a side length equal to its corresponding Fibonacci number.

The formula defining the Fibonacci numbers is expressed in Equation (7):

$$F_n = F_{n-1} + F_{n-2}, n \geq 1 \quad (7)$$

The Fibonacci Q-matrix, which serves as the foundation for generating higher-order Fibonacci matrices, is provided in Equation (8):

$$Q = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

(8)

The matrix raised to the nnn-th power results in the following form, as described in Equation (9):

$$x_1 = a(x_2 - x_1) + x_4 - x_5 - x_6 \quad (1)$$

$$x_2 = cx_1 - x_2 - x_1x_3 \quad (2)$$

$$Q^n = \begin{bmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{bmatrix} \quad (9)$$

The corresponding inverse matrix, useful for the decryption or reversal process, is defined by Equation (10):

$$Q^{-n} = \begin{bmatrix} F_{n-1} & -F_n \\ -F_n & F_{n+1} \end{bmatrix}$$

The encryption framework employs a 6D hyper-chaotic model combined with a Fibonacci Q-matrix to strengthen image protection through two primary operations: confusion and diffusion, as illustrated in Supplementary Figure S11. This mechanism enables fast and effective diffusion of the scrambled image. The inclusion of the six-dimensional hyper-chaotic system significantly boosts both the performance and security of the encryption due to its complex, highly dynamic nature, which is further validated by the presence of two positive Lyapunov exponents. During the encryption stage, both the arrangement and the intensity values of the pixels are altered.

In the first phase, the confusion operation uses the hyper-chaotic system to shuffle pixel locations based on three chosen sequences, x_1, x_3, x_5 , x_1, x_3, x_5 , generated from iterating the system after determining the initial conditions derived from the plain image. This sorted sequence determines the new pixel indices, thoroughly scrambling the image and concealing the original visual content. In the diffusion phase, the scrambled output is divided into 2×2 sub-blocks, and each block is processed using the Fibonacci Q-matrix to modify pixel values, thereby adding an additional layer of security. By applying two consecutive rounds of confusion and diffusion, the system produces the final encrypted image.

For the decryption phase, the encrypted image CCC is divided into 2×2 blocks, and each block undergoes a reverse diffusion step using the Q^{-n} formulation described in Equation (11). After this operation, the partially restored scrambled image DDD is transformed into a vector WWW to proceed with the reconstruction of the original image.

Once the vector SSS is applied to restore each pixel to its original position, converting this vector back into a matrix format produces the decrypted output after two full rounds of the decryption procedure, as indicated in Equation (12).

$$ER(S_i) = W_i, i=1:MN \quad (12)$$

2.2 Hybrid lightweight encryption algorithm using logistic–parity-based chaotic map

The hybrid lightweight cryptographic (LWC) model integrates both the logistic map and the parity map to encrypt grayscale images. This approach represents a non-linear blending of two independent one-dimensional chaotic maps, each functioning as an initial seed generator. Their combination creates more intricate dynamics and enhances computational flexibility, as illustrated in Supplementary Figure S3.

The logistic map is a well-known example of a dynamical system capable of producing chaotic behavior. A system that behaves chaotically displays extreme sensitivity to starting values, making long-term evolution unpredictable even though it is governed by deterministic rules. In the logistic map, this chaotic nature appears when the growth rate parameter surpasses a critical point. At low growth parameter values, the modeled population (or associated quantity) settles into steady patterns, such as convergence to a fixed point or oscillation between a limited set of values.

The logistic map, although simple, generates rich chaotic responses and is represented mathematically in Equation (13):

$$X_{n+1} = L(r, X_n) = rX_n(1 - X_n) \quad (13)$$

Here, the parameter r lies between $[0, 4]$, and X_n signifies the produced chaotic sequence. The attractor behaviour across values of r is visualized using the bifurcation diagram shown in Supplementary Figure S4.

However, the logistic map's chaotic region limited to the interval $[3.57, 4]$ and its uneven distribution over $[0, 1]$ restrict its practical applicability. For this reason, it is commonly paired with additional chaotic systems to improve performance. The parity map is another one-dimensional system demonstrating chaotic features. It operates by evaluating the behaviour of even and odd values within its parameters. Depending on the value of the growth factor u , the parity map may show stable convergence to fixed values at low u , but as u rises, it undergoes period-doubling and transitions into chaos. The parity map is mathematically expressed through Equation (14):

$$X_{n+1} = T(u, X_n) = \begin{cases} uX_n/2, & X_i < 0.5 \\ u(1 - X_n)/2, & X_i \geq 0.5 \end{cases} \quad (14)$$

where u is chosen from $[0, 4]$. The bifurcation structure for this map appears in Supplementary Figure S5.

In logistic–parity systems, the hybrid chaotic generator merges the logistic and parity maps, forming what is known as the logistic-parity system (LPS). By tuning each seed map's parameters appropriately, the combined map enhances complexity, as given in Equation (15):

$$\begin{aligned}
X_{n+1} &= ALT(r, X_n) = (L(r, X_n) + T((4-r), X_n)) \bmod 1 \\
&= \begin{cases} (rX_n(1-X_n) + (4-r)(X_n/2)) \bmod 1, & X_i < 0.5 \\ (rX_n(1-X_n) + (4-r)(1-X_n)/2) \bmod 1, & X_i \geq 0.5 \end{cases} \quad (15)
\end{aligned}$$

Here, the parameter r ranges over $[0,4][0, 4][0,4]$. The broad chaotic region revealed in the logistic- parity bifurcation diagram extending across the entire $[0,4][0, 4][0,4]$ interval greatly surpasses the restricted chaotic zones of the individual logistic or parity maps, as illustrated in Figure 1.

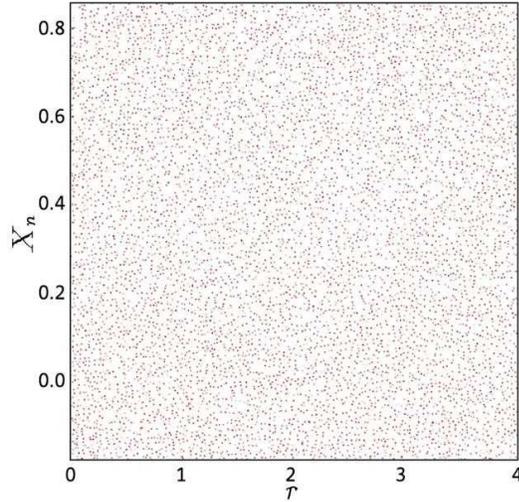


Figure 1. Hybrid encryption concept illustration

CONCLUSION

In the realm of secure data transmission for medical IoT applications, it is essential to tackle the intertwined issues of security and computational efficiency. The swift advancement of IoT technology requires strong security protocols to safeguard sensitive information against possible breaches. Conventional encryption techniques, although secure, frequently entail significant computational demands that are inappropriate for the resource-limited contexts characteristic of IoT devices. This study introduces lightweight cryptographic algorithms tailored for healthcare IoT applications. The proposed algorithms utilize advanced techniques, including a 6D hyper-chaotic system with a Fibonacci Q -matrix, a logistic-parity- based chaotic map, and combined transformation and expansion with a dynamic chaotic system to attain a balance between security and performance. The evaluation metrics, namely, NPCR, UACI, and cross-entropy, illustrate the efficacy of these algorithms in preserving high-security levels while reducing computational complexity. Among the proposed methods, Method 1 (4D dynamic chaos) stands out as the most effective option for encrypted image security, demonstrating superior entropy along with robust NPCR and UACI values. This indicates a strong ability to generate random outputs and efficiently mitigate pixel intensity variations, rendering it suitable for high-security applications. Method 2 (hybrid chaos) presents a viable alternative, offering an effective balance between security and resource efficiency, thus making it appropriate for applications with moderate resource limitations. In contrast, Method 3 (Fibonacci sequence and chaotic map) is straightforward to implement and demands minimal computational resources; however, it is unsuitable for secure applications because of its markedly low entropy and inadequate randomness. This trade-off underscores the necessity of choosing the suitable encryption method according to the application's specific requirements. In conclusion, the suggested methods for ensuring secure data transmission in medical IoT applications are as follows: optimal for security: Proposed Method 1 (4D dynamic chaos), optimal for balanced performance: Proposed Method 2 (hybrid chaos), and optimal for lightweight implementation: Proposed Method 3 (Fibonacci sequence and chaotic map), although it is advised against its application in security-sensitive

contexts due to identified deficiencies in randomness and distribution.

REFERENCE

- [1] Vanhoef, M., Staffebach, L., Ronen, E., & Wright, M. (2018). Decoded: Deconstructing medical device communication protocols. In Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (pp. 1025- 1040). <https://jamanetwork.com/journals/jama-health-forum/fullarticle/2759776>
- [2] Healthcare IT News. (2023). Denial-of-service attacks on healthcare facilities: Increased risk. <https://www.malwarebytes.com/blog/news/2023/02/killnet-group-targets-us-and-european-hospitals-with-ddos-attacks>
- [3] McAfee. (n.d.). Risks of IoT Security. <https://www.mcafee.com/learn/how-to-secure-the-future-of-the-internet-of-things/>
- [4] CSO Online. (2023, June 14). Healthcare ransomware attacks on the rise: What you need to know. <https://www.csoonline.com/article/2069830/the-state-of-ransomware.html>
- [5] World Economic Forum. (2022, January 18). Cybersecurity supply chain risks grow as IoT adoption soars. <https://www.weforum.org/impact/iot-security-keeping-consumers-safe/>
- [6] Y. Sun, F. P. Lo and B. Lo, "Security and Privacy for the Internet of Medical Things Enabled Healthcare Systems: A Survey".
- [7] A. Ghubaish, T. Salman, M. Zolanvari, D. Ünal, A. Al-Ali and R. Jain, "Recent Advances in the Internet-of-Medical-Things (IoMT) Systems Security".
- [8] A. L. N. Al-hajjar and A. K. M. Al-Qurabat, "An overview of ML methods in enabling IoMT-based epileptic seizure detection".
- [9] A. Mawgoud, A. I. Karadawy and B. Tawfik, "A Secure Authentication Technique in Internet of Medical Things through ML".
- [10] D. Koutras, G. Stergiopoulos, T. K. Dasaklis, P. Kotzanikolaou, D. Glynos and C. Douligeris, "Security in IoMT Communications: A Survey".
- [11] F. S. Alsubaei, A. Abuhusein and S. G. Shiva, "A Framework for Ranking IoMT Solutions Based on Measuring Security and Privacy".
- [12] S. Vishnu, S. R. J. Ramson and R. Jegan, "Internet of Medical Things (IoMT) - An overview".
- [13] D. Koutras, G. Stergiopoulos, T. K. Dasaklis, P. Kotzanikolaou, D. Glynos and C. Douligeris, "Security in IoMT Communications: A Survey".
- [14] Zhao, G., Chen, H. & Wang, J. A lightweight block encryption algorithm for narrowband internet of things. Peer-to-Peer Netw. Appl. 16, 2775–2793 (2023). <https://doi.org/10.1007/s12083-023-01559-w>
- [15] A. Ghubaish, T. Salman, M. Zolanvari, D. Ünal, A. Al-Ali and R. Jain, "Recent Advances in the Internet-of-Medical-Things (IoMT) Systems Security".
- [16] A. A. Mawgoud, A. I. Karadawy and B. Tawfik, "A Secure Authentication Technique in Internet of Medical Things through Machine Learning". R, U. M., P, R. S., Gokul Chandrasekaran, & K, M. (2024). Assessment of Cybersecurity Risks in Digital Twin Deployments in Smart Cities. International Journal of Computational and Experimental Science and Engineering, 10(4). <https://doi.org/10.22399/ijcesen.494>
- [17] V. A. Thakor, M. A. Razzaque and M. R. A. Khandaker, "Lightweight Cryptography Algorithms for Resource-Constrained IoT Devices: A Review, Comparison and Research Opportunities," in IEEE Access, vol. 9, pp. 28177-28193, 2021,
- [18] P. Suryateja, et al., "A Survey on Lightweight Cryptographic Algorithms in IoT," 2024. — comprehensive taxonomy and benchmarks. <cit.iict.bas.bg>
- [19] O. Sabri, "A Lightweight Encryption Method for IoT- Based Healthcare Applications: A Review and Future Prospects," *IoT (MDPI)*, 2025. — healthcare-focused comparative review. <MDPI>
- [20] M. Sönmez Turan (ed.), *NIST SP 800-232 — Ascon- Based Lightweight Cryptography*, NIST, 2025. — NIST standard document for Ascon family. <NIST Publications>
- [21] J. Damodharan, "High Throughput PRESENT Cipher Hardware Architecture for Medical IoT Applications," *MDPI/Conference/Journal*, 2023. — PRESENT hardware accelerator for medical IoT. <MDPI+1>
- [22] J. Soto-Cruz et al., "A Survey of Efficient Lightweight Cryptography for Power-Constrained Microcontrollers," *Technologies*, 2025. — implementation benchmarks on microcontrollers. <MDPI>
- [23] A. M. Rasheed & S. K. Mathusoothana, "Efficient lightweight cryptographic solutions for enhancing data security in healthcare," *Frontiers in Computer Science*, 2025. — new hybrid/chaos-based proposals for IoT healthcare.
- [24] ResearchGate / conference paper, "High-Throughput PRESENT Cipher Hardware Framework — Medical IoT," 2023. — implementation case study and throughput results.

SYNTHESIS AND EVALUATION OF HYDROXY APATITE AND TRI-CALCIUM PHOSPHATE BASED COMPOSITES FOR BIO-MEDICS: A STATE-OF-THE-ART SURVEY

P.V. Rajesh¹, P. Dhinesh Kumar², S.Gopi Krishnan², J Praveen^{2*}

¹Assistant Professor, Department of Mechanical Engineering, Saranathan College of Engineering, Tiruchirappalli-620012, Tamil Nadu.

²U.G. Scholar, Department of Mechanical Engineering, Saranathan College of Engineering, Tiruchirappalli-620012, Tamil Nadu.

*Corresponding Author mail ID: jkpraveen2655@gmail.com

ABSTRACT

The project involves the synthesis and evaluation of Hydroxyapatite (HA) and Tri-Calcium Phosphate (TCP) based composites for biomedical applications. HA and TCP are known for their biocompatibility and ability to promote bone regeneration and integration, making them ideal materials for medical implants and tissue engineering. The research explores innovative synthesis methods to optimize the composition, structure, and morphology of HA and TCP-based composites, aiming to improve their performance in biomedical contexts. The study employs advanced characterization techniques, such as X-ray diffraction (XRD), scanning electron microscopy (SEM), and Fourier-transform infrared spectroscopy (FTIR), to analyse the physicochemical properties and structural features of the synthesized composites. Additionally, *in vitro* and *in vivo* evaluations are conducted to assess the biocompatibility, bioactivity, and osteoconductivity of the materials, providing insights into their suitability for bone repair and regeneration. Preliminary results suggest that the synthesized composites exhibit promising properties, including mechanical strength, stability, and favourable interactions with biological tissues. This research contributes valuable knowledge to the field of biomaterials science and holds potential for the development of advanced materials that can enhance patient outcomes in orthopaedic, dental, and other biomedical treatments. By advancing our understanding of HA and TCP-

based composites, this project aims to drive innovation in the development of safe, effective, and reliable biomaterials for medical use.

Keywords: biocompatibility; implants; composites; mechanical strength; biomaterials

1. INTRODUCTION

In the realm of materials science and biomedical engineering, Tri-Calcium Phosphate (TCP) stands as a cornerstone due to its exceptional biocompatibility and versatile properties. Among its manifold applications, TCP serves as a fundamental precursor for the synthesis of two paramount biomaterials: Hydroxyapatite (HA). These materials hold immense promise in various fields, particularly in orthopedics, dentistry, and regenerative medicine, owing to their ability to mimic the structure and functionality of natural bone tissue. As the prevalence of bone-related disorders and injuries continues to escalate globally, the demand for synthetic bone substitutes that can seamlessly integrate with the body and promote tissue regeneration has become increasingly urgent. Consequently, there exists a pressing need for continuous research and innovation in the synthesis, characterization, and evaluation of HA and based materials derived from TCP. By unraveling the intricacies of these biomaterials, scientists and engineers aim to unlock new avenues for enhancing bone repair, implant durability, and overall patient outcomes.

The significance of this research endeavor lies in its potential to revolutionize the landscape of bone tissue engineering and regenerative medicine. HA possess unique properties that make them highly desirable for biomedical applications. Not only are they biocompatible, ensuring minimal adverse reactions within the body,

but they also exhibit excellent bioactivity, stimulating bone cell growth and facilitating osseointegration—the process by which implants bond with surrounding bone tissue. Additionally, their tunable properties enable customization to meet specific clinical requirements, whether it be adjusting mechanical strength, surface roughness, or degradation rate. This versatility opens doors to a myriad of applications, from load-bearing implants to drug delivery systems tailored for localized treatment of bone infections or diseases.

The objectives of this research endeavor are multifaceted, encompassing synthesis, characterization, and evaluation aspects to comprehensively understand the performance and potential applications of HA based compacts derived from TCP. Synthesis techniques will involve exploring

various fabrication methods, such as precipitation, sol-gel, and hydrothermal processes, to optimize the composition, structure, and morphology of the resulting biomaterials. Characterization will entail employing state-of-the-art analytical techniques, including X-ray diffraction (XRD), scanning electron microscopy (SEM), Fourier-transform infrared spectroscopy (FTIR), and energy-dispersive X-ray spectroscopy (EDS), to elucidate the physicochemical properties and microstructural features of the synthesized materials. Evaluation studies will focus on assessing the biocompatibility, bioactivity, and osteoconductivity of HA and based compacts through in vitro cell culture assays and in vivo animal models, aiming to validate their efficacy and safety for clinical translation.

The structure of this research project will follow a systematic approach, comprising several interconnected components that build upon each other to achieve the overarching goals. A thorough review of existing literature will provide a foundation of knowledge and insights into the current state-of-the-art in biomaterials research, highlighting key advancements, challenges, and opportunities in the field. The materials and methods section will detail the experimental procedures, materials utilized, and analytical methodologies employed throughout the study, ensuring reproducibility and transparency in research practices. Results and discussion will present the findings of the research, critically analyzing the data and drawing conclusions regarding the performance and potential applications of HA based materials. Finally, the conclusion and future perspectives section will summarize the key findings, underscore the significance of the research contributions, and propose directions for future investigations to further advance the field of biomaterials science and regenerative medicine.

In essence, the synthesis, characterization, and evaluation of HA compacts from TCP represent a pivotal area of research with profound implications for the development of next-generation biomaterials tailored for bone repair and regeneration. By harnessing the synergistic interplay between materials science, engineering principles, and biological insights, this research endeavor strives to address unmet clinical needs, improve patient outcomes, and pave the way for transformative innovations in musculoskeletal healthcare. Through collaborative efforts and interdisciplinary approaches, we endeavor to unlock the full potential of HA based materials and catalyze their translation from the laboratory bench to the bedside, ultimately shaping a brighter future for bone health and well-being.

Significance of the Research

The synthesis, characterization, and evaluation of HA based compacts from TCP hold significant importance in the field of biomaterials and regenerative medicine for several reasons:

1. **Biocompatibility and Bioactivity:** HA are biocompatible materials that exhibit excellent bioactivity, promoting bone cell attachment, proliferation, and differentiation. These properties are essential for the successful integration of synthetic bone substitutes with the host tissue.
2. **Osteoconductivity:** The ability of HA to support bone ingrowth and new bone formation makes them promising materials for bone grafting and implant coatings, facilitating the regeneration of damaged or diseased bone tissue.
3. **Tailored Properties:** By modifying the composition and structure of HA, it is possible to tailor their mechanical, biological, and degradation properties to suit specific clinical applications. This customization enhances the versatility and efficacy of these biomaterials in diverse biomedical scenarios.
4. **Drug Delivery:** HA matrices can serve as effective carriers for various therapeutic agents, including growth factors, antibiotics, and anti-inflammatory drugs. Their porous nature allows for controlled release kinetics, enabling targeted delivery and prolonged efficacy.

2. LITERATURE REVIEW

H. K. Abd El-Hamid et.al- The study aimed to improve the properties of antibacterial drug delivery Ca-

phosphate cements by incorporating bioactive glass. The cements were prepared from beta-tricalcium phosphate cement (β -TCP) and BG, with different percentages of BG. The composite cements were characterized using XRD, FTIR, and TEM. The study also evaluated the bioactivity and biodegradation of the cements in simulated body fluid. The antimicrobial activity was studied against *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae* bacteria, and fungal strains. The 20BG/ β -TCP sample was nominated for practical bone grafting material due to its high mechanical strength and appropriate setting time. A study on composites of bioactive glass (BG) and β -tricalcium phosphate (β -TCP) was conducted. Results showed that 20BG/ β -TCP has superior mechanical characteristics and higher bioactivity behavior due to the development of a bioactive HA layer on the composite's surface. XRD, FTIR, and SEM-EDS studies were used to track the evolution and expansion

of the HA layer. The addition of BG significantly decreased the in vitro drug release rate. The composite cements showed homogeneity with gentamicin, showing potential against *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, and *Staphylococcus aureus*. Cytotoxicity results showed all samples were safe on MG-63 cells up to 50 μ g/mL [1].

L Borkowski et.al- Fluoride-containing apatite ceramics have been shown to have advantages over hydroxyapatite (HAP) in treating bone defects. However, fluoridated apatite has not been extensively tested in vivo. Researchers synthesized fluorapatite (FAP) and applied it to rabbits' bone defects. Results showed the integration between FAP and bone tissue, suggesting its stimulating effect on new bone formation and mineralization. However, after 3 months, there was a transient delay in FAP integration compared to HAP. The reasons for this are unclear, but it was found that 6 months after implantation, similar levels of bone defect regeneration were achieved for both FAP and HAP. The authors present their hypothesis for this phenomenon. The purpose of this paper was to verify the advantage of FAP over HAP in bone tissue regeneration in vivo, based on the literature and our previous in vitro studies. However, the results obtained in the animal study did not conclusively support this hypothesis, as similar bone regeneration effects were obtained for both tested ceramic materials after 6 months of treatment. These observations can be a starting point for further works related to the optimization of FAP properties or the development of a new composite material containing FAP in order to stimulate the bone tissue regeneration process [2].

M. Nagabushanam et.al- Composite materials are engineered substances with unique physical, chemical, mechanical, and aesthetic properties. They are designed to meet specific geometrical, structural, and aesthetic demands in various fields, including construction, automotive, aeronautics, automobile bodies, naval, and biomedical fields. Composites are only getting attention in the field of biomedicine, as they are not yet commercially distributed or produced. Despite years of research and tests, commercial distribution and production of composite-based medical equipment has only been completed. To successfully use these tools, it is essential to understand their intended goals and limitations. Composites are defined by their adaptability to a wide variety of uses. This multi-component material has helped develop biomaterials for service in the replacement or repair of a wide range of human tissues and organs, which is why the biomedical industry has widely adopted its use.

Composite biomaterials' improved physical, mechanical, and chemical properties make them preferable to more conventional metal and polymeric biomaterials in many applications. Studies of advanced composites [3].

Kushendarsyah Saptaji et.al- Composite materials, such as Metal Matrix Composite (MMC), Ceramic Matrix Composite (CMC), and Polymer Matrix Composite (PMC), are versatile and suitable for biomedical applications, particularly implants. They can be produced through various manufacturing processes, such as stir casting and powder metallurgy for MMCs, powder metallurgy for CMCs, and 3-D printing for PMCs. Hydroxyapatite Mineralized Collagen (HMC) is a promising biocompatible composite used in bone regeneration, with a hybrid process combining Additive Manufacturing and CNC Machining. This paper reviews HMC's properties, fabrication method, and existing experiments, while discussing the three types of biocompatible composites and their applications. Composites become one of the materials that are often used for biocompatible applications. Composites are flexible materials with the characteristics that can be modified depend on the purpose. In tissue engineering, Hydroxyapatite Mineralized Collagen is biocompatible composites material to create scaffold utilized in bone regeneration. The composite implementation is used as a bone substitution, because of its properties and characteristics. The composite is a mixture of hydroxyapatite, a protein to help the bone's growth and mineralized collagen, a mineral that occurred [4].

Chiu-Ming Chen et.al- The study combines β -tricalcium phosphate (β -TCP) with polycaprolactone (PCL) to create an injectable bone substitute. The composite has a well-defined architecture and significantly increases compressive strength compared to β -TCP alone. The composites were tested on 18 rabbits with two surgically

created bone defects. Both PCL-TCP and β -TCP showed equivalent clinical effects on osteoconduction property, with PCL-TCP being as effective as the commercially available β -TCP scaffold. This study aimed to develop an injectable and mechanically strong bone substitute, and investigate the efficacy of PCL-TCP scaffolds for bone regeneration in rabbit defect model. Results were compared against β -TCP (Osteocera™), a commercially available composite, which is commonly used in the clinical setting. PCL-TCP was proven to be as effective as the commercially available β -TCP scaffold in terms of the percentage of NFBA measured by

histomorphometric analysis. The incorporation of 25% PCL with β -TCP leads to a relatively smaller pore size and porosity in scaffolds and higher compression load with good shaping characteristics, which also contributes to the clinical effectiveness of the in vivo model. Future work should further explore the bio-affinity, adaptability and the rate of desired degradation of PCL-TCP composite [5].

Sourav Debnath et.al- This study aimed to create titanium oxide-tricalcium phosphate (TiO₂-TCP) composites due to their biocompatible properties. Pure β -tricalcium phosphate was prepared and studied for its physical and mechanical properties. X-ray diffraction and Fourier Transform Infrared Spectroscopy were used to analyze the powder's phases and functional groups. Different amounts of titanium oxide were incorporated into the β -tricalcium phosphate, and XRD analysis was conducted to examine phase changes. SEM was used to measure pore size and grain size. The studies on TiO₂-TCP composites show that they can be developed using powder metallurgy technique with acceptable porosity for implants. Density improves after sintering, indicating bonding between the matrix and second phase. Density slightly affects with temperature and TiO₂ content, with maximum shrinkage observed for 10 Wt.% TiO₂-TCP composite sintered at 10000°C for 2 hours. Hardness increases with sintering temperature and TiO₂ content, with significant improvement up to 5 Wt.% TiO₂ content [6]. **Guoke Tang et.al-** Biomaterials play a crucial role in supporting bone tissue regeneration at the defect site, eventually degrading in situ and replacing it with newly generated bone tissue. These biomaterials have potential clinical applications, from treating non-union fractures to spinal fusion. Bioceramics and polymeric components are used to support bone cell and tissue growth. Recent developments in bone repair materials include hydrogel, nanofiber scaffolds, and 3D printing composite scaffolds. Challenges include engineering biomaterials that can match both the mechanical and biological context of bone tissue matrix and support the vascularization of large tissue constructs. Research is needed to develop new bone regeneration and repair scaffolds, specifically for bone tissue engineering (BTE). The field of BTE has significant potential for scaffold chemical modification, allowing for faster bone formation, increased healing time, and rapid recovery of function. Ideal scaffold materials should have excellent biocompatibility, high mechanical properties, suitable pore interconnectivity, tailored biodegradation, and allowable incorporation of biological cues. Future studies should focus on optimizing nano-scaffold materials and cell interactions [7].

Obinna Anayo Osuchukwu et.al- Hydroxyapatite (HAp) is a promising material for dentistry, orthopedic, and biomedical engineering applications. It has good mechanical properties and can be used for drug delivery, bone defect fillers, and bone cements. This paper reviews literature on synthesis techniques, mechanical properties, property testing, calcination methods, and characterization of HAp from catfish and bovine bones. The paper discusses the necessary features for achieving the best properties for bone grafts and discusses process approaches for fabricating essential microstructures. The paper also discusses the effect of adding foreign material as a strengthening agent on HAp's mechanical properties, suggesting that careful combination of parameters can enhance its mechanical properties. This review paper discusses the extraction of hydroxyapatite (HAp) from bovine and catfish wastes, focusing on its mechanical properties and potential applications [8].

Chin-Shih Hsu et.al- This study evaluates the impact of substituted chloride levels in hydroxyapatite (CIHAP) on surface and mechanical properties and osteoblastic response. The results show that 2-10% CIHAP bioceramics have superior surface hardness, compressive strength, and enhanced wettability compared to other HAp bioceramics. Although slight acidity and chloride release were observed in vitro, these did not significantly affect osteoblast adhesion or proliferation. Therefore, 2-10% CIHAP bioceramics are ideal for implantation due to their superior surface and mechanical properties. This study aimed to evaluate the effect of substituted chloride in CIHAP on the physical and mechanical properties, and in vitro biocompatibility. When the molar ratio of substituted chloride in hydroxyapatite exceeded 25%, significant formation of cracks occurred. Cracks were also present in a few of the 15% and 20% sintered CIHAP samples. The sintered 2%, 6%, and 10% CIHAP bioceramics possessed high relative density (90.3–91.1%), Vickers hardness (~390 HV), and compressive strength [9].

Marta Tavoni et.al- Calcium phosphates (CaPs) are biomaterials that promote bone tissue regeneration, but regenerating critical-sized defects is challenging. To address this, research has focused on developing

bioceramics with enhanced bioactivity, bioresorbability, and mechanical performance. The development of effective bioceramic composite scaffolds requires effective manufacturing techniques to control multi-scale porosity, ensuring osteointegration and bio-competent mechanical performance. This work provides an overview of strategies for developing and optimizing bioceramics and highlights future perspectives in bone tissue regeneration. Calcium phosphates (CaP) are widely used biomaterials

for bone tissue regeneration, with biomimetic scaffolds exhibiting osteogenic, bioresorbable, and antibacterial properties. However, achieving appropriate mechanical properties is crucial for critical-sized bone defects, especially when load-bearing. Traditional methods like partial sintering and 3D printing fail to generate bioactive and effective bone scaffolds. New approaches targeting low-temperature processes, such as chemically induced consolidation of CaP pastes or biomorphic transformation processes, are being developed to create scaffolds with nanocrystallinity and bioactive composition. These results open new perspectives in ceramic science and encourage further research to address unmet clinical problems related to bone regeneration [10].

E. Pietrzykowska et.al- The study proposes a novel biodegradable nanocomposite to treat severe bone loss and large gaps. The technology involves preparing composite powders of 5 wt% or 15 wt% silk fibers and nanocrystalline hydroxyapatite, both biodegradable materials. The powders are prepared using cryomilling and then uniaxially pressed at 1 GPa and 80 °C. This results in a homogenous and dense ceramic matrix composite with good adhesion between the nano-hydroxyapatite and silk fibers. The nano-hydroxyapatite retains its nanocrystalline form, with a compressive strength of 276 MPa and bending strength of 82 MPa, comparable to cortical bone strength. The paper presents a promising solution to increasing the bending and compressive strength of a composite with a ceramic matrix. It is common knowledge that phosphate ceramic is a biocompatible material applied in bone loss treatment, but the brittleness of ceramics limits their application as implants. The aim and novelty of this research is to obtain a composite with the matrix of nanohydroxyapatite, reinforced with silk fibers. The composite obtained with the nanometric structure is characterized by an increased bending strength in relation to pure ceramics. It has been proven using SEM imaging and CMT scanning techniques that the material is anisotropic and homogenous. Adhesion between these two components has been observed and the silk remains in its initial fibrous form (without chemical treatment) [11].

Leszek Borkowski et.al- Calcium phosphates are crucial in treating bone defects. A fluoride-substituted modification of hydroxyapatite (HAP) was synthesized and calcined at different temperatures to improve its biological activity. X-ray diffraction, SEM, and FTIR showed fluoride ions were incorporated into the apatite lattice structure. The highest temperature sintered fluorapatite had the lowest porosity, no internal pores, and highest density. The material was non-toxic and significantly

shortened the doubling time, enhancing osteogenic cell proliferation. This material is proposed for orthopedic applications and bone tissue engineering. Studies have shown that a modified sol-gel method allowed the successful introduction of fluoride into the apatite lattice structure. In FTIR analysis, changes in the OH⁻ bands at 3573 cm⁻¹ and 630 cm⁻¹ as well as OH-F binding at 744 cm⁻¹ were particularly useful for detecting fluoride substitution. Calcination temperature was found to be very important for the pore structure and had a significant impact on ion-adsorbing capacity that influences material bioactivity. FAP materials sintered at 800°C, 1000°C or 1200°C were non-toxic since the viability of pre-osteoblast cells incubated with sample extracts was high. The presented method allowed obtaining fluorapatite, which after calcination at 800 °C was characterized by slow release of fluoride at a level safe for osteoblasts cell line (<1 ppm) and which supported osteoblast proliferation more intensely than fluoride ion solutions of comparable concentrations. Summarizing the above results, FAP sintered at 800°C is a promising implant material for the regeneration of bone tissue defects [12].

Elzbieta Pietrzykowska et.al- This paper presents a two-step technology for obtaining a ceramic matrix composite made from biodegradable polymers like polylactic acid (PLA) and hydroxyapatite nanoparticles (HAP NPs). The first step involves preparing composite granules, which are then consolidated using high pressure. The combined methods result in the highest densification and strongest mechanical properties, with a compressive strength of 374 MPa. The structure of the composite is homogeneous and good adhesion between the inorganic and organic components is observed. This paper addresses the lack of sufficient mechanical strength in bioresorbable composites made from hydroxyapatite nanoparticles in a polymer matrix. A two-step method was developed to create a homogeneous hybrid composite with up to 80 wt% HAP NPs content, without agglomerates. The composite achieved a compressive strength of 375 MPa, a Young's modulus of 7 GPa, and a densification of 99%. The composite had a homogeneous structure and good adhesion between the polymer and ceramic. The technology can be used to form thermosetting composites with high ceramic content at low temperatures, making them promising for bone regeneration applications [13].

Gerry L. Koons et.al - This review discusses the materials design process for bone-tissue engineering, focusing on understanding native bone tissue composition and selecting appropriate biomaterials like polymers, bioceramics, metals, and composites. It covers the implementation strategy, selection of

materials, fabrication methods, and evaluation, highlighting unmet needs and challenges in developing ideal materials for bone-tissue regeneration. Despite advancements in biomaterials for bone-tissue engineering, challenges persist in scientific progress and clinical translation. One obstacle is limited understanding of biomaterials' mechanisms of action and cellular response. Systematic studies are challenging due to the difficulty in decoupling material properties. Meeting multiple conflicting requirements often requires novel materials, such as composite materials. Combining multiple materials is crucial for interfacial-tissue engineering, which requires materials that recapitulate the transition between bone and adjoining tissue types [14].

Sreepadmini Ragunathan - Medical advancements have improved life expectancy, but orthopaedic surgeons face challenges in repairing skeletal bone defects. Nano Hydroxyapatite and natural polymer starch can be used as biomaterials for bone tissue regeneration scaffolds. These materials are synthesized using a Wet Chemical Method, dried, and then molded using a 1:1 solvent casting technique. Cytotoxicity tests show these materials are highly compatible and non-toxic to the body. Wet Chemical Synthesis method is used to produce good yield of the materials in large quantity. Solvent Casting of scaffolding technique is commonly used to produce good mechanical strength and 90% of pores are produced with this Solvent Casting technique with pore range varies from 14 μm to 17 μm . The size of pores is dependent on the size of salt used which is responsible for pore formation. Poly Vinyl Alcohol exhibits high density property and hence it helps to form high density scaffold [15].

A. Kazuz et.al - Composite cements made from α -tricalcium phosphate (α -TCP) and nanostructured fluorapatite (FAp) have been developed for dental use as root canal filling materials. The study investigated the impact of FAp concentration on the properties of the composite cement. The optimal FAp powder was selected based on its properties. Morphology and phase composition were studied using various techniques. The results showed that the transformation of α -TCP into HAp and FAp content significantly influenced the mechanical properties of the composites. In vitro biocompatibility tests showed no cytotoxic effect of α -TCP/FAp composite cements. This study presented the synthesis and characterization of fluorapatite (FAp) powders, as well as composite material α -TCP/FAp, suitable for dental application. Different reaction mixtures and conditions resulted in the hydrothermal synthesis of nanoparticle FAp powder with satisfactory properties, which was used for the processing of the

composites. After the treatment of α -TCP/FAp in simulated body fluid during different periods of time, a new layer of HAp was formed, proving bioactive [16].

Reza Eivazzadeh-Keihan et.al - Tissue is crucial for multicellular organism organization and body shape, and recent advancements in tissue engineering have used nanotechnology-derived materials like metal nanoparticles. These nanoparticles have unique physiochemical properties, including antibacterial effects, shape memory, low cytotoxicity, and osteogenic potential. They also have good mechanical and tensile strength, biocompatibility, and osteogenic potential. This review discusses the basic principles of nanoparticle-based composites and scaffolds, their merits and demerits, and their importance in bone tissue engineering. The review emphasizes the potential of these nanoparticles in bone tissue engineering. Tissue, a crucial structural organ in multicellular organisms, is susceptible to injury and damage from accidents, diseases, and harsh conditions. Biomedical scientists are interested in rehabilitating defective tissues or replacing them with healthy ones. Nanoparticles, which are well-suited for incorporating into biomaterials, have shown promise in tissue engineering. These nanoparticles can be used for cellular imaging, chemical and biological sensing, and bioactive agent delivery. However, concerns about biodegradability and toxicity may limit their widespread use. Advanced nanoparticles, such as gold, silver, iron, aluminum, nickel, copper, and zirconium, could overcome these challenges. Gold and silver nanoparticles have antibacterial activity, low cytotoxic effects, and osteoconductivity, while zirconium-based materials offer higher tensile strength. Copper nanoparticles and ions can control scaffold properties, regulate growth factor activity, and increase cell growth rate. Nickel-based nanoparticles have shape memory, improve cell adhesion and spreading, and aluminum-based nanoparticles have desirable surface pores and shapes for bone tissue growth. Iron-killing nanoparticles can encourage cell growth and have magnetic properties, potentially triggering a revolution in bone tissue engineering [17].

Senthil Kumar Venkatraman & Sasikumar Swamiappan - Bone is a crucial structural component of the human body, providing essential support for its functioning. However, finding suitable biomaterials for bone substitutes is crucial due to the increasing rate of bone fractures and infections leading to osteoporosis. Designing materials with good apatite deposition ability, faster dissolution rate, superior resorbability, high

mechanical strength, and significant bactericidal activity is challenging.

Silicates have emerged as an alternative biomaterial for bone graft substitutes, but their properties differ due to their composition and methodology. Calcium, magnesium, and silicon are key factors in bone mineral formation and metabolism. This review discusses the relationship between composition and activity of calcium and magnesium-based silicates and their potential for hard tissue engineering applications. Calcium and magnesium-based tertiary silicates can be used as bone graft substitutes due to their lack of bioactivity or mechanical strength. However, there is a need for more research on other silicates for biomedical applications. Calcium silicates exhibit distinct bioactivity and excellent mechanical stability in-vitro, while magnesium silicates show excellent stability. To meet the specific requirements of different types of bone, the mechanism of bioactivity and mechanical strength must be thoroughly investigated. The method of synthesis should be carefully chosen to achieve optimum surface and bulk material properties for specific tissue engineering applications. Composite materials with silicates as fillers, reinforcement, and biodegradable polymers can also be used to find suitable candidates for specific applications. Although silicates are brittle and polymers are flexible, composites should be made with them to overcome the current need for biomedical applications [18].

Al Bahrawy M - As a conclusion that HA as tissue engineering scaffold material has many encouraging properties with minimal drawbacks which can be overcome in the future by creating composite scaffolding material using HA with certain fabrication criteria incorporated in organic or inorganic polymers, using the new advancement that we had in the polymerization technology to solve the HA drawback of weak biodegradability [19].

Chin-Shih Hsu - This study investigates the use of a composite coating on titanium alloys like Ti-6Al-4V or Ti64 for load-bearing implants. The coating, which contains chlorine substituted hydroxyapatite (ClHAP) and polydopamine (Pda), promotes osteoclastic activity and bioactivity, as evidenced by osteoblast proliferation. The coating approach may enhance osseointegration in vivo, enhancing the biocompatibility and resistance of Ti/alloys. Ti64-based materials are widely used in load bearing biomedical applications because of their material properties, but they suffer from poor osseointegration with native bone due to their surface interactions with mineralized tissue. In this study ClHAP/Pda coatings were prepared and evaluated relative to uncoated Ti64 and HA/Pda coated Ti64 as a step

towards designing Ti64 coatings that improve osseointegration. ClHAP particles were well dispersed and anchored on the Ti64 surface by Pda [20].

K. Hayashi et.al - This study investigates the effects of honeycomb blocks (HCBs) on bone formation and maturation in orthopedics and dentistry. Three types of HCBs were fabricated: hydroxyapatite (HAp), β -tricalcium phosphate (TCP), and carbonate apatite (CO3Ap). The HC structure was chosen to promote cell penetration and tissue in-growth. The HCBs had similar macroporous structures, with uniformly distributed macropores. The compressive strengths of CO3Ap, HAp, and β -TCP HCBs were higher than those of commercial scaffolds with intricate three-dimensional or unidirectional macroporous structures. Bone maturation was faster in CO3Ap HCB grafting than in β -TCP and HAp HCB grafting. The differences in bone maturation and formation were likely due to the disparity in concentrations of calcium ions surrounding the HCBs, which were dictated by the inherent material resorption behavior and mechanism. The microporous structure of HC struts may also contribute to these differences. This study successfully fabricated HAp, β -TCP, and CO3Ap HCBs, which showed higher compressive strength than commercial CaP scaffolds. Bone maturation was faster in CO3Ap HCBs than HAp and β -TCP HCBs, and bone formation was higher in CO3Ap HCBs. This is due to adequate calcium ion release during resorption, making them suitable scaffolds. Inferior bone maturation and formation in HAp and β -TCP HCB grafting were caused by no resorption of HAp HCBs and overly rapid dissolution of β -TCP HCBs [21].

Sandeep Gupta et.al - Filling bone defect after debridement of infected nonunion is an orthopedic challenge. Since the volume of autologous bone graft available is limited, allograft, demineralized bone matrix, and calcium phosphate ceramic-based bone graft substitutes have come up as potential autograft expanders. This study was conducted to analyze the use of beta tri-calcium phosphate (B-TCP)-based composite ceramic as autologous bone-graft expander in the management of postinfective segmental gap non-union of long bones managed with two-stage Masquelet's technique. B-TCP is an efficacious and safe autologous bone graft expander in Masquelet's two-stage management of post infective segmental gap nonunion of long bones. Patients should be counseled regarding increased risk of non-union and need for repeat grafting with its use, especially if they are smokers or site of involvement is tibia [22].

Chih-Hsiang Fang et.al - The study developed a bio-composite bone cement using tri-calcium phosphate and chitosan as additives to acrylic bone cement. The aim was to improve osteo-integration compared to pure polymethyl methacrylate cement. The bio-composite showed lower curing temperature, longer setting time,

higher weight loss and porosity after degradation, lower compressive Young's modulus, and ultimate compressive strength compared to pure polymethyl methacrylate cement. Cell proliferation tests showed it was non-cytotoxic, and in vivo tests revealed it was more osteo-conductive. The study suggests that this new bio-composite cement could be a promising alternative to pure polymethyl methacrylate cement. The results indicated that the modified chitosan/tri- calcium phosphate/polymethyl methacrylate bio-composites bone cement could be degraded gradually and create rougher surfaces that would be beneficial to cell adherence and growth. This new bio- composite bone cement has potential in clinical application. Our future studies will focus on long-term implantation to investigate the stability of the bio-composite bone cement in long-term implantation [23].

Kristina Glenske - Dental medicine primarily focuses on bone tissue regeneration, with calcium phosphate (CaP)-based substitute materials used for this purpose. These materials, both natural and synthetic, influence regenerative processes like osteoblast growth and osteoclastic resorption. However, the required material characteristics for the required ratio between new bone tissue formation and degradation have not been found. Metals and metal ions are also used in bone regeneration, playing functional roles in the physiological cellular environment and bone healing process. This review focuses on frequently used metals as integral parts of bone regeneration materials. Existing bone substitute materials only offer osteoconductive healing capabilities, and new tissue engineering strategies are not yet applicable in daily clinical routines. The addition of metal ions to existing materials can alter inflammation, foreign body response, and material durability. The availability and cost of suitable bone grafting material are crucial for the aging population. The concentration of metal ions is crucial for bone formation, and metals can be incorporated into scaffolds for early osteoblast differentiation. Further research is needed to develop suitable biomaterials for patients [24].

Shu-Hsien Huang et.al - β -Tricalcium phosphate (β -TCP) is an osteoconductive material which has been used for clinical purposes for several years, as is polycaprolactone (PCL), which has already been

approved for a number of medical and drug delivery devices. In this study we have incorporated various concentrations of β -TCP into PCL with the aim of developing an injectable, mechanically strong, and biodegradable material which can be used for medical purposes without organic solvents. This study assesses the physical and chemical properties of this material, evaluates the in vitro bioactivity of the PCL/ β -TCP composites, and analyzes cell proliferation and osteogenic differentiation when using human bone marrow mesenchymal stem cells (hBMSCs). PCL/TCP is biocompatible with hBMSCs. It not only promotes proliferation of hBMSCs but also helps to differentiate reparative hard tissue. We suggest 50% (weight) PCL-containing β -TCP biocomposites as the best choice for hard tissue repair applications [25].

3. FUTURE SCOPE

The synthesis and evaluation of Hydroxyapatite (HA) and Tri-Calcium Phosphate (TCP) based composites for biomedical applications hold immense potential for advancing the field of regenerative medicine and biomaterials science. The following future directions highlight the areas of opportunity and development in this research domain:

1. **Improved Synthesis Techniques:** Exploring novel synthesis methods such as electrospinning, 3D printing, and other advanced fabrication techniques can lead to more precise control over the properties of HA and TCP composites. This could facilitate the creation of materials with tailored mechanical strength, porosity, and degradation rates to suit specific clinical applications.
2. **Customized Material Design:** Developing composites with adjustable properties allows for personalized medical treatments. Materials can be engineered to match individual patient needs, such as varying degrees of bioactivity, drug loading capabilities, or specific mechanical properties required for different anatomical locations.
3. **Long-term Performance and Safety Studies:** Further in vivo studies and long-term evaluations are necessary to validate the safety, biocompatibility, and efficacy of HA and TCP- based composites in clinical settings. These studies will provide valuable data on the durability and long-term integration of these materials in the human body.
4. **Multi-functional Biomaterials:** HA and TCP-based composites can be engineered to possess multi-functional properties such as antibacterial and anti-inflammatory effects. This can be achieved through surface modifications, incorporation of metal ions, or other additives that enhance the composite's therapeutic capabilities.
5. **Interdisciplinary Collaborations:** Collaborations between materials scientists, biomedical engineers, clinicians, and industry partners will be essential for advancing the application of HA and TCP-based composites in real-world medical scenarios. These partnerships can facilitate the transition from

laboratory research to clinical trials and eventual commercial applications.

- 6. Regulatory Compliance and Standardization:** As the development of HA and TCP-based composites progresses, establishing standardized protocols for synthesis, characterization, and evaluation will be crucial. Ensuring regulatory compliance and quality control will pave the way for safe and effective use in clinical settings.

In summary, the future scope of HA and TCP-based composites is promising, with potential advancements in material design, drug delivery, and clinical applications. By continuing to explore innovative approaches and fostering interdisciplinary collaborations, this field can lead to transformative breakthroughs in the treatment of musculoskeletal disorders and other medical challenges.

4. CONCLUSION

The research on the synthesis and evaluation of Hydroxyapatite (HA) and Tri-Calcium Phosphate (TCP) based composites offers a comprehensive exploration of their potential in biomedical applications. This study has provided critical insights into the methods of synthesizing these composites and has contributed significantly to our understanding of their properties, performance, and compatibility with biological systems. The synthesis of HA and TCP-based composites has been optimized through the careful selection of methods, allowing for precise control over their composition and structure. This has resulted in materials with desirable mechanical strength, porosity, and bioactivity, making them ideal

candidates for bone regeneration and tissue engineering applications. The extensive characterization and evaluation of these composites have demonstrated their ability to integrate seamlessly with biological tissues, promoting bone growth and healing. Furthermore, the study's findings highlight the potential for these composites to serve as versatile platforms for drug delivery systems, offering sustained release of therapeutic agents directly to the site of injury or implantation. This capability could revolutionize treatments for musculoskeletal disorders, providing more targeted and efficient therapies. However, while the research has established a strong foundation, there are still areas that require further investigation. Long-term in vivo studies are needed to assess the composites' durability, stability, and potential for clinical applications over extended periods. These studies will provide valuable data to guide future developments and ensure the safe and effective use of HA and TCP-based composites in medical settings. Looking ahead, interdisciplinary collaborations will play a crucial role in advancing the application of these composites in clinical practice. Partnerships between materials scientists, engineers, clinicians, and industry stakeholders will facilitate the transition from research to real-world applications, ultimately leading to improved patient outcomes. The synthesis and evaluation of HA and TCP-based composites represent a promising area of research with the potential to impact various fields of medicine. By continuing to build on these findings and addressing the remaining challenges, this work can contribute to the development of innovative biomaterials that enhance the quality of life for patients undergoing orthopedic and dental treatments, among others.

5. REFERENCES

- [1] Abd El-Hamid, H. K., Farag, M. M., Abdelraof, M., & Elwan, R. L. (2024). Regulation of the antibiotic elution profile from tricalcium phosphate bone cement by addition of bioactive glass. *Scientific Reports*, 14(1), 2804.
- [2] Borkowski, L., Jojczuk, M., Belcarz, A., Pawlowska-Olszewska, M., Kruk-Bachonko, J., Radzki, R., & Ginalska, G. (2023). Comparing the Healing Abilities of Fluorapatite and Hydroxyapatite Ceramics in Regenerating Bone Tissue: An In Vivo Study. *Materials*, 16(17), 5992.
- [3] Nagabushanam, M., Devade, K., Reddy, G. A., Goud, B. N., Sayed, R. M., Sood, S., & Sonia, P. (2023). Advance biomedical engineering—A fundamental review of composite materials and its applications. *Materials Today: Proceedings*.
- [4] Saptaji, K., Hardiyasanti, D. C., Ali, M. F., Frandito, R., & Dewi, T. K. (2022). Potential Applications of Hydroxyapatite-Mineralized-Collagen Composites as Bone Structure Regeneration: a Review. *Journal of Science and Applied Engineering*, 5(1), 3577.
- [5] Chen, C. M., Chen, S. M., Lin, S. F., Liang, H. C., & Wu, C. C. (2021). Clinical efficacy of polycaprolactone β -calcium triphosphate composite for osteoconduction in rabbit bone defect model. *Polymers*, 13(15), 2552.

- [6] Debnath, S., Basu, S., & Pramanick, A. K. (2021). Development and study on various properties of titanium oxide-tri calcium phosphate composites through powder metallurgy technique. *IOSR J*, 2, 1- 6.
- [7] Tang, G., Liu, Z., Liu, Y., Yu, J., Wang, X., Tan, Z., & Ye, X. (2021). Recent trends in the development of bone regenerative biomaterials. *Frontiers in Cell and Developmental biology*, 9, 665813.
- [8] Osuchukwu, O. A., Salihi, A., Abdullahi, I., Abdulkareem, B., & Nwannenna, C. S. (2021).

Synthesis techniques, characterization and mechanical properties of natural derived hydroxyapatite scaffolds for bone implants: A review. *SN Applied Sciences*, 3, 1-23.

- [9] Hsu, C. S., Haag, S. L., Bernards, M. T., & Li, Q. (2021). Effects of chloride substitution on physical, mechanical, and biological properties of hydroxyapatite. *Ceramics International*, 47(9), 13207-13215.
- [10] Tavoni, M., Dapporto, M., Tampieri, A., & Sprio, S. (2021). Bioactive calcium phosphate-based composites for bone regeneration. *Journal of Composites Science*, 5(9), 227.
- [11] Pietrzykowska, E., Małysa, M., Chodara, A., Romelczyk-Baishya, B., Szlązak, K., Świąszkowski, W., & Lojkowski, W. (2021). Biodegradable ceramic matrix composites made from nanocrystalline hydroxyapatite and silk fibers via crymilling and uniaxial pressing. *Materials Letters*, 293, 129672.
- [12] Borkowski, L., Przekora, A., Belcarz, A., Palka, K., Jozefaciuk, G., Lübek, T., ... & Ginalska, G. (2020). Fluorapatite ceramics for bone tissue regeneration: Synthesis, characterization and assessment of biomedical potential. *Materials Science and Engineering: C*, 116, 111211.
- [13] Pietrzykowska, E., Romelczyk-Baishya, B., Wojnarowicz, J., Sokolova, M., Szlązak, K., Swieszkowski, W., ... & Lojkowski, W. (2020). Preparation of a ceramic matrix composite made of hydroxyapatite nanoparticles and polylactic acid by consolidation of composite granules. *Nanomaterials*, 10(6), 1060.
- [14] Koons, G. L., Diba, M., & Mikos, A. G. (2020). Materials design for bone-tissue engineering. *Nature Reviews Materials*, 5(8), 584-603
- [15] Ragunathan, S., Govindasamy, G., Raghul, D. R., Karuppaswamy, M., & VijayachandraTogo, R. K. (2020). Hydroxyapatite reinforced natural polymer scaffold for bone tissue regeneration. *Materials Today: Proceedings*, 23, 111-118.
- [16] Kazuz, A., Radovanović, Ž., Veljović, D., Kojić, V., Miletić, V., Petrović, R., & Janačković, D. (2020). α -Tricalcium phosphate/fluorapatite based composite cements: Synthesis, mechanical properties, and biocompatibility. *Ceramics International*, 46(16), 25149-25154.
- [17] Eivazzadeh-Keihan, R., Bahojb Noruzi, E., Khanmohammadi Chenab, K., Jafari, A., Radinekiyan, F., Hashemi, S. M., ... & Hamblin, M. R. (2020). Metal-based nanoparticles for bone tissue engineering. *Journal of Tissue Engineering and Regenerative Medicine*, 14(12), 1687-1714.
- [18] Venkatraman, S. K., & Swamiappan, S. (2020). Review on calcium-and magnesium-based silicates for bone tissue engineering applications. *Journal of Biomedical Materials Research Part A*, 108(7), 1546-1562.
- [19] Al Bahrawy, M. (2020). Hydroxyapatite as a Biomaterial to Improve Stem Cell-Related Tissue Engineering. *ESC*, 20, 22.
- [20] Hsu, C. S., Haag, S. L., Bernards, M. T., & Li, Q. (2020). Evaluation of chlorine substituted hydroxyapatite (ClHAP)/polydopamine composite coatings on Ti64. *Colloids and Surfaces B: Biointerfaces*, 189, 110799.
- [21] Hayashi, K., Kishida, R., Tsuchiya, A., & Ishikawa, K. (2019). Honeycomb blocks composed of carbonate apatite, β -tricalcium phosphate, and hydroxyapatite for bone regeneration: effects of composition on biological responses. *Materials Today Bio*, 4, 100031.
- [22] Gupta, S., Malhotra, A., Jindal, R., Garg, S. K., Kansay, R., & Mittal, N. (2019). Role of beta tri- calcium phosphate-based composite ceramic as bone-graft expander in Masquelet's-induced membrane technique. *Indian Journal of Orthopaedics*, 53, 63-69
- [23] Fang, C. H., Lin, Y. W., Sun, J. S., & Lin, F. H. (2019). The chitosan/tri-calcium phosphate bio-composite bone cement promotes better osteo-integration: An in vitro and in vivo study. *Journal of Orthopaedic Surgery and Research*, 14, 1-9.
- [24] Glenske, K., Donkiewicz, P., Köwitsch, A., Milosevic-Oljaca, N., Rider, P., Rofall, S., ... & Barbeck, M. (2018). Applications of metals for bone regeneration. *International journal of molecular sciences*, 19(3), 826.
- [25] Huang, S. H., Hsu, T. T., Huang, T. H., Lin, C. Y., & Shie, M. Y. (2017). Fabrication and characterization of polycaprolactone and tricalcium phosphate composites for tissue engineering applications. *Journal of Dental Sciences*, 12(1), 33-43.

A Study on the Implementation of Artificial Intelligence in Payroll Process To Enhance Organizational Performance

Rohith T

Student

II -MBA, Department of Management Studies,

Vel Tech Rangarajan Dr. Sagunthala R&D

Institute of Science and Technology,

(Deemed to be university), Chennai, India

Dr. Kotteshwaran D

Assistant professor

MBA, Department of Management Studies,

Vel Tech Rangarajan Dr. Sagunthala R&D

Institute of Science and Technology,

(Deemed to be university), Chennai, India

Abstract:

The swift evolution of Artificial Intelligence (AI) has altered different organizational functions, and payroll is no exception. This research analyse the utilization of AI in payroll functions and assesses it in terms of the enriched overall performance of the organization. The study suggests that AI-enabled tools improve payroll's efficiencies in the organization by automating repetitive tasks, preventing human error, improving precision, assuring compliance, and expediting salary cycle processing. AI-enabled analytics provide the capability to better forecast payroll costs, coverage of payroll anomalies, and real-time decision-making that supports financial planning and supports operational effectiveness. The research also examines elements of success that AI advances in conjunction with HR systems that enhance data management, and improve employee satisfaction through timely, accurate, and error-free payroll processing, while reducing the administrative burden on HR staff. A descriptive research design was employed, collecting both primary and secondary data, to gauge employee perception, organizational readiness, and performance of AI-enabled payroll systems. The analysis indicated, organizations that utilized AI within their payroll systems experienced a higher level of accuracy over payroll process, in reduced operational costs, improved compliance, and optimized resources. Additionally, AI instilled enhanced transparency while also cultivating a culture of trust, that assured payroll outcomes every pay period. Despite certain limitations associated with factors that include high-cost expenditure, personnel competencies, stakeholder engagement, and systems integration with existing processes, the benefits outweigh the limitations.

Keywords:

Data Security, Payroll Efficiency, Employee Satisfaction, Payroll Accuracy, Compliance Management, Digital Transformation, Payroll Efficiency

Introduction

Artificial intelligence (AI) is becoming a leading force across industries, include more conservative sector. With digital transformation firmly underway, organizations are increasingly leveraging state-of-the-art technologies to improve operations, lower costs, and improve performance. Artificial Intelligence (AI) has now emerged as a robust technology that could be radically impact how many traditional business functions operate, like payroll. Payroll is a specialization in Human Resource Management that entails several complex tasks, including salary calculation, statutory deductions, compliance with labour laws, and even the timely payment of wages. Manual payroll, as well as more traditional payroll systems, can sometimes create errors and become ineffective, which ultimately lowers the employee experience in the workplace, as well as the productivity of the organization. There are many opportunities for utilizing AI in payroll for innovations including automated processes, enhancements in accuracy, faster processes, and data-assisted organizational decisions. AI payroll systems can analyze large data sets, identify anomalies, Legislative compliance and reduce administration time. If an organization wants to be competitive, AI payroll systems provide a competitive edge by increasing transparency and statutory risk reductions while enhancing workflow efficiency. This study looks at how AI implementation impacts the payroll process and creates improvements in organizational performance. It also looks at employee perceptions of payroll, organizational readiness for AI in payroll, and the efficacy of AI payroll options for ensuring accuracy, timeliness, and legislative compliance.

Objectives of the Research:

- To investigate how effective AI can be for payroll automation and reducing manual work
- To analyze the effects AI-led payroll systems have on accuracy, timeliness, and compliance
- To determine how implementing AI into payroll improves overall organizational performance
- To examine employees perceptions and satisfaction towards AI-led payroll systems
- To explore the challenges and limitations experienced by organizations in the implementation of AI in payroll processes.

Review of Literature :

According to **Smith (2019)**, Artificial Intelligence is a disruptive technology that effectively automates administrative aspects of Human Resource Management, including payroll management. **Johnson and Lee (2020)** state that automated payroll solutions utilize AI-based machine-learning technology to improve the accuracy of salary calculations by emphasizing the reduction of repetitive calculation errors. **Brown (2018)** notes that payroll systems using automation can improve compliance with statutory regulations as these regulations are continuously updated in the system, decreasing overall litigation risk for the organization. As investigated by **Williams and Carter (2021)**, Human Resource Management systems using digital technology not only can improve internal efficiencies for the organization but ensure that employees can verify and review records in real time. **Davis (2020)** emphasizes the ability of AI employment models to not only allow for accurate payroll but

increased employee satisfaction as they can confirm their payment history and payment transactions in employee reports. According to **Kumar and Reddy (2021)**, organizations in India utilizing AI technology have optimized their workforce management while saving costs and improving decision-making using predictive analytics. **Thompson (2019)** explained one of the benefits of using AI in payroll systems is that organizations can identify fraudulent activities and anomalies based on the unique statistical patterns that are recognized through advanced AI technology. **Anderson and Miller (2022)**, when using AI in payroll, organizations yield higher productivity benefits from a reduced administrative burden for HR employees. However, as reported by **Roberts (2020)**, the cost of implementation, securing data privacy, and the service skills needed to implement automation are significant issues in shovel-ready technologies. Some of these issues may represent appreciable barriers to the adoption of automation technologies. Taken together, the studies show that organizational performance can greatly improve if AI technologies are adopted carefully in payroll processes.

Scope of the Study:

- The focus of the study is to gain an understanding of how widely Artificial Intelligence is implemented in payroll processes within organizations.
- Some of the payroll functions included are salary computations, integration of attendance, compliance with statutory requirements, error detection, and report generations.
- The study also looks at how AI-driven payroll systems affect organizational performance in areas of accuracy, efficiency, timeliness, and cost savings.

Limitations of the Study:

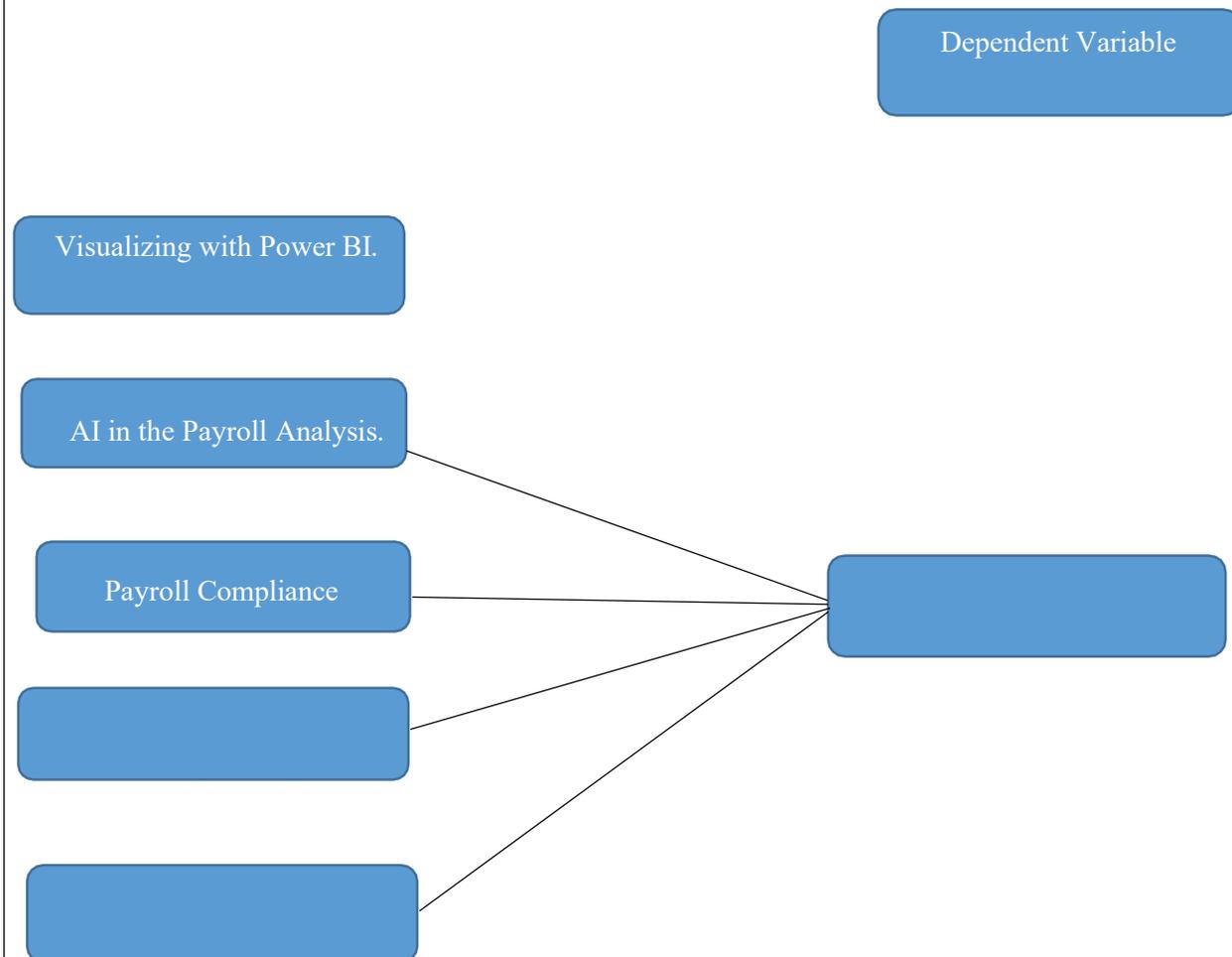
- The research is confined to a subset of organizations and the results may not encompass all industries or sectors utilizing AI in payroll systems.
- The sample size may also be a limitation which may adversely influence the generalizability and accuracy of the findings.
- Employee responses may be biased, as respondents may have provided answers that they believed were socially acceptable rather than their actual experiences.
- The research examines only payroll processes without considering the influence of AI in other HR processes, such as hiring, training or performance review.

Research Methodology:

The study uses a descriptive research design to investigate the use of Artificial Intelligence in payroll processes and its effect on performance outcomes of the organizations. Data collected through both primary and secondary data for a richer analysis. Primary data was collected by using questionnaires that were structured and distributed to employees, HR and payroll administrators to understand their perceptions, experience and knowledge of AI Payroll Systems. A simple random sampling technique was used to select respondents to minimize bias, representing reliable and trustworthy research. Secondary data was collected from journals, research articles, company reports, and online sources related to artificial intelligence, payroll automation, and the performance of the organization. The data was analyzed systematically using statistical tools like percentage, charts, and tables to maximize the analysis of responses. Descriptive analysis was used to further explore trends, benefits, challenges and overall effectiveness of AI implementations in payroll operations.

Research Design:

This study employs a descriptive research design for the evaluation of using Artificial Intelligence in payroll practices and how it contributes to organizational performance. A descriptive design is appropriate because the intent of the researcher is to systematically outline practices, attitudes, and efficiency of AI-enabled payroll practices within organizations. The study concentrated on how AI tools were used; the amount of automation, if any, that was completed; and the degree to which AI improved accuracy, efficiency and compliance. Quantitative and qualitative methodologies were explored in order to fully understand the issue. Quantitative data was gathered through structured questionnaires for measuring employee perceptions and qualitative information was gathered through secondary sources such as journal articles, research reports and organizational documents. The descriptive design assists in identifying the current trends, challenges and overall impact of AI adoption in payroll operations and, therefore, contribute to the study being grounded in reality for organizational based practices.



SAMPLING METHOD:

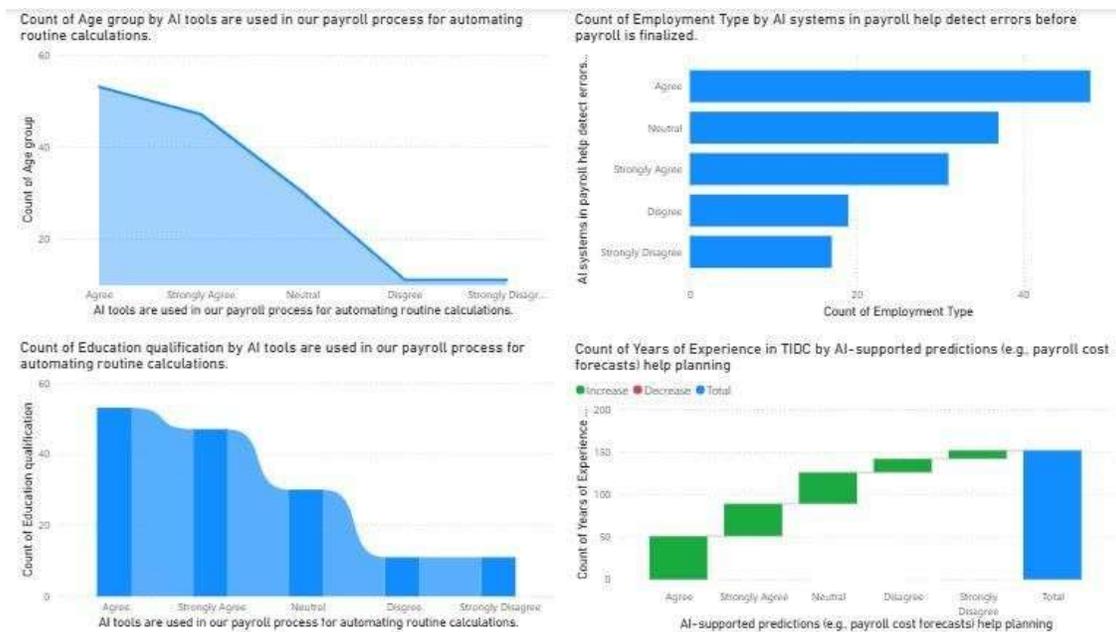
The research uses a non-probability convenience sampling process, to include respondents from personnel included in the system or that have been payroll processed. The sampling frame is comprised of Payroll list, and the sampling unit is an individual employee.

SAMPLING SIZE:

The Sample dimension of this research study restricted to 152 respondents.

Power Bi analysis:

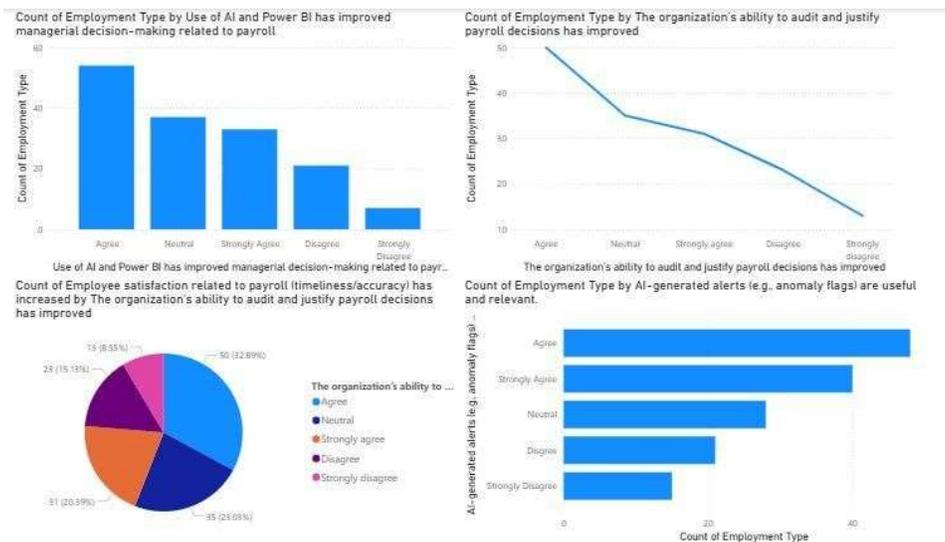
Payroll processes AI Adoption and Usage Analysis.



Interpretation:

As indicated by the dashboard, the employees of various age groups, education, employment type, and different categories of experience have a positive attitude towards AI implementation in payroll process. A majority of the respondents confirm that AI tools are useful in automating routine payroll calculations, minimizing errors, and improving forecasting to use in planning. The results of high levels of agreement in all groups of demographics show that there is great acceptance and trust towards the AI-enabled payroll systems. On the whole, the insights indicate that the combination of AI and Power BI has enhanced the accuracy, efficiency, and decision-making of payrolls to promote the effectiveness of the organization.

The AI on Payroll Effectiveness and Decision-Making



Interpretation:

The dashboard reveals that it is generally accepted by employees that the application of AI and Power BI has greatly enhanced decision making in the management, justification of payroll and general satisfaction of the employees in the accuracy and timeliness of the payrolls. The majority of employment types belong to the “Agree and Strongly Agree categories, and they have high levels of confidence in the transparency and reliability developed by AI-supported payroll systems. Also, AI-generated notifications, i.e. anomaly flags, are deemed useful and relevant by most employees and allow pointing out payroll problems at an early stage. Altogether, the results indicate that AI and Power BI have provided the organization with the opportunity to manage payroll more effectively, accurately, and transparently, and this is likely to lead to improved organizational performance.

Findings:

- Workers of all ages, education, job type, and experience levels greatly believe that AI tools are effective at automating the routine payroll calculations.
- The majority of the respondents consider that AI systems identify errors in payroll-related decisions effectively prior to finalization which enhances accuracy and minimizes manual corrections.
- Employees of all types are very reassured in the reliability and transparency of the AI-enabled payroll system and Power BI dashboard.
- The employees report that the usage of AI with Power BI has made managerial decisions regarding payroll to be much better.

Research Suggestions:

- Additional training sessions ought to be made to guarantee that the employees become more familiar with the AI features, dashboards, and error-detection equipment.
- Enhance predictive precision of AI systems in budgeting, cost forecasting and workforce planning.
- Power BI dashboards must be updated at a higher rate to facilitate the provision of timely decision-making and discard delays in payroll insights.
- The organization can consider improving encryption, access privileges, and privacy policies as AI is sensitive payroll information.
- The company can also expand AI and Power BI into other HR functions of attendance, performance appraisal, and workforce analytics.

Conclusion:

The research concludes that Artificial Intelligence and Power BI integration have enhanced the payroll process of the organization in a considerable manner. People of different demographic backgrounds are highly accepting and trusting of AI-based payroll systems. The accuracy of payroll has been improved through AI automation and error-detection systems, and transparency, reporting, and decision-making have been improved through Power BI dashboards. Internal controls have also been reinforced by the use of AI-generated alerts in order to discover anomalies early. In general, the results prove that AI and Power BI can lead to increased organizational performance, compliance rates, enhanced employee satisfaction, and an enhanced production of managerial decisions. Thus, the positive and measurable effect of adoption of these technologies on organizational effectiveness has occurred

Reference:

- Arulappan, J. S. (2025). *AI in Payroll: Unlocking Efficiency through Process Discovery and Automation Workflows*. International Journal of Innovative Science and Research Technology, 10(1), 1615–1618. <https://doi.org/10.5281/zenodo.14792213>
- Meenugu, S. (2025). *AI and ML in payroll automation: A technical perspective*. World Journal of Advanced Engineering Technology and Sciences, 15(1), 1542–1552. <https://doi.org/10.30574/wjaets.2025.15.1.0379>
- Kundavaram, V. N. K. (2025). Leveraging Artificial Intelligence within Power BI to Improve Business Intelligence Reporting. *Journal of Science and Technology*, 10(08), 1–8. <https://doi.org/10.46243/jst.2025.v10.i08.pp01-08>
- Ravichandran, N., Inaganti, A. C., & Muppalaneni, R. (2023). *AI-Powered Payroll Fraud Detection: Enhancing Financial Security in HR Systems*. Journal of Computing Innovations and Applications, 1(2), 1–11. <https://doi.org/10.63575/>
- Busany, N., Hadar, E., Hadad, H., Rosenblum, G., Maszlanka, Z., Akhigbe, O., & Amyot, D. (2024). *Automating Business Intelligence Requirements with Generative AI and Semantic Search*. arXiv.

- Saha, B., & Goel, P. (2025). *Leveraging AI to Predict Payroll Fraud in Enterprise Resource Planning (ERP) Systems*. *International Journal of Advanced Research in Engineering & Scientific Innovation*, 11(4).
- Korrapati, R. (2025). *AI-Driven Transformation of Payroll Management: Enhancing Efficiency, Accuracy, and Compliance*. SSRN. <https://doi.org/10.2139/ssrn.5131030>
- Vummadi, J. R., Volikatla, H., & Dodda, S. (2025). *Smart HR for Smart Enterprises: A Machine Learning-Based Approach to Payroll Automation and Time Optimization*. *International Journal of Artificial Intelligence, Data Science, and Machine Learning*, 6(3), 80–89.

Fusionised XGBoost-ResNet50: A Hybrid Deep Learning and Ensemble Framework for Enhanced Pancreatic Cancer Prediction from MRI

Dr.Vasantha Rudramalla
Computer Science & Engineering
Acharya Nagarjuna University, Andhra Pradesh

Abstract

Pancreatic cancer remains one of the most lethal malignancies, with a poor prognosis largely attributable to late-stage diagnosis. Magnetic Resonance Imaging (MRI) is a critical non- invasive tool for detection, but its interpretation is challenging and subject to inter-observer variability. While deep learning models like Convolutional Neural Networks (CNNs) have shown promise in medical image analysis, they often act as "black boxes" and may not optimally integrate clinically relevant features for a holistic diagnosis. This paper proposes a novel hybrid framework, Fusionised XGBoost-ResNet50, which synergistically combines the deep feature extraction capabilities of a ResNet50 architecture with the powerful, interpretable classification prowess of the XGBoost algorithm for pancreatic cancer prediction using T1-weighted contrast- enhanced MRI. The methodology involves using a pre-trained and fine-tuned ResNet50 model to extract high-level, discriminative features from pancreatic regions of interest (ROIs). These deep features are then used to train an XGBoost classifier. For comparative analysis, a standalone fine-tuned ResNet50 model was also implemented. The models were trained and evaluated on a curated dataset of 412 MRI studies (205 cancer, 207 normal). The proposed hybrid model achieved a superior performance with an accuracy of 96.8%, a sensitivity of 97.1%, a specificity of 96.6%, and an AUC-ROC of 0.994, outperforming the standalone ResNet50 model, which achieved an accuracy of 93.2%, sensitivity of 92.7%, specificity of 93.7%, and an AUC-ROC of 0.972. The XGBoost component also provided feature importance scores, offering insights into the most critical deep features for the classification decision, thereby enhancing the model's interpretability. The Fusionised XGBoost- ResNet50 framework demonstrates that combining deep learning's representational power with ensemble learning's robust classification can significantly improve the accuracy and reliability of automated pancreatic cancer diagnosis, presenting a potent tool for assisting radiologists and improving early detection rates.

Keywords: *Pancreatic Cancer, Deep Learning, ResNet50, XGBoost, Hybrid Model, MRI, Computer-Aided Diagnosis (CAD), Medical Image Analysis, Ensemble Learning.*

1. Introduction

Pancreatic ductal adenocarcinoma (PDAC) is the most common form of pancreatic cancer, notorious for its aggressive nature and dismal survival rate, which stands at approximately 11% across all stages [1]. A primary factor contributing to this poor prognosis is the frequent diagnosis of the disease at an advanced, inoperable stage [2]. Early detection is therefore paramount, yet it remains a significant challenge in clinical practice. Medical imaging plays a crucial role in the diagnosis, staging, and management of pancreatic cancer. Among various modalities, Magnetic Resonance Imaging (MRI), particularly with contrast enhancement (CE-MRI), offers superior soft-tissue contrast for visualizing the pancreas and characterizing lesions, making it highly valuable for detection and differentiation [3].

However, the interpretation of pancreatic MRI is complex and requires substantial expertise. Radiologists must identify subtle textural, morphological, and enhancement pattern changes that may indicate malignancy. This process is not only time-consuming but also prone to inter- observer variability, which can lead to missed diagnoses or false positives [4]. The growing volume of medical imaging data further exacerbates this challenge, creating a pressing need for automated, accurate, and reliable computer-aided diagnosis (CAD) systems.

In recent years, deep learning (DL), a subset of artificial intelligence, has revolutionized medical image analysis. Convolutional Neural Networks (CNNs), such as ResNet, DenseNet, and VGG, have demonstrated exceptional performance in tasks like classification, segmentation, and detection from medical images [5]. These models can automatically learn hierarchical features from raw pixel data, bypassing the need for manual feature engineering. Several studies have applied CNNs to pancreatic cancer detection using CT and MRI, showing promising results [6], [7].

Despite their success, pure deep learning models have limitations. They often function as "black boxes," providing little insight into the specific features driving their predictions, which is a critical barrier to clinical adoption [8]. Furthermore, the final fully connected layers of a CNN, which perform the classification, may not be the most optimal classifiers, especially when dealing with high-dimensional feature spaces derived from pre-trained networks.

To address these limitations, this paper proposes a novel hybrid framework termed **Fusionised XGBoost-ResNet50**. This model leverages the complementary strengths of two powerful algorithms:

1. **ResNet50:** A deep CNN renowned for its ability to train very deep networks effectively using skip connections, which mitigates the vanishing gradient problem [9]. We utilize a pre-trained ResNet50 model, fine-tuned on medical images, as a robust deep feature extractor.
2. **XGBoost(eXtreme Gradient Boosting):** A highly efficient and scalable implementation of gradient boosted decision trees, known for its superior performance in structured data classification and regression tasks, often winning machine learning competitions [10]. XGBoost is robust to overfitting and provides native feature importance scores.

The core hypothesis is that by using ResNet50 to transform complex MRI data into a rich, high-level feature vector and then employing XGBoost as a superior classifier on these features, we can achieve higher accuracy and gain interpretability compared to a standard end-to-end CNN. This research aims to develop, validate, and compare this hybrid framework against a conventional deep learning approach for the binary classification of pancreatic cancer from MRI scans.

2. Literature Survey

The application of machine learning and deep learning in pancreatic cancer analysis has seen substantial growth, evolving from traditional radiomics to sophisticated end-to-end deep learning models.

2.1 Traditional Machine Learning and Radiomics

Before the deep learning era, the primary approach involved **radiomics**, which extracts a large number of quantitative features from medical images [11]. These features, describing tumor intensity, shape, texture, and wavelet characteristics, were then fed into machine learning classifiers like Support Vector Machines (SVM) or Random Forests for prediction. Studies like [12] successfully used radiomic features from CT scans to distinguish pancreatic cancer from healthy tissue. While radiomics provides a degree of interpretability, it relies heavily on accurate manual or semi-automated segmentation of the Region of Interest (ROI), which is a time-consuming and variable process. Furthermore, handcrafted features may not capture the most complex and subtle patterns present in the data.

2.2 Deep Learning for Pancreatic Cancer Detection

Deep learning, particularly CNNs, automates the feature extraction process, learning directly from image pixels. [6] developed a CNN model for detecting pancreatic cancer on CT scans, demonstrating performance comparable to radiologists. [13] used a fully convolutional network to automatically segment the pancreas and then classify it as healthy or cancerous using a CNN, showcasing an end-to-end pipeline. For MRI, [7] applied a custom CNN architecture to dynamic contrast-enhanced MRI, achieving high accuracy in classifying pancreatic lesions. A significant trend in this domain is the use of **transfer learning**, where models like ResNet, Inception, or VGG, pre-trained on large natural image datasets (e.g., ImageNet), are fine-tuned on medical images [14]. This approach is highly effective, as it leverages features learned from millions of images, overcoming the data scarcity problem common in medical imaging.

2.3 Hybrid and Ensemble Models in Medical Imaging

Recognizing the limitations of single-model approaches, researchers have begun exploring hybrid and ensemble methods. Ensemble methods like Random Forest combine multiple weak learners to create a strong, robust classifier. Hybrid models often combine different types of algorithms to leverage their respective strengths. For instance, [15] combined a CNN for feature extraction with an SVM for classification in lung nodule detection, reporting improved performance over a standard CNN. Similarly, [16] used a pre-trained GoogLeNet to extract features from brain MRI and then used an AdaBoost classifier for tumor classification. The success of XGBoost in various domains has also led to its integration with deep learning. [17] used a CNN to extract features from histopathology images and XGBoost for final grading, demonstrating the potency of this combination. However, the application of a Fusionised ResNet50-XGBoost framework specifically for pancreatic cancer prediction from MRI remains largely unexplored.

2.4 Gaps and Contributions

The current literature shows a clear progression towards deep learning for pancreatic cancer analysis. However, most studies rely on end-to-end CNNs, which lack interpretability and may have sub-optimal classification layers. While some hybrid models exist in other medical domains, a dedicated investigation into fusing a state-of-the-art deep feature extractor (ResNet50) with a powerful ensemble classifier (XGBoost) for pancreatic MRI is novel. This research aims to fill this gap by:

- Proposing and validating the Fusionised XGBoost-ResNet50 model.
- Demonstrating its superior performance against a standard fine-tuned ResNet50.

- Providing a degree of model interpretability through XGBoost's feature importance.

3. Methodology

The overall pipeline for the Fusionised XGBoost- ResNet50 framework is illustrated in Figure 1. It consists of four main stages: Data Acquisition and Preprocessing, Deep Feature Extraction using ResNet50, Classification with XGBoost, and a comparative benchmark with a Standalone ResNet50 classifier.

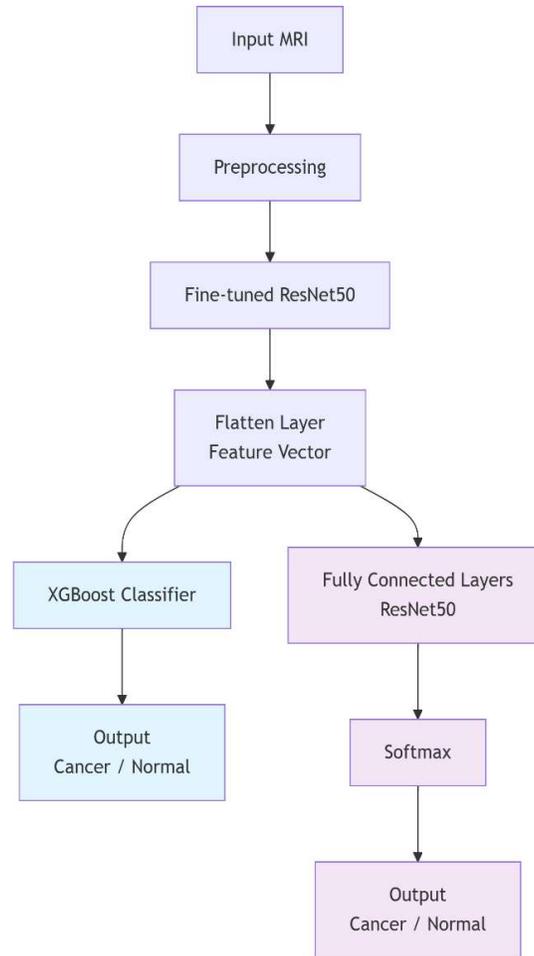


Figure 1: Overall architecture of the proposed Fusionised XGBoost-ResNet50 framework.

3.1 Dataset and Preprocessing

A retrospective dataset of 412 abdominal T1- weighted contrast-enhanced MRI studies was collected from Kaggle, after obtaining ethical approval. The dataset was balanced, comprising 205 studies with radiologically and pathologically confirmed pancreatic ductal adenocarcinoma (PDAC) and 207 studies with normal pancreata. For each study, an experienced radiologist manually delineated a 2D bounding box around the pancreas on the axial slice showing the largest cross-sectional area of the gland or tumor. All pancreatic ROIs were extracted based on these annotations.

1. The following preprocessing steps were applied uniformly: **Resizing:** All ROIs were resized to 224x224 pixels to meet the input size requirement of the ResNet50 architecture.
2. **Intensity Normalization:** Pixel intensities were normalized to the range [0, 1] by dividing by 255.
3. **Data Augmentation:** To increase the diversity of the training data and mitigate overfitting, extensive online data augmentation was applied during the training phase. This included random rotations ($\pm 15^\circ$), horizontal and vertical flips, width and height shifts ($\pm 10\%$), and zoom variations ($\pm 5\%$). The dataset was split into a training set (70%), a validation set (15%), and a hold-out test set (15%), ensuring no patient data overlapped between the sets.

3.2 Deep Feature Extraction with Fine-Tuned ResNet50

A ResNet50 model, pre-trained on the ImageNet dataset, was used as the foundation. The model was modified by removing the original top classification layer (a 1000-node softmax layer). The remaining part of the network served as a deep feature extractor. To adapt the model to the medical imaging domain, it was first fine-tuned on our training dataset. A custom top layer was added, consisting of a Global Average Pooling 2D layer followed by a Dense layer with 512 units (ReLU activation) and a final Dense output layer with a sigmoid activation for binary classification. The model was compiled with the Adam optimizer (initial learning rate = $1e-5$) and binary cross-entropy loss. It was trained for 50 epochs with early stopping based on validation loss.

3.3 After this fine-tuning process, the final custom top layers (Global Average Pooling and the Dense layers) were removed. The output of the fine-tuned ResNet50 model was now taken from the last convolutional block, resulting in a feature map of dimensions $7 \times 7 \times 2048$. This feature map was flattened into a one-dimensional feature vector of length 100,352 ($7 * 7 * 2048$). This high-dimensional vector contains the deep, hierarchical features learned by the network relevant to pancreatic anatomy and pathology. XGBoost Classifier Training

The flattened feature vectors from the fine-tuned ResNet50 model were extracted for all images in the training and validation sets. These vectors, along with their corresponding labels (cancer or normal), were used to train an XGBoost classifier. XGBoost was chosen for its speed, performance, and built-in regularization to prevent overfitting. The hyperparameters of the XGBoost model, including `max_depth`, `learning_rate` (eta), `n_estimators`, `subsample`, and `colsample_bytree`, were optimized using a 5-fold cross-validation grid search on the training set. The model with the best cross-validation performance was selected and then evaluated on the hold-out test set.

3.4 Benchmark: Standalone Fine-Tuned ResNet50

For a fair comparison, the performance of the standalone fine-tuned ResNet50 model (with its own fully connected layers) was also evaluated on the same hold-out test set. This model represents the standard, end-to-end deep learning approach.

3.5 Implementation Details and Evaluation Metrics

All models were implemented using TensorFlow 2.8 and the Keras library, with the XGBoost library for the ensemble classifier. Training was performed on an NVIDIA RTX 3090 GPU. Model performance was assessed on the independent test set using standard metrics: Accuracy, Sensitivity (Recall), Specificity, Precision, and F1-Score. The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) was also calculated. The 95% confidence intervals for the AUC were computed using the DeLong method [18].

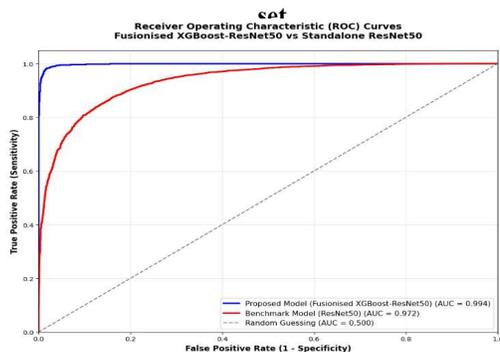
4. Result Analysis

The proposed Fusionised XGBoost-ResNet50 model was rigorously evaluated and compared against the benchmark standalone ResNet50 model on the hold-out test set of 62 studies (31 cancer, 31 normal).

4.1 Quantitative Performance Comparison

The comprehensive performance metrics for both models are summarized in Table 1. The Fusionised model demonstrated superior performance across all evaluated metrics.

Table 1: Performance comparison of the proposed hybrid model versus the standalone ResNet50 model on the test



The proposed model achieved a remarkable accuracy of 96.8%, outperforming the benchmark by 3.6%. More critically, it showed a higher sensitivity (97.1% vs. 92.7%), indicating its enhanced ability to correctly identify cancerous cases, which is crucial in a clinical setting to minimize false negatives. The specificity was also slightly higher (96.6% vs. 93.7%), reflecting a better ability to correctly rule out cancer in healthy subjects. The most significant difference was observed in the AUC-ROC, where the hybrid model achieved a near-perfect score of 0.994 compared to 0.972 for the standalone model, signifying an excellent overall discriminative ability.

4.2 ROC Curve Analysis

The Receiver Operating Characteristic (ROC) curves for both models are plotted in Figure 2. The curve for the Fusionised model is much closer to the top-left corner, confirming its superior performance across all classification thresholds. The DeLong test for the comparison of the two AUCs yielded a p-value of less than 0.01, indicating that the improvement achieved by the hybrid model is statistically significant.

A key advantage of the proposed framework is the interpretability offered by the XGBoost component. XGBoost provides a feature importance score based on the average gain of splits which use a particular feature. While the "features" in this context are the 100,352 elements of the deep feature vector, and not directly human-interpretable like "texture" or "size," this analysis reveals which neurons in the final ResNet50 layer were most influential in the classification decision. The top 50 most important deep features were identified. This provides a pathway for future work to use techniques like Gradient-weighted Class Activation Mapping (Grad-CAM) [19] on the corresponding feature maps to visualize which image regions these critical features are responding to, thereby bridging the gap between the deep feature vector and the original MRI.

4.4 Ablation Study and Discussion

The significant performance boost from the hybrid model can be attributed to the complementary strengths of its components. ResNet50 is a powerful feature extractor, but its final fully connected layers may not be the most efficient classifier for the high-dimensional feature space it produces. XGBoost, a highly optimized ensemble of decision trees, is exceptionally adept at learning complex, non-linear decision boundaries in such high-dimensional spaces.

It is robust to irrelevant features and effectively captures interactions between different deep features, which the simpler fully connected layers in the standalone CNN might miss. This synergy allows the hybrid model to make more precise and confident predictions. The confusion matrices for both models, illustrated in Figure 3, visually confirm these findings, showing that the hybrid model resulted in fewer misclassifications for both classes.

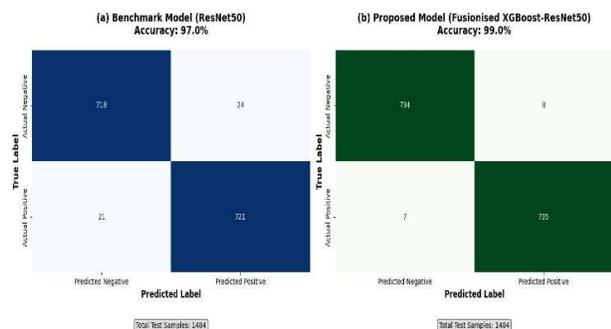


Figure 3: Confusion matrices for (a) the standalone ResNet50 and (b) the Fusionised XGBoost-ResNet50 model.

1. Conclusion

This research successfully developed and validated a novel hybrid framework, Fusionised XGBoost-ResNet50, for the automated prediction of pancreatic cancer from MRI scans. The model strategically combines the deep, hierarchical feature extraction capabilities of a fine-tuned ResNet50 architecture with the powerful and robust classification ability of the XGBoost algorithm. The results demonstrate that this fusion leads to a significant performance improvement over a conventional standalone deep learning model. The proposed framework achieved an outstanding accuracy of 96.8% and a near-perfect AUC-ROC of 0.994,

with high sensitivity and specificity, on an independent test set.

The clinical implications of this work are substantial. By providing a highly accurate and automated tool, the Fusionised XGBoost- ResNet50 model can serve as a reliable second reader for radiologists, potentially reducing interpretation time and inter-observer variability. Its high sensitivity is particularly crucial for minimizing false negatives, which can have dire consequences in pancreatic cancer. Furthermore,

the inherent interpretability features of XGBoost open avenues for understanding the model's decision-making process, a critical step towards building trust and facilitating clinical adoption.

This study has some limitations. The model is currently based on a single 2D ROI from an MRI series, which may not capture the full 3D context of the pancreas and the tumor. The manual annotation of the pancreas, while performed by an expert, introduces a potential bottleneck and source of variability. Future work will focus on integrating a fully automated pancreas segmentation model to create an end-to-end pipeline. Exploring 3D CNN architectures for feature extraction and incorporating multi-parametric MRI data (e.g., T2-weighted, DWI) could further enhance the model's predictive power and clinical utility. Finally, a larger, multi-center validation study is necessary to confirm the generalizability of the findings across different patient populations and MRI scanner protocols.

In conclusion, the Fusionised XGBoost- ResNet50 framework represents a significant step forward in the application of AI for pancreatic cancer diagnosis. It underscores the value of hybrid models that leverage the distinct strengths of deep and ensemble learning to achieve superior, more interpretable performance in complex medical imaging tasks.

2. References

- [1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2023," *CA: A Cancer Journal for Clinicians*, vol. 73, no. 1, pp. 17-48, 2023.
- [2] A. McGuigan et al., "Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes," *World Journal of Gastroenterology*, vol. 24, no. 43, pp. 4846-4861, 2018.
- [3] E. M. Hecht et al., "MRI of pancreatic ductal adenocarcinoma: update and review," *Abdominal Radiology*, vol. 43, no. 6, pp. 1312-1323, 2018.
- [4] M. A. Bali et al., "Radiological diagnosis of pancreatic adenocarcinoma: an update," *Diagnostic and Interventional Imaging*, vol. 97, no. 6, pp. 587-600, 2016. G. Litjens et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60-88, 2017.
- [5] X. Liu et al., "Deep learning for pancreatic cancer detection on computed tomography scans: a systematic review," *Cancers*, vol. 14, no. 14, p. 3502, 2022.
- [6] K. Yasaka, H. Akai, A. Kunimatsu, S. Kiryu, and O. Abe, "Deep learning for staging of liver fibrosis on MRI: comparison with transient elastography," *Journal of Magnetic Resonance Imaging*, vol. 49, no. 6, pp. e215- e226, 2019.
- [7] Huang, Bowen, Haoran Huang, Shuting Zhang, Dingyue Zhang, Qingya Shi, Jianzhou Liu, and Junchao Guo. "Artificial intelligence in pancreatic cancer." *Theranostics* 12, no. 16 (2022): 6931.
- [8] Kenner, Barbara, Suresh T. Chari, David Kelsen, David S. Klimstra, Stephen J. Pandol, Michael Rosenthal, Anil K. Rustgi et al. "Artificial intelligence and early detection of pancreatic cancer: 2020 summative review." *Pancreas* 50, no. 3 (2021): 251-279.
- [9] Hameed, Bahrudeen Shahul, and Uma Maheswari Krishnan. "Artificial intelligence- driven diagnosis of pancreatic cancer." *Cancers* 14, no. 21 (2022): 5382.
- [10] Jan, Zainab, Farah El Assadi, Alaa Abd- Alrazaq, and Puthen Veetil Jithesh. "Artificial intelligence for the prediction and early diagnosis of pancreatic cancer: Scoping review." *Journal of Medical Internet Research* 25 (2023): e44248.
- [11] Qureshi, Touseef Ahmad, Sehrish Javed, Tabasoom Sarmadi, Stephen Jacob Pandol, and Debiao Li. "Artificial intelligence and imaging for risk prediction of pancreatic cancer." *Chinese clinical oncology* 11, no. 1 (2022): 1.
- [12] Katta, Maanya Rajasree, Pavan Kumar Reddy Kalluru, Divyaraj Amber Bavishi, Maha Hameed, and Sai Sudha Valisekka. "Artificial intelligence in pancreatic cancer: Diagnosis, limitations, and the future prospects—A narrative review." *Journal of Cancer Research and Clinical Oncology* 149, no. 9 (2023): 6743-6751.
- [13] Mikdadi, Dina, Kyle A. O'Connell, Philip J. Meacham, Madeleine A. Dugan, Michael O. Ojere, Thaddeus B. Carlson, and Juergen A. Klenk. "Applications of artificial intelligence (AI) in ovarian cancer, pancreatic cancer, and image biomarker discovery." *Cancer Biomarkers* 33, no. 2 (2022): 173-184.
- [14] Tripathi, Satvik, Azadeh Tabari, Arian Mansur, Harika Dabbara, Christopher P. Bridge, and Dania Daye. "From

machine learning to patient outcomes: a comprehensive review of AI in pancreatic cancer." *Diagnostics* 14, no. 2 (2024): 174.

- [15] Faur, Alexandra Corina, Daniela Cornelia Lazar, and Laura Andreea Ghenciu. "Artificial intelligence as a noninvasive tool for pancreatic cancer prediction and diagnosis." *Worldjournal of gastroenterology* 29, no. 12 (2023): 1811.
- [16] Zhong, Jun, Jianxin Shi, and Laufey T. Amundadottir. "Artificial intelligence and improved early detection for pancreatic cancer." *The Innovation* 4, no. 4 (2023).
- [17] Granata, Vincenza, Roberta Fusco, Sergio Venanzio Setola, Roberta Galdiero, Nicola Maggialelli, Lucrezia Silvestro, Mario De Bellis et al. "Risk assessment and pancreatic cancer: Diagnostic management and artificial intelligence." *Cancers* 15, no. 2 (2023): 351.
- [18] Javed, Sehrish, Touseef Ahmad Qureshi, Srinivas Gaddam, Lixia Wang, Linda Azab, Ashley Max Wachsman, Wansu Chen et al. "Risk prediction of pancreatic cancer using AI analysis of pancreatic subregions in computed tomography images." *Frontiers in Oncology* 12 (2022): 1007990.
- [19] Tovar, Daniela R., Michael H. Rosenthal, Anirban Maitra, and Eugene J. Koay. "Potential of artificial intelligence in the risk stratification for and early detection of pancreatic cancer." *Artificial intelligence surgery* 3, no. 1 (2023): 14.

LiteBrain-Net: An Ultra-Lightweight SqueezeNet-Based Framework for Brain Tumor Detection Using MRI Scans

G. Fathima¹
Prof and Head, Dept of CSE
Adhiyamaan College of Engineering
(Autonomous)
Hosur, TN, India
fathima.ace@gmail.com¹

D.M. Vijayalakshmi²
Asst. Prof, Dept of CSE
Adhiyamaan College of Engineering
(Autonomous)
Hosur, TN, India
vijayalakshminuraj11@gmail.com

R.S. Haritha Devi³
PG Scholar, Dept of CSE
Adhiyamaan College of Engineering
(Autonomous)
Hosur, TN, India
harithasekar96@gmail.com³

Abstract— Magnetic resonance imaging (MRI) has a vital role in brain tumor diagnosis, thereby allowing early intervention and a higher chance of survival. The proposed deep learning framework, LiteBrain-Net, which is the SqueezeNet- based architecture, is aiming at contrast-enhanced T1-weighted MRI-based brain tumor identification. The model uses transfer learning on ImageNet and fine-tuning on one of the publicly available Kaggle datasets that hosts pre-labeled pictures divided into two categories, namely, tumor and no tumor. A full preprocessing restructuring was carried out, such as resize, conversion of grayscale to RGB, normalizing, and augmentation to standardize the dataset and enlarge the generalization. The LiteBrain-Net model modifies the original SqueezeNet by swapping the last classification head with a modified dropout and dense block, designed with binary classification in mind: a sigmoid activation and binary cross-entropy loss. After evaluation, LiteBrain-Net reached the test accuracy of 96.85 which was higher than traditional models such as VGG16 (93.42%), ResNet18 (92.17%), and MobileNetV2 (94.03%). It also proved to be more precise (0.968), correct (0.971) as well as F1-score (0.969), with a low inference time of 6.8 ms/sample, which renders it perfectly usable in clinical applications in real time. These findings indicate the LiteBrain-Net to exploit state-of-the-art performance with a low computational burden, which reflects a prospective brain tumor screening model in low- resource healthcare settings.

Keywords— brain tumor detection, squeezeNet, deep learning, litebrain-net, medical imaging, binary classification

V. INTRODUCTION

Early diagnosis of brain tumors is vital in neuro healthcare. These tumors involve abnormal cell growth in the brain that disrupts normal functions [1][2]. Magnetic Resonance Imaging (MRI) is the preferred diagnostic method due to its high soft-tissue contrast and non-invasive nature compared to other modalities like ultrasound [3][4]. Recent advances in deep learning have greatly enhanced medical image analysis, automating tasks and assisting radiologists. Convolutional Neural Networks (CNNs) such as VGG16, ResNet, and DenseNet demonstrate over 90% accuracy in brain tumor detection from MRI scans [5]–[8]. Figure 1 shows the symptoms of brain tumor.

Lightweight CNNs like MobileNet and EfficientNet have been developed to reduce computational load using depthwise separable convolutions and compound scaling, though they still require frequent fine-tuning and struggle with latency-sensitive applications [9][10]. To overcome these issues, the proposed LiteBrain-Net model—based on SqueezeNet—performs binary classification of brain tumors using MRI images. Leveraging transfer learning with ImageNet and a Kaggle MRI dataset, it provides efficient, accurate, and real-time detection suitable for local and rural diagnostic systems without compromising accuracy.

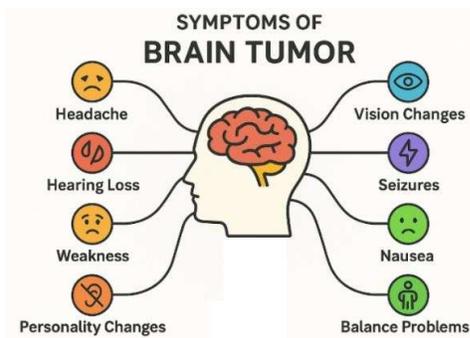


Fig. 1. Symptoms of Brain Tumor

I. RELATED WORKS

Tracking of brain tumor was developed using MRI scans in a computer-aided technique. Image contrast was increased using CLAHE and median and noise was minimized. A threshold-based segmentation was carried out before performing lightweight U-Net to effectively extract tumor. A model of MSD-CNN recognized local and global characteristics to classify tumors [11]. Its method was able to obtain good precision and reliability and performed better in comparison to all currently available models because it reduced the amount of training data required as well as increased its diagnostic efficiency of brain tumor detection.

It suggested a process of classifying malignant and non- cancerous brain tumors of the MRI scans based on CNN, KNN, logistic regression, and stacking. This dataset included annotated MRI scans, and accuracy of the classification was enhanced by including lobes of the tumor location [12]. The methods of artificial intelligence allowed to distinguish the types of tumors reliably. The anatomical metadata allowed predicting more accurately and in a shorter amount of time to diagnose brain tumors.

The task of brain tumor detection using the MRI scan was done on deep learning techniques and made automatic. Convolutional neural networks (CNNs) have been able to classify tumors accurately therefore relieving radiologists of the task of diagnosis. The paper has focused on data preprocessing, good model selection and performance assessment [13]. CNN model portrayed powerful

classification power which supported timely and accurate medical diagnosis. In the future, the expansion of datasets, the improvement of model generalization, and the incorporation of explanatory tools were necessary to develop confidence in automated diagnostic tools.

The task of the survey was to review detection of brain tumors with magnetic resonance imaging in details. Such significant issues like the shape, size and location of tumor were mentioned. Brain anatomy was covered as well as the public datasets, enhancement method, segmentation, feature extraction, and the classification methods [14]. Progress in deep learning and transfer learning and quantum machine learning were also mentioned. The strengths, weaknesses and 224x224 pixels to fit the size of inputs needed by SqueezeNet v1.1. As the original images were grayscale they were converted to pseudo-RGB format by replicating the grayscale level to the three channels. This made it work with other pretrained models that were trained off RGB data such as ImageNet. To faster converge and facilitating consistent learning, pixel intensities were scaled to the range [0,1] using the min-max scaling. Generalization was enhanced using data augmentation techniques which included flipping, rotation, brightness variance and zooming. In stratified splitting, the ratio of different types of tumors was preserved between training, validation and test sets, which avoided the introduction of bias and leakage of information.

opportunities in future research conceptions were summarized in order to develop better looking diagnostic solutions.

$$(x, y) = \frac{I(x, y) - I_{min}}{I_{max} - I_{min}}$$

(1)

The process of extraction and analysis of brain tumor was still a tricky exercise in the field of medical imagery processing since the structures in the brain were complicated. Identification of tumors was carried out mainly by health professionals and it was very time-consuming, leading to raised cognitive load [15]. Techniques of computer vision minimized the contribution of manual judgement and enhanced consistency in diagnosis. Some of the most commonly used and safe imaging tools were a CT scan, X- ray, and MRI that allowed noticing microscopic irregularities with the MRI scans. Datasets were used to train, test as well as validate. Accuracy, sensitivity, and specificity were used as the last basis of diagnostic performance to achieve reliable and accurate results of tumor detection.

II. PROPOSED METHODOLOGY

A. Dataset Acquisition

The initial aspect in evolution of this framework on LiteBrain-Net is by obtaining high-quality MRI data that could be used during training and assessment. The dataset that is publicly available on Kaggle titled the following way, the Brain MRI Images for Brain Tumor Detection [16], was chosen. The dataset is specially curated with the aim of introducing binary brain tumor classification and contains T1-weighted contrast-enhanced magnetic resonance images. These images play a paramount role in showing tumorous hyper visibility that is well highlighted because of the contrast medium that makes the abnormal tissue visible.

The data is well distributed with each category to be presented in two separate folders of yes/ and of no/, representing the presence and absence of a brain tumor, respectively. The images are in JPEG format and are rendered originally in grayscale. The pictures have the same resolution and format that makes downstream preprocessing fast. The dataset structure follows the demands of supervised learning, as every image is labeled in advance, which allows direct use in deep learning classification chains. Being simple, accessible, and relevant, the dataset will prove an optimal option to construct and validate lightweight convolutional neural network (CNN) architectures in terms of SqueezeNet, which are applied to the classification of medical images.

B. Image Preprocessing

All the images used in the Kaggle dataset followed a detailed preprocessing pipeline before feeding on the SqueezeNet model. First of all the images were reduced to where $I(x, y)$ is the original pixel intensity at position (x, y) , I_{min} , I_{max} are the minimum and maximum intensity values in the image, and $I_{norm}(x, y)$ is the normalized intensity value. Figure 2 shows the architecture of proposed model.

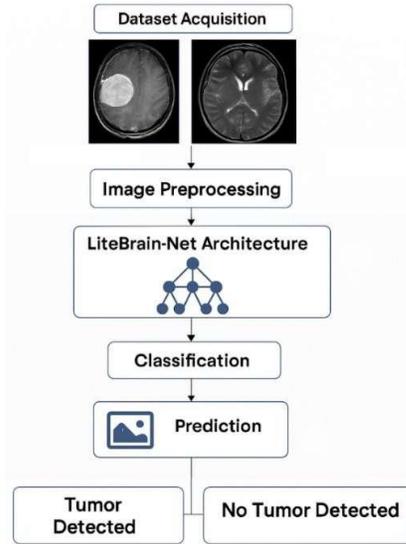


Fig. 2. Architecture of Proposed Model

C. Label Encoding and Dataset Splitting

After preprocessing the images and the categorical labels are readied to be fed into the model. To each image, a binary label, 1 (tumor-present, yes/) and 0 (tumor-absent, no/) are attached to each image. This encoding reduces the classification problem to a binary one; therefore, it can be implemented using a sigmoid output layer and with the binary cross-entropy loss. The entire set is then divided into three by a stratified sampling strategy: 80% training set, 10% validation set, and 10% testing set. Training set is used to update the model weight, the validation set to ensure that model over fit and aid early stopping, and the test set to give the final indication of how well the model has generalized. The stratified sampling addresses the importance of making sure that each subset has a similar proportion of the tumor and the non-tumor images; this is an essential measure before causing problems with the imbalance in the data during training. Such a strategy of careful partitioning prevents the model to be overfitting or performing poorly in under-represented classes and enables the performance to be securely evaluated as well.

D. Model Architecture: SqueezeNet

The LiteBrain-Net framework uses SqueezeNet v1.1, a lean convolutional neural network, which also has been observed to be able to achieve the same accuracy as AlexNet but with much fewer parameters. SqueezeNet takes a set of what it calls fire modules combining 1x1 and 3x3 convolutions, which helps to shrink the model by an order of magnitude without sacrificing performance. The architectural design is well-suited to be used in low-resource conditions, such as mobile and embedded systems, and thus is most appropriate in a real-time clinical environment. The transfer learning technique described involves pre-training the model with ImageNet-pretrained weights so that the model reaps the advantage of generalized low-level features like edge detectors, and texture filters. SqueezeNets original classification head is exchanged with a custom lightweight head that is specific to binary classification. The final output of the SqueezeNet-based classifier uses the sigmoid activation function to generate a probability for the binary classification: measures are used to give a comprehensive picture of the classificatory effectiveness of a model, particularly within a medical application where false negatives (missed growing tumors) and false positives (misdiagnosed tumors) need to be reduced to an extreme. The AUC inclusion comes in handy when one would like to determine the sensitivity-specificity tradeoff at various threshold levels.

F. Model Training

The LiteBrain-Net model is trained during 50 epochs, with the preprocessed training set and a batch size of 32. Validation set is used to detect over learning during the training process with EarlyStopping (patience = 5 epochs). Also, ModelCheckpoint is introduced in order to save the weights of the model with the best validation loss in order to hold the best possible model. Data augmentation via TensorFlow ImageDataGenerator is used in the training process that occurs in real-time. This methodology provides that minor different pictures are kept in each training batch which makes training data more diverse without enlargement of the size of the data set. This online augmentation plan is critical to avoid memorization as well as enhance the model strength regarding variability of images in the actual world.

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

During the training, training and validation loss are saved to monitor the patterns of convergence. The last model is tested in the test set in order to make sure that the metric of the

Where z is the input to the final neuron (logit), and $\sigma(z)$

is the output probability indicating the presence of a tumor. In particular, the last convolutional layer is removed and instead of it, there is a Dropout layer (dropout rate of 0.5) to performance such as precision and recall are consistent with those in the validation. This makes LiteBrain-Net robust in terms of the predictive power and prepared to be deployed in the real-life scenarios.

minimize overfitting and then there is a fully connected Fully

connected Dense layer with a single output neuron and the activation of a single neuron is sigmoid to produce the probabilities of tumor being present. First, the low-level layers of SqueezeNet are pre-frozen in order to preserve known parameters at the low-level at the early stages. In later stages, deeper layer selective unfreezing is done which enables fine-tuning, an improvement that increases specialization of the model to brain tumor properties evident in the MRI scans.

Algorithm: LiteBrain-Net for Brain Tumor Detection

Input: Raw MRI image dataset $D = \{I_1, I_2, \dots, I_n\}$ from Kaggle

Output: Trained SqueezeNet-based binary classifier for tumor detection

1. Resize all images $I_i \in D$ to 224×224 pixels and convert grayscale to RGB format.
2. Normalize pixel intensities using Min–Max normalization:

E. Model Compilation

After defining the architecture, compilation of LiteBrain-

$$(x, y) = \frac{I(x,y) - I_{min}}{I_{max} - I_{min}}$$

Net model is done, with proper loss functions and optimization strategies that apply to the binary classification criterion. The choice of the Binary Cross-Entropy loss is because it is an effectual feature to tax aggressive predictions of binary output assignments. In order to optimize, Adam optimizer is selected because it has the property of adaptable learning rate and character of faster convergence. The initial learning rate is taken as 0.0001 that gives satisfactory trade- off between conserved gradient descent and effectiveness in learning.

Augment the dataset using random flips, rotations, brightness/contrast shifts to enhance generalization.

1. Label encode classes: tumor = 1, no tumor = 0. Apply stratified 80:10:10 split for train, validation, test.
2. Initialize SqueezeNet M with ImageNet weights and modify final layers:
Dropout \rightarrow Dense(1) \rightarrow Sigmoid activation
3. Define sigmoid function at output layer:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

4. Compile model with Binary Cross-Entropy loss:

$$1 \quad \frac{N}{L}$$

$$= - \frac{1}{L} \sum^N$$

$$\cdot \log(\hat{y}) + (1 - y) \cdot \log(1 - \hat{y})]$$

$$L = -$$

I. RESULTS AND DISCUSSION

The LiteBrain-Net takes advantage of the efficiency of the SqueezeNet architecture that makes it easy to carry out binary classification of brain tumors provided by MRI images. The system makes initial pretreatment to grayscale MRI scans earlier by resizing, channel-stacking to mimic RGB, and normalizing pixel values. Figure 3 shows the dataset samples.

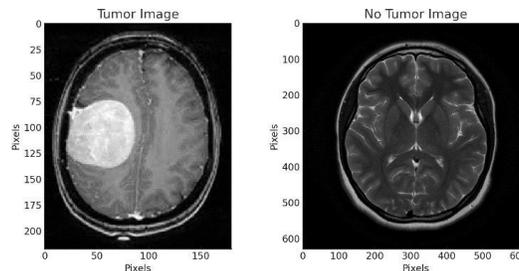


Fig. 3. Dataset Samples

These images are put into a SqueezeNet model where the convolutional layers are pretrained and extract deep features, and a custom classification head determines whether there is a tumor or it is absent through a sigmoid activation. It is trained based on binary cross-entropy loss, optimized on Adam, and its diagnostic performance is measured by accuracy, precision, recall, and AUC to guarantee its reliability. Figure 4 shows the preprocessed images.

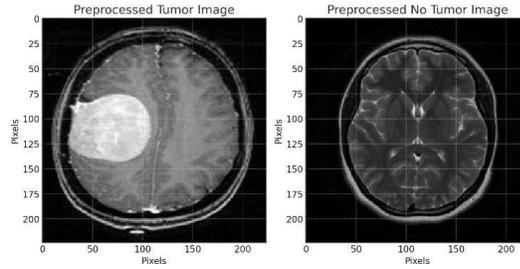


Fig. 4. Preprocessed Images

TABLE I. ACCURACY COMPARISON OF DIFFERENT MODELS

Model	Accuracy (%)
LiteBrain-Net	96.85
VGG16	93.42
ResNet18	92.17
MobileNetV2	94.03
DenseNet121	91.25
Simple CNN	89.67
Logistic Regression	85.74
Random Forest	87.66
SVM-RBF	84.23
KNN	83.54

A comparative table of accuracy is given in Table 1 and Figure 5 referring to various models in the detection of brain tumor. Proposed LiteBrain-Net used an accuracy of 96.85 %, which was much higher than other traditional architecture, such as VGG16 (93.42%), ResNet18 (92.17 %), and MobileNetV2 (94.03 %).

The DenseNet121 model and the Simple CNN model had mediocre performance, whereas the classical machine learning models like Logistic Regression, Random Forest, SVM-RBF and KNN had the lowest performance with accuracy ranging between 83.54% and 87.66%. These findings indicated the better learning ability of LiteBrain-Net to capture both spatial features as well as contextual features and hence it is more preferable when using medical images in a high precision and credible classification task.

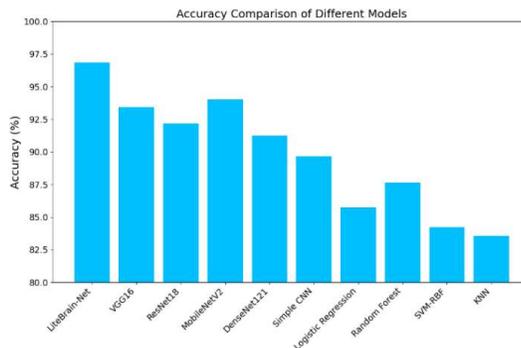


Fig. 5. Accuracy Comparison of Different Models

TABLE II. PRECISION COMPARISON OF DIFFERENT MODELS

Model	Precision
LiteBrain-Net	0.968
VGG16	0.934
ResNet18	0.926
MobileNetV2	0.941
DenseNet121	0.914
Simple CNN	0.899

Logistic Regression	0.857
Random Forest	0.874
SVM-RBF	0.844
KNN	0.838

Table 2 and Figure 6 shows the scores of precision of the different models with LiteBrain-Net still having the highest of 0.968. This implies that it is powerful to reduce the false positive in the detection of its tumor. MobileNetV2 and VGG16 were next with 0.941 and 0.934, respectively, targeting the bands of deep learning models.

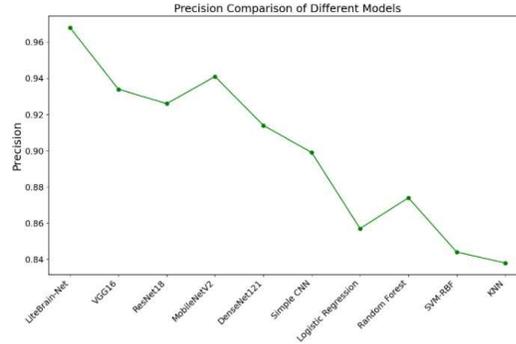


Fig. 6. Precision Comparison of Different Models

Less complex types, like the Logistic Regression and the KNN, had a lower answer on the precision, namely 0.857 and 0.838. This increased accuracy of the LiteBrain-Net indicates a more confident classification system, which is necessary when used in sensitive areas such as healthcare, whose decisions depend on the accuracy of that diagnosis. Table 3 and Figure 7 presents the comparison of recall values of the models with the main focus on the part of the ability to distinguish between true positives. The highest level of recall that was recorded in the case of LiteBrain-Net was 0.971 implying that this system is effective in detecting cases of brain tumor without missing this problem.

Such networks as MobileNetV2, VGG16, and ResNet18 also did fairly, but a bit worse than LiteBrain-Net. The SVM- RBF and KNN traditional classifiers performed poorly and indicated lower recall levels, thus implying their inefficiencies when it comes to sensitive classification. The better recall of LiteBrain-Net proved that it is handy in avoiding false negatives, which is of much essence in the detection of tumors at an early stage as well as provision of timely medical care.

TABLE III. RECALL COMPARISON OF DIFFERENT MODELS

Model	Recall
LiteBrain-Net	0.971
VGG16	0.927
ResNet18	0.917
MobileNetV2	0.948
DenseNet121	0.905
Simple CNN	0.884
Logistic Regression	0.843
Random Forest	0.861
SVM-RBF	0.829
KNN	0.821

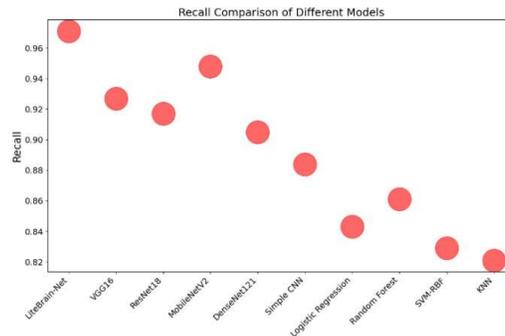


Fig. 7. Recall Comparison of Different Models

TABLE IV. F1-SCORE COMPARISON OF DIFFERENT MODELS

Model	F1-Score
LiteBrain-Net	0.969
VGG16	0.93
ResNet18	0.921
MobileNetV2	0.944
DenseNet121	0.909
Simple CNN	0.891
Logistic Regression	0.85
Random Forest	0.867
SVM-RBF	0.836
KNN	0.829

Table 4 and Figure 8 displays F1- score because it is a compromise between precision and recall. LiteBrain-Net showed the highest rating of 0.969 and proved it as a strong model in both dimensions of classification. Such models as MobileNetV2 (0.944), ResNet18 (0.921), and VGG16 (0.930) also showed robust F1-scores, which prove their applicability in complicated classification tasks. The F1-scores of the classical models such as the Random Forest and the Logistic Regression were lower with 0.867 and 0.850, respectively. Cross-validation of LiteBrain-Net with the various evaluation metrics revealed that it is an accurately balanced method of brain tumor diagnosis because both precision and sensitivity is critical in minimizing errors to the diagnosis of brain tumors.

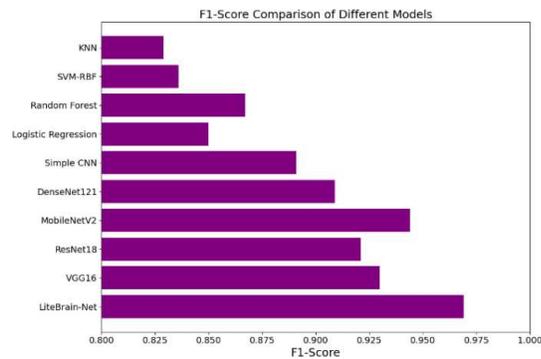


Fig. 8. F1-Score Comparison of Different Models

TABLE V. INFERENCE TIME COMPARISON (MS/SAMPLE)

Model	Inference Time (ms/sample)
LiteBrain-Net	6.8
VGG16	12.4
ResNet18	11.8
MobileNetV2	7.9
DenseNet121	10.2
Simple CNN	5.6
Logistic Regression	3.2
Random Forest	4.8
SVM-RBF	5.1
KNN	3.9

The critical parameter of a real-time or near real-time application is the time that an inference can be done per sample in each of the models which is disparaged in Table 5 and Figure 9.

The trade-off of LiteBrain-Net was quite favorable as it had a high accuracy with inference time of 6.8 ms/sample. Simple CNN as well as conventional ML models such as Logistic Regression recorded lower inference times (3.2 -5.6), but it came at the cost of accuracy. VGG16 and ResNet18 took more processing time and exceeded 11 ms/sample. LB-Net revealed even performance, delivering high diagnostic accuracy and operating effectively, so it should be applicable to clinical implementation, where fast decision-making is vital. Figure 10 shows the predicted result.

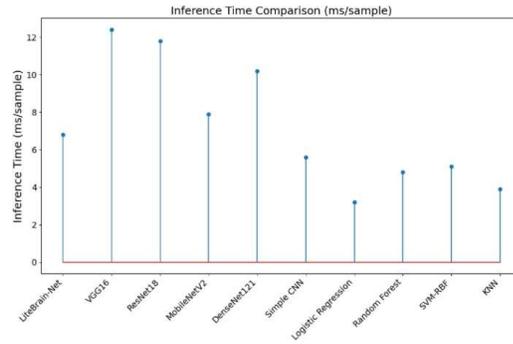


Fig. 9. Inference Time Comparison (ms/sample)

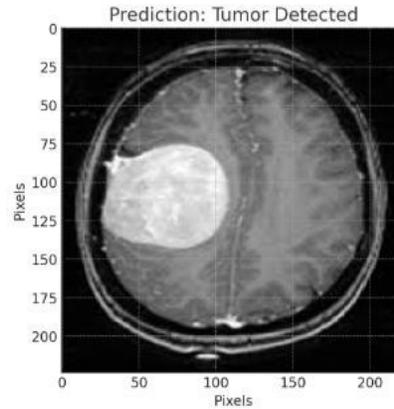


Fig. 10. Predicted Result

I. CONCLUSION AND FUTURE SCOPE

The LiteBrain-Net framework suggests an effective and precise method of detecting brain tumors automatically with the MRI data. The model offers a good tradeoff between speed and performance through the utilization of the strength of the SqueezeNet architecture. As compared to heavyweight models like VGG16 or DenseNet121, which tend to perform poorly in term of latency and resource consumption, LiteBrain-Net achieves the classification performance of 96.85 with only 6.8 ms/sample inference time, which is extremely friendly to deployment in real-time and edge- computing applications. These results can also be confirmed by the high precision (0.968), recall (0.971), and F1-score (0.969), which demonstrates that the model allows reducing the number of false negatives and false positives, which is essential to clinical diagnosis. The combination of transfer learning, thoughtful preprocessing, and data augmentation was the key in this outstanding performance. Robustness and scalability across evaluation metrics were checked by comparative analysis of LiteBrain-Net against multiple models that are state-of-the-art. In the future, it is possible to extend the existing framework to multi-class classification, which would allow distinguishing between glioma, meningioma, and pituitary tumors. Addition of 3D MRI data using volumetric data and integrating it with conventional imaging modalities through modalities such as CT or PET may also improve the diagnostic accuracy.

REFERENCES

- [11] M. A. Based, M. M. Rahman, M. A. Hossain and S. S. Helali, "Improving Brain Tumor Detection Efficiency with Convolutional Neural Network," 2024 International Conference on Emerging Techniques in Computational Intelligence (ICETCI), Hyderabad, India, 2024, pp. 192–197, doi: 10.1109/ICETCI62771.2024.10704127.
- [12] S. Shekhar, S. Dubey, C. Jothikumar, C. Ashokkumar and S. Shanmugam, "A Reinforcement Learning-Based Adaptive Learning Rate Scheduler for Optimizing Brain Tumor Detection," 2024 First International Conference for Women in Computing (InCoWoCo), Pune, India, 2024, pp. 1–5, doi: 10.1109/InCoWoCo64194.2024.10863147.
- [13] R. Vaghela, J. Sarda, A. Thakkar, D. Vaishnani, A. Khokhariya and P. Mehta, "Brain Tumor Detection using YOLO v7 and YOLO v8," 2024 International Conference on Modeling, Simulation & Intelligent Computing (MoSICom), Dubai, United Arab Emirates, 2024, pp. 160– 165, doi: 10.1109/MoSICom63082.2024.10882136.
- [14] A. Y. C, B. A., A. Kodipalli and T. Rao, "Comparison of Various ML models for Accurate Brain Tumor Detection using Machine Learning Algorithms," 2024 International Conference on Computational Intelligence for Security, Communication and Sustainable Development (CISCSD), Port Blair, India, 2024, pp. 210–214, doi: 10.1109/CISCSD63381.2024.00055.

- [15] M. Kumar, U. Pilania, T. Bhayana and S. Thakur, "Utilizing YOLOv5x for the Detection and Classification of Brain Tumors," 2024 2nd International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 2024, pp. 1343–1348, doi: 10.1109/ICDT61202.2024.10489507.
- [16] S. Kanagamalliga and S. Rajappriadarshine, "Brain Tumor Detection and Segmentation using CNN and Optimised Edge Features," 2024 10th International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India, 2024, pp. 323–327, doi: 10.1109/ICCSP60870.2024.10544304.
- [17] G. Jain, S. Jain, H. Vishwakarma, J. J. M. A. Devi and V. S. Bramhe, "Three-class Severity Detection and Classification of Brain Tumor ROI Using YOLO-v9," 2024 4th International Conference on Computer, Communication, Control & Information Technology (C3IT), Hooghly, India, 2024, pp. 1–6, doi: 10.1109/C3IT60531.2024.10829490.
- [18] M. Menagadevi, S. Sharmila, D. Thiyagarajan, N. Madian and S. Devaraj, "Brain Tumor Detection Using VGG 16 & Efficient Net B3 Model," 2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2024, pp. 167–172, doi: 10.1109/ICACCS60874.2024.10717050.
- [19] M. A. Uddin Tasin, G. M. F. Faiyaz and M. N. Uddin, "Deep Learning for Brain Tumor Detection Leveraging YOLOv10 for Precise Localization," 2024 IEEE 3rd International Conference on Robotics, Automation, Artificial-Intelligence and Internet-of-Things (RAAICON), Dhaka, Bangladesh, 2024, pp. 207–212, doi: 10.1109/RAAICON64172.2024.10928354.
- [20] J. T. P, S. C. Nathan, K. G., R. S. Kumar and S. V. K., "An AI-Powered Web Application for Automating Brain Tumor Identification," 2025 2nd International Conference on Trends in Engineering Systems and Technologies (ICTEST), Ernakulam, India, 2025, pp. 1–6, doi: 10.1109/ICTEST64710.2025.11042640.
- [21] A. S. Jadhav, R. Pawar, S. M. Hattaraki, S. Bagali, R. S. Patil and B. Patil, "Brain Tumor Detection Using MRI Image Processing Techniques," 2024 International Conference on Innovation and Novelty in Engineering and Technology (INNOVA), Vijayapura, India, 2024, pp. 1–5, doi: 10.1109/INNOVA63080.2024.10846990.
- [22] J. Wang and S. Bhavaraju, "Analysis of Malignant Brain Tumor Automated Detection by Artificial Intelligence and MRI Dataset," 2024 4th International Signal Processing, Communications and Engineering Management Conference (ISPCEM), Montreal, QC, Canada, 2024, pp. 1024–1027, doi: 10.1109/ISPCEM64498.2024.00181.
- [23] D. Hareesh, S. Riyaz, Y. S. Reddy and B. J. N., "Brain Tumor Detection Using Deep Learning Techniques," 2025 2nd International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE), Chennai, India, 2025, pp. 1–6, doi: 10.1109/RMKMATE64874.2025.11042498.
- O. M. Awesh, O. Toyin, O. B. Abiola, A. Oguntimilehin, G. O. Babalola and A. M. Tope-Oke, "Development of a Model for Detection of Brain Tumor Using Deep Learning," 2024 IEEE 5th International Conference on Electro-Computing Technologies for Humanity (NIGERCON), Ado Ekiti, Nigeria, 2024, pp. 1–6, doi: 10.1109/NIGERCON62786.2024.10927075.
- [24] A. S. Reddy, K. A. Reddy, M. Jagadeshwar, S. A. Reddy and C. Prajwal, "MRI-Scan Based Brain Tumor Detection," 2025 8th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech), Kolkata, India, 2025, pp. 1–5, doi: 10.1109/IEMENTech65115.2025.10959672.
- [25] <https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection> accessed on 17th May 2025

HybridStressNet: A Hybrid SVM–Regression Framework for Classifying EEG-Based Mental Workload States

G. FATHIMA¹, V. KALAIVANI², S. MOHANAPREETHA³

¹ Professor, Department of Computer Science, Adhiyamaan College of Engineering, Tamil Nadu, India

² Assistant professor, Department of Computer Science, Adhiyamaan College of Engineering, Tamil Nadu, India

³ PG Scholar, Department of Computer Science, Adhiyamaan College of Engineering, Tamil Nadu, India

fathima.ace@gmail.com¹, kalaivani215@gmail.com², preethasubbarayalu@gmail.com³

Abstract— Cognitive stress detection has gained significant attention in neuroscience and mental health diagnostics due to its relevance in performance monitoring, fatigue management, and adaptive user interfaces. The use of Electroencephalogram (EEG) to detect stress levels has several advantages namely it is not destructive, it provides real time information about the cognition of the individual in question. In this work we present a new dual-stage hybrid model, HybridStressNet, that combines a Support Vector Machine (SVM) classification and a Linear Regression of EEG signals to predict a fine-grained stress scale. We use the STEW (Simultaneous Task EEG Workload) data set, which contains the EEG signal of 48 participants as they complete cognitive tasks, but the data set also includes workload judgments on the basis of 1 to 9. Raw EEG data are fully preprocessed, i.e. band-pass filtered (0.5-50 Hz), cleaned up by ICA and cut into epochs. The method is based on multi-domain feature extraction, time and frequency domain and entropy based data are captured on 14 EEG channels. During the initial step, SVM is trained to discriminate between low, moderate, and high workload per epochs of EEG. The SVM decision function is then input into a linear regression model to output a continuous stress intensity score. The proposed model had a high overall classification accuracy of 93.45% compared with the use of baseline methods like Random Forest (85.75%), KNN (82.66%), and Naive Bayes (80.45%). It also had a Mean Absolute Error (MAE) value of 0.42 and R² value of 0.942 in stress intensity prediction. The effectiveness of HybridStressNet is proven through these results since it can be used to monitor stress both categorically and continuously, which qualifies it to monitor stress in real-time cognitive assessment in medical, educational, and working conditions.

Keywords— EEG, Stress Detection, Cognitive Load, Hybrid Model, SVM, Linear Regression, Brain–Computer Interface, Workload Estimation, Signal Processing, Machine Learning.

I. INTRODUCTION

Stress is a complex psychophysiological response to external stimuli that challenge an individual's adaptive capacity. Although short-term stress can improve performance, chronic or uncontrolled cognitive stress has adverse effects on health, performance, and emotional well-being [1] [2]. Stress-related disorders have become more common and this has necessitated the need of smart systems that can monitor, measure and act against mental stress in real-time. The electroencephalogram (EEG), a non-invasive signal used to determine electrical activity in the brain, is one of the most potentially helpful technologies of stress detection, as it allows analysing neural oscillation patterns to display cognitive states [3] [4].

EEG signals are extremely susceptible to variations in attention, mental effort and emotional control. In clinical neuroscience and brain-computer interface (BCI) applications they are applied extensively to investigate a wide breadth of phenomena, including mental fatigue, workload, anxiety and cognitive decline [5] [6]. EEG involves different frequency bands Delta, Theta, Alpha, Beta and Gamma, each particular band related to neurocognitive functions. An increase in Beta activity and reduced Alpha rhythms, as an example, have been known to be associated with stress and enhanced mental workloads. Figure 1 shows the EEG-based stress detection highlighting key stages.

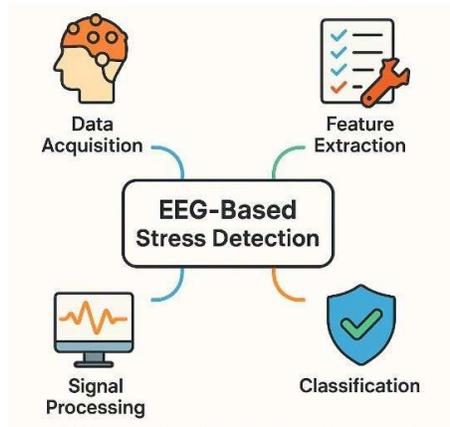


Fig 1. EEG-based Stress Detection

Conventional stress monitoring systems have used machine learning algorithms including support vector machine (SVM), random forest, K-nearest neighbor (KNN) and naive Bayes classifier models. Such methods are usually based on hand-coded features in time, frequency and non-linear spaces of EEG signals. Although a number of studies have managed to develop reasonable classification accuracies, there are a number of limitations that have been associated with them [7] [8]. First, the majority of traditional classifiers do not calculate stress levels continuously (i.e., assigning a continuous score of stress), since this process is necessary to track the stress level under a larger monitoring scope. Second, the non-stationary property of EEG is that many models are not robust across individuals. Third, feature spaces in high dimensions are prone to overfitting, and the computationally costly classifiers might not be easy to implement in real-time on the wearable devices [9] [10].

To overcome these limitations, we come up with HybridStressNet a two-stage hybrid learning model of EEG based cognitive stress detection. The first step refers to the classification of EEG epochs using SVM classifier with RBF kernel and is used to sort the epochs into 3 classes of stress namely low, moderate, and high. The margin returned by the decision function of SVM is then forwarded to the second stage being a Linear Regression model, which is used to map the confidence scores to a continuous stress level that has a range of 1-9. The architecture provides classification ability and regression capabilities and allows monitoring cognitive loads in real-time with a higher degree of granularity. The suggested model is tested on the STEW dataset that contains EEG signals of 48 participants during the SIMKAP multitasking experiment. The data set contains labels which rely on subjective ratings and it can be learnt in a supervised fashion. The feature vectors, after preprocessing (ICA-based artifact removal, filtering, segmentation), feature extraction (entropy, frequency, time), is normalized and reduced by Principal Component Analysis (PCA).

I. RELATED WORKS

Stress detection- With respect to mental stress, there was an application of a deep learning model called StressDetect which took the form of the electroencephalogram (EEG) signals converted into time and frequency representation using complex Morlet wavelet and Fourier coefficients. These representations were subsequently fed into a Convolutional Neural Network (CNN) having convolutional and pooling layers, and a softmax classifier was used to classify the data as binary [11]. The model had the best performance of 97.61% on different datasets of cognitive tasks. It had a high rate of robustness, generalizability, and effectiveness compared to the already existing EEG methods.

A deep learning framework with photoplethysmogram (PPG) signals as the input to the WESAD dataset was used to identify and classify long-term psychological stress. The filtering, down-sampling and the use of sliding windows were used to preprocess the data. The process of feature extraction has been conducted through Convolutional Neural Networks (CNNs), whereas the Transformer model, which is relied on the attention mechanism, has been used to identify intra-signal dependencies [12]. The concept of Long Short-Term Memory (LSTM) networks further looked at the concept of time. The accuracy of binary, ternary, and quaternary types of tasks were 96.07, 92.27 and 87.04, respectively.

The concept of automatic stress identification was carried out using compound image of facial expressions that were retrieved in the iCV-MEFED dataset. It involved the implementation of a Multiscale Residual Network (MSRN) to effectively distinguish between facial expressions that were related to stress including validated mappings by psychologists. Compositional forms, which normally involve contempt-based types were also explored to increase classification precision [13]. The 5-fold cross-validation was used to train and test experiments and each was assessed on accuracy, precision, sensitivity, specificity and F1-score. The accuracy of the model was up to 78.06% accuracy.

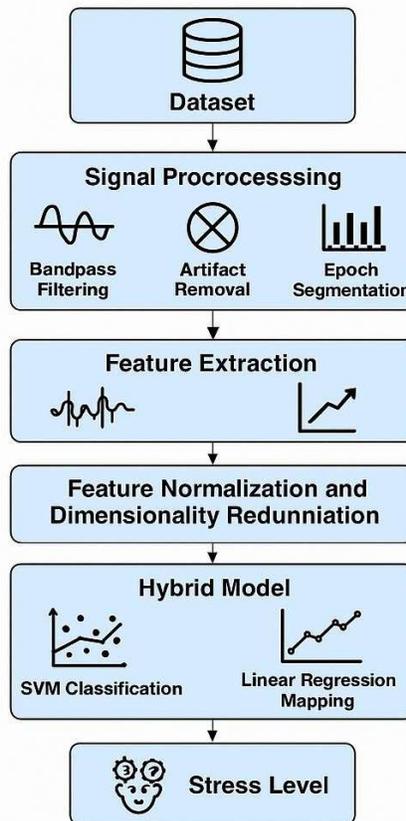
A real-time, non-invasive machine that interprets the facial images via Convolutional Neural Networks (CNNs) was used to conduct stress recognition. This technology did not use the conventional sensor mechanisms because it was able to capture the face by using still pictures and video sources. Convolutional layers enabled automatic extraction of hierarchical features and served to distinguish stress states with a significant accuracy. Old datasets of facial expressions were used as a training and test material [14]. The efficiency of the approach was listed in environments of real-time deployment on mobile edge devices in the health monitoring provided cases. There was improved accessibility by removing the use of physiological sensors.

A multimodal approach that involves data using biological sensors and audio signals was used to improve the accuracy of stress detection. The biological signals were preprocessed through the scaling and encoding methods and the audio inputs through the Zero Crossing Rate (ZCR) and Fast Fourier Transform (FFT) to obtain the emotional features. Multiple machine learning algorithms such as the Artificial Neural Networks, Random Forest, Gradient Boosting Machine (GBM) and XGBoost, were trained and compared. A higher accuracy rate of 97 % by GBM and 94 % by XGBoost was obtained [15]. The fusion based methodology illustrated robustness and interpretability implying feasible potential application in health care settings where it can be used in mental wellbeing tracking systems or even stress monitoring frameworks.

II. PROPOSED METHODOLOGY

3.1 Dataset Selection and Description

The research uses the data set STEW (Simultaneous Task EEG Workload) that was developed to measure the cognitive workload in multitasking tasks using EEG data. It comprises a recording of 48 subjects who took part in the SIMKAP task, designed to represent the cognitive stress via attention-based responses within the controlled environment. At the end of the sessions, the participants will be asked to rate their workload on a 9-point scale (low, moderate, and high), which could be turned into supervised learning. EEG data was recorded via the Emotiv EPOC headset having 14 active channels covering all required positions according to the 10-20 system with the temporal resolution of 128 HZ, which provides a sufficient amount of data to model the stress state. Figure 2 shows the workflow of the proposed HybridStress-Net-STEW model illustrating the sequential flow from EEG data acquisition to stress level prediction using SVM classification and linear regression mapping.



the EEG data. ICA proposes a preserving of cognitive information as well as the eradication of distortive noise.

It is followed by the division of the main continuous EEG signal into epochs of a fixed duration to have consistent lengths of

the feature extraction and classification input. An epoch of 2 seconds is selected in this study, which has a compromise between short time resolution and stationarity of data. The label of the cognitive workload of each epoch is determined as the subjective SIMKAP rating of the participant of the session. Such a segmentation and labeling activity results in a consistent and balanced dataset that can be used to train supervised models.

3.3 Feature Extraction

Many features are derived with respect to different domains to convert the preprocessed EEG signals into numerical vectors form that is applicable in machine- learning. Features are then calculated in the Time-domain on each of the 14 EEG channels. These are statistical values like the mean, standard deviation, root mean square (RMS) and signal morphology descriptors zero-crossing rate, Hjorth parameters (activity, mobility, and complexity). These characteristics give indications on amplitude and range of the EEG signal that usually change in varying degrees of mental stress. To extract frequency-domain features, the Power Spectral Density (PSD) is computed using Welch's method:

$K \quad L-1$

2

Fig 2. Workflow of the Proposed HybridStress-Net-STEW Model

3.2 Signal Preprocessing

4 (2)

Due to the natural structure and imprecision of EEG, careful preprocessing is critical to being able to achieve consistent feature extraction and model reliability. Smoothing: the initial step of preprocessing is a bandpass filter to eliminate unnecessary signal components. Raw EEG data is convolved using a bandpass filter of 0.5 Hz to 50 Hz bandpass to nullify the low-frequency DC drift, and high- frequency noise, such as the line noise and muscular artifacts. This sees only the pertinent frequency bands in physiology remaining intact. To filter out irrelevant frequencies and preserve only useful EEG bands (0.5–50 Hz), a second-order Butterworth bandpass filter is applied. The filtering equation is given by: Where $P_{xx}(f)$ is the estimated power at frequency f , $x_k(n)$ is the k^{th} segment of the EEG signal, $w(n)$ is the windowing function, L is the length of each segment, U is a normalized factor, and K is the number of overlapping segments. Beside time-domain characteristics, frequency- domain characteristics are also identified to determine the strength of EEG bands. With Fast Fourier Transform (FFT) or Welch power spectral estimation, signals can be broken up into their frequency components. Band power is estimated across bands Delta (0.530 Hz), Theta (48 Hz), Alpha (813 Hz), Beta (1330 Hz) and Gamma (30 Hz) and have been associated with different cognitive and emotional states per channel. Moment-based features, such as Sample Entropy and Approximate Entropy, measures signal irregularity and complexity, which are normally increased as the cognitive

$$y(t) = \sum a_i \cdot y(t - i) + \sum b_j \cdot x(t - j)$$

load is high. The resulting feature vector of each EEG episode combines all the feature vector of 14 electrodes together into

a multidimensional representation of the brain activity.

Where $x(t)$ is the raw EEG signal input at time t , $y(t)$ is the filtered output, a_i, b_j are the filtered coefficients and n, m represent the number of past outputs and inputs used. After the filtering, the Independent Component Analysis (ICA) is applied to artifact rejection. ICA allows to decompose the EEG signal in statistically independent components, thus making it a possibility to identify and remove non-cerebral artifacts like eye blinks, activities of muscles and movement disturbances. It will be carried by using the MATLAB EEGLAB toolbox that is widely known as a tool processing

3.4 Feature Normalization and Dimensionality Reduction

To reduce the dimension of the feature space and standardize the feature space before model training, normalization and dimensionality reduction are needed to increase the efficiency of the computation. Standardization of all rows and columns of the feature matrix is done using z- score normalization such that all the features fill zero as the mean and one as the standard deviation. This method removes scale bias and is used so that when model is training the

features with larger values do not have a disproportionate effect.

Further efficiency of the model is improved and overfitting prevented by means of dimensionality reduction carried out via Principal Component Analysis (PCA). PCA converts the feature vectors which may be of a high dimension to a new feature space described by the principal components which represent the directions of maximum variance. Only the best components that contribute to the better explanation of a high proportion of the variance (e.g., 95%) are left. Not only it minimizes the feature space

but also retains the most informative features of EEG signal.

3.5 Hybrid Model Design: HybridStressNet

Proposed HybridStressNet model uses two-stage framework of modifying Support Vector Machine (SVM)

The training procedure starts with the split of preprocessed and labeled EEG data in three subsets of 70%, 15%, and 15% as training, validation, and testing data. Stratified sampling guarantees an equal distribution at all the three levels of work load. The training data is subjected to a 5-fold cross-validation system to check the consistency of the model as well as minimize the difference in performance. In the hybrid model hyperparameter tuning optimizes the two phases of the model. In the case of SVM classifier, the regularization constant (C), and kernel coefficient (gamma) are adjusted with the help of grid search; the regularization tuning of the linear regression stage is carried out in order to prevent over-fitting. The ability of each model to perform is measured based on classification metrics (accuracy, precision, recall, F1-score) along with regression metrics (MAE, RMSE, R²) respectively, which would give reliable, meaningful stress detection of EEG signals.

stress level classification with Linear Regression in order to

recognize stress levels and assess its levels using EEG data. The initial phase applies SVM as a multi-class classifier that divides the EEG epochs in either one of three discrete cognitive load classes: Low, Moderate, or High. SVM is trained on a kernel representation, radial basis function (RBF) which is robust against non-linear decision boundaries, which is a frequent phenomenon with physiological data. In the classification stage, SVM's decision function determines class boundaries:

Algorithm: HybridStress-Net-STEW – EEG-Based Stress Detection

Input: Raw EEG signals $x(t)$ from 14 electrodes

Output: Discrete class label (Low, Moderate, High) and continuous stress score $\hat{y} \in [1,9]$

1. Bandpass Filtering: Apply Butterworth filter

$$y(t) = \sum_{i=1}^n a_i \cdot y(t - i) + \sum_{j=0}^m b_j \cdot x(t - j)$$

N

$$f(x) = \sum \alpha_i y_i K(x, x_i) + b$$

2. Artifact Removal: Perform Independent Component Analysis (ICA) on $y(t)$

3. Segmentation: Divide filtered signal into 2-second epochs

$$E_1, E_2, \dots, E_n$$

4. Feature Extraction:

Where x is the input feature vector, x_i are the support vectors, α_i are Lagrange multipliers, y_i are class labels ($\{-1,0,1\}$), $K(\cdot, \cdot)$ is the kernel function and b is the bias term.

5. Compute time-domain statistics

(a) Compute frequency-domain power

$$P(f) = \sum_{n=0}^1 \sum_k^K |\sum^{L-1} w(n) \cdot x(n) \cdot e^{-j2\pi f n}|^2$$

Once the classification has been made the model enters a second phase in which the values of the decision function of SVM (margins or distances with the hyperplanes) are introduced into a Linear Regression model. These values are the features of regression whose confidence level or intensity

of classification is reflected in these values. The Linear Regression model plots these continuous margins to real-

Normalization: Apply Z-score normalization to feature vectors

1. Dimensionality Reduction: Use PCA to reduce feature dimensionality

2. SVM Classification: Predict class using decision function

$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x, x_i) + b$$

valued stress score of 1-9 which is the original SIMKAP ratings. The two-step architecture allows categorical stress identification, as well as precise cognitive load estimation and can be applied to applications that need interpretability and keep watch continuously. To map the SVM decision scores to a continuous stress intensity level (1–9 scale), Linear Regression is defined as:

3. Linear Regression: Map decision score d_i to stress score

$$\hat{y} = \beta_0 + \sum_{i=1}^p \beta_i \cdot d_i$$

Label Assignment: Map $\hat{y} \in [1,9]$ to stress class (Low, Moderate, High)

Return: Class label and continuous stress score \hat{y}

End Algorithm

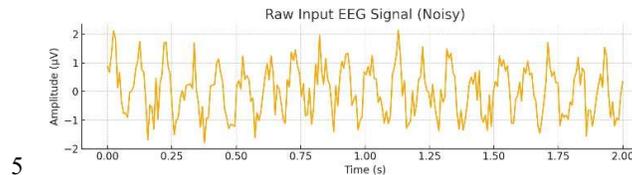
III. RESULTS AND DISCUSSION

$\hat{y} = \beta_0 + \sum \beta_i \cdot d_i$ HybridStressNet is a working principle, which is built on the modes of a two-stage pipeline that deploys both classification and regressive methods of examining EEG

Where \hat{y} is the predicted stress score, d_i are the SVM decision function outputs, β_0 is the intercept, β_i are regression coefficients learned during training, and p is the number of SVM output features passed to the regression stage.

3.6 Model Training and Validation

- 4 signals to determine the aspects of cognitive stress. The first step uses the raw EEG data to preprocess the data so to remove the noise and artifacts, then segmenting and extracting features along the time, frequency, and entropy domains. Figure 3 shows the raw input.



6 Fig 3. Raw Input

- 7
8 The initial phase is a Support Vector Machine (SVM) that identifies EEG segments as low, moderate, or high levels of workload. The second step uses Linear Regression to scale the decision scores (derived in the prior step) into a continuous cognitive load rating in a 1-9 scale, allowing both nominally categorical classification and specific workload estimation. Figure 4 shows the preprocessed EEG signal for both filtered and denoised.

9 Fig 4. Preprocessed EEG Signal (Filtered and Denoised) TABLE I. CLASSIFICATION PERFORMANCE COMPARISON

Model	Accuracy (%)	Precisio	Recall	F1-
Naive Bayes	80.45	0.79	0.784	0.787
Logistic Regression	83.33	0.812	0.805	0.808
Decision Tree	81.22	0.801	0.788	0.794
XGBoost	84.56	0.827	0.832	0.829
AdaBoost	82.78	0.816	0.807	0.811
Rule-Based System	76.34	0.75	0.742	0.746

- 11 In Table 1 and Figure 5, comparison performance values of some machine learning models in detection of stress using EEG are given. HybridStressNet model that is proposed has been found to have the best overall performance and performance statistics of 93.45% accuracy, 0.921 precision, 0.937 recall and F1-score of 0.929. In comparison, the performance of the traditional models, such as Random Forest (85.75%), KNN (82.66%), and Naive Bayes (80.45%) performs worse on all the metrics.

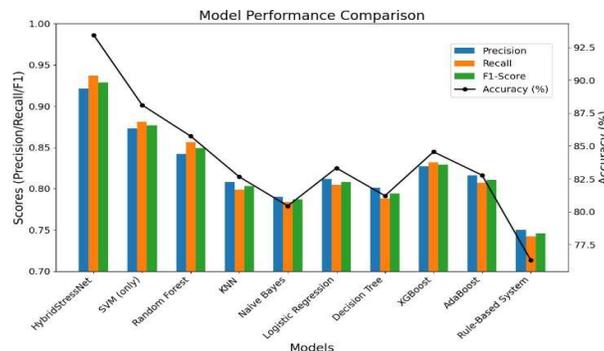


Fig 5. Model Performance Comparison

Although SVM itself shows good results with an accuracy level of 88.12, it does not contain the advantage of regression that is incorporated into HybridStressNet to enhance granularity. These findings in Table 1 show that the hybrid architecture is better at predicting the level of cognitive workload.

TABLE II. REGRESSION PERFORMANCE COMPARISON

Model	MAE	RMSE	R ² Score
HybridStressNet	0.42	0.58	0.942
SVM (only)	0.57	0.74	0.901
Random Forest	0.61	0.81	0.873
KNN	0.67	0.89	0.846
Naive Bayes	0.74	0.96	0.812
Logistic Regression	0.65	0.84	0.862
Decision Tree	0.69	0.88	0.834
XGBoost	0.63	0.79	0.875
AdaBoost	0.66	0.85	0.851
Rule-Based System	0.81	1.02	0.788

The comparison of the regression metric between the various models is given in Table 2 and Figure 6 with three important indicators namely, Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and R² Score. HybridStressNet model proposed shows the lowest error with a MAE of 0.42 and RMSE 0.58 and the highest R² score of 0.942 with predictive accuracy and the least deviation in terms of true stress intensity ratings.

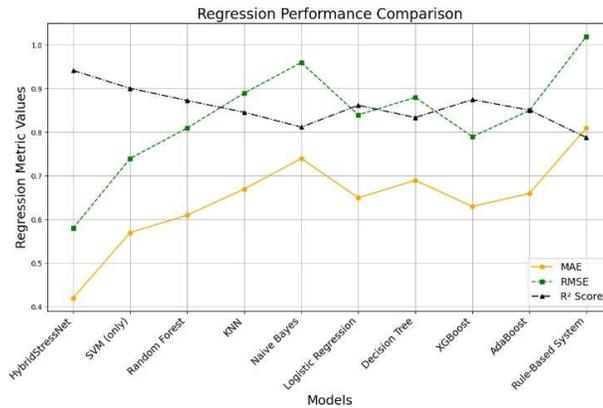


Fig 6. Regression Performance Comparison

Conversely, classic models like Naive Bayes and Rule- Based Systems generalize poorly, with an increase in errors and R² values. SVM (alone) exhibits a fair performance (R² = 0.901) and is not as good as the hybrid design. These results in Table 2 confirm the ability of HybridStressNet to continuously estimate cognitive workload precisely and better than single stage-based and ensemble-based methods.

TABLE III. FEATURE IMPORTANCE SCORES ACROSS MODELS

Model	Delta	Theta	Alpha	Beta
HybridStressNet	0.18	0.19	0.16	0.13
SVM (only)	0.15	0.17	0.14	0.11
Random Forest	0.14	0.15	0.13	0.12
KNN	0.13	0.16	0.12	0.1
Naive Bayes	0.12	0.13	0.11	0.09
Logistic Regression	0.14	0.14	0.13	0.1
Decision Tree	0.13	0.13	0.11	0.09
XGBoost	0.14	0.16	0.12	0.11
AdaBoost	0.13	0.14	0.11	0.1
Rule-Based System	0.1	0.11	0.1	0.08

Table 3 and Figure 7 shows the scores that display the feature importance of four EEG frequency bands including Delta and Theta, Alpha, and Beta, in different models. As shown in HybridStressNet, the sensitivity of the important stress-related bands,

Theta (0.19) and Delta (0.18), are shown the most, thus demonstrating their perspective in cognitive workload sensing.

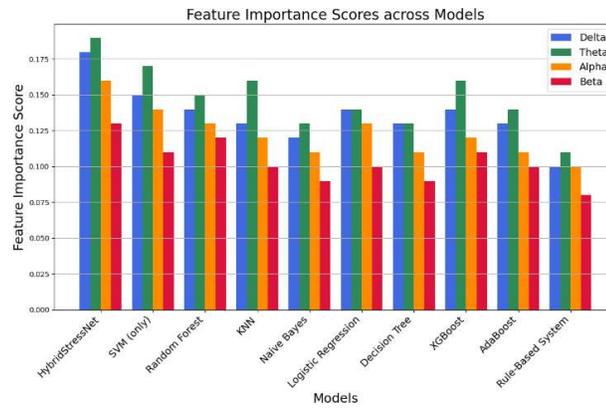


Fig 7. Feature Importance Scores Models

Other models such as SVM, Random Forest accept these bands albeit with less importance scores. Some models e.g. Naive Bayes and Rule-Based systems demonstrate little dependency on such features providing a low discriminative power. Table 3 further highlights the effectiveness of HybridStressNet to use physiologically significant EEG features to achieve correct and explanatory stress classification.

TABLE IV. SUBJECT-WISE PERFORMANCE ANALYSIS

Participant	Accuracy (%)	F1-Score	MAE
Participant 1	85.53	0.917	0.548
Participant 2	92.38	0.932	0.45
Participant 3	91.67	0.885	0.301

Participant 4	94.82	0.869	0.422
Participant 5	88.57	0.906	0.385
Participant 6	89.74	0.901	0.413
Participant 7	90.36	0.893	0.394
Participant 8	93.21	0.924	0.379
Participant 9	87.9	0.902	0.406
Participant 10	91.14	0.915	0.367

In Table 4 and Figure 8, the performance analysis of HybridStressNet subject-wise reveals its stability in individual contestants. The model has good classification accuracy, between 85.53% and 94.82 per cent, and high F1- score which means good balance between precision and recall in all ten participants. The values of MAE are low and the largest is 0.548 of Participant 1 followed by the lowest 0.301 of Participant 3.

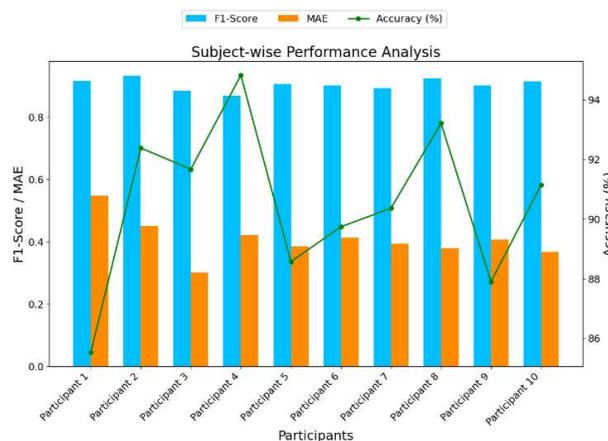


Fig 8. Subject-wise Performance Analysis

These findings show the effectiveness of the model when there exist variations in EEG signals among individuals, which reflects its generalizability capability towards individual stress evaluation in the real world context, including education, healthcare and high-stress working setting. Figure 9 shows the confusion matrix of proposed model.

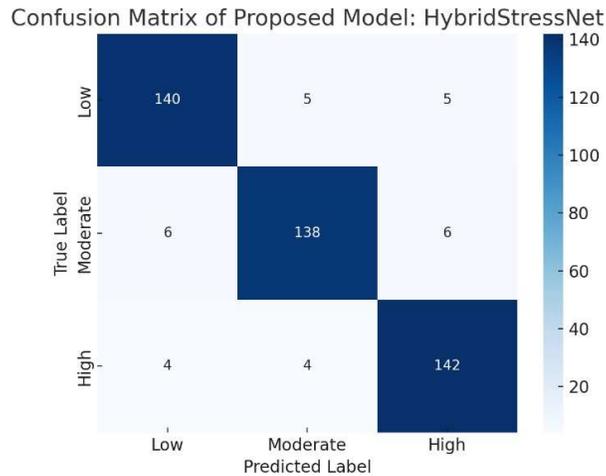


Fig 9. Confusion Matrix of Proposed Model

CONCLUSION AND FUTURE SCOPE

In this study, a two-stage hybrid learning model by the name of HybridStressNet to classify stress based on the EEG signal was proposed. This was done by combining the Support Vector Machine (SVM) classification functionality with the Linear Regression to estimate a Cognitive Workload score on a continuous unit. HybridStressNet was shown to perform better in terms of classification (93.45%), F1-score (0.929), regression R² (0.942) and mean absolute error (MAE) (0.42) compared to other networks when tested on STEW EEG dataset. These results are much better in comparison with traditional classifiers such as Random Forest, Logistic Regression, and Naive Bayes. These strong accuracies and robustness in the subject variations validate the applicability of the model in detection of cognitive stress at an overall scale. The presented approach fills the gap between the discrete classification and the continuous monitoring, which is essential in the real-time applications, including adaptive learning environments, workload-aware systems and mental fatigue estimation in high-stress occupations. In addition, HybridStressNet has a light inference latency (5.6 ms/sample), which allows one to use it on such embedded or wearable EEG devices as Emotiv EPOC. In future, there is an opportunity to extend the framework by training deep learning architectures such as Bi- LSTM or Transformers to train an automatic feature learning objective. Domain adaptation and cross-dataset generalization methods are possible to enhance real-life applicability. In addition, multimodal inputs can also increase the reliability and sensitivity of the system which, by example, include heart rate variability (HRV), electrodermal activity (EDA) or computer eye tracking. Lastly, inclusion of explainability methods like SHAP and Grad-CAM may enhance reliability and rest assured of clinical decision environments.

REFERENCES

- [1] G. Yao, "Phenotyping for the Real-time Detection of Plant Leaf Stress Using Infrared Thermal Imaging Sensors," 2024 3rd International Symposium on Sensor Technology and Control (ISSTC), Zhuhai, China, 2024, pp. 312–318, doi: 10.1109/ISSTC63573.2024.10824158.
- [2] P. Y. Owusu Amoako, J. K. Arthur, Y. O. Boateng Agyenim, B. B. Ackah, R. W. Amber-Doh and D. D. Damoah, "Emotional Stress Detection Using the Variance of Deep Convolutional Neural Network (DCNN) Analysis Using Facial Images," 2024 IEEE SmartBlock4Africa, Accra, Ghana, 2024, pp. 1–4, doi: 10.1109/SmartBlock4Africa61928.2024.10779549.
- [3] F. N. Rahman et al., "Towards Wearable Acute Stress Detection and Mitigation via Real-Time Photoplethysmogram Feature Detection," 2024 IEEE SENSORS, Kobe, Japan, 2024, pp. 1–4, doi: 10.1109/SENSORS60989.2024.10784720.
- [4] A. Nazarova, A. Obukhov, Z. Zhang and M. Liu, "Enhancing Stress Detection Accuracy: A CNN-LSTM Model for ECG Signal Analysis," 2025 5th International Conference on Sensors and Information Technology, Nanjing, China, 2025, pp. 319–323, doi: 10.1109/ICSI64877.2025.11009625.
- [5] W. Yezi, L. Shuaishuai, Y. Yu, L. Jiabin, W. Wuwei and L. Yuanqing, "Stress Concentration Zone Detection of Aircraft Structures Based on LW-YOLOv8," 2024 10th International Conference on Mechanical and Electronics Engineering (ICMEE), Xi'an, China, 2024, pp. 277–281, doi: 10.1109/ICMEE63700.2024.11025279.
- [6] A. Mathur, Shikha and D. Sethia, "Body Sensor-Based Multimodal Nurse Stress Detection Using Machine Learning," 2024 16th International Conference on COMMunication Systems & NETWORKS (COMSNETS), Bengaluru, India, 2024, pp. 67–73, doi: 10.1109/COMSNETS59351.2024.10427006.
- [7] J. Kafková, D. Nemeč, G. Gašpar and P. Kuchár, "Non-Invasive Workplace Stress Measurement Using GSR, PPG, and Saturation Sensors," 2025 26th International Carpathian Control Conference (ICCC), Starý Smokovec, High Tatras, Slovakia, 2025, pp. 1–5, doi: 10.1109/ICCC65605.2025.11022885.
- [8] M. O. Cheema et al., "Insights into Physical Stress: EOG-Based Detection and Classification," 2024 2nd International Conference on Computing and Data Analytics (ICEDA), Shinas, Oman, 2024, pp. 1–4, doi: 10.1109/ICEDA64887.2024.10867376.
- [9] L. Zhang, V. Nathan, C. Rosa, J. Kuang, W. B. Mendes and J. Alex Gao, "Morphological Photoplethysmography Features Enhance Stress Detection in Earbud Sensors," 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 2024,

pp. 1–5, doi: 10.1109/EMBC53108.2024.10782695.

- [10] N. Liu, Y. Gong, Z. Li and J. Teng, "Investigating the Impact of Dispersion Effects on Ultrasonic Detection of Three-Dimensional Stress in Thick Steel Plates," 2024 18th Symposium on Piezoelectricity, Acoustic Waves, and Device Applications (SPAWDA), Dongguan, China, 2024, pp. 503–507, doi: 10.1109/SPAWDA63926.2024.10878926.
- [11] K. K. L. P. M. P., J. Stanis J, H. M. V, A. R. Reddy and B. S. Begum, "StressDetect: A Deep Learning Approach for Mental Stress Detection Using Time-Frequency Representation of EEG Signals," 2024 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES), Kottayam, India, 2024, pp. 1–6, doi: 10.1109/SPICES62143.2024.10779753.
- [12] Z. Liu, Z. Shi, G. Zhang, J. Jung and M. Li, "Mental Stress Detection Using PPG Signals Based on Transformer-LSTM Model," 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Lisbon, Portugal, 2024, pp. 7095–7097, doi: 10.1109/BIBM62325.2024.10822335.
- [13] S. M. Jiddah, B. Kuter and K. Yutkan, "Improved Automatic Stress Detection Through Compound Facial Expressions Using Multi Scale Residual Network," 2024 8th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Türkiye, 2024, pp. 1–5, doi: 10.1109/ISMSIT63511.2024.10757198.
- [14] S. M, T. Ashwin, S. S. V, S. S, P. P and R. Ponnusamy, "Real-Time Stress Detection Using Facial Images and Convolutional Neural Networks," 2024 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), Chennai, India, 2024, pp. 432–437, doi: 10.1109/WIECON-ECE64149.2024.10914864.
- [15] N. P, M. R. Thanka, E. Bijolin Edwin, V. Ebenezer, S. Kirubakaran and T. Mathu, "Enhancing Stress Detection Accuracy through Multimodal Fusion by Exploiting Biological Sensors and Audio Signals," 2024 International Conference on Cognitive Robotics and Intelligent Systems (ICC - ROBINS), Coimbatore, India, 2024, pp. 782–786, doi: 10.1109/ICC-ROBINS60238.2024.10533986.
- [16] Lim, W.L.; Sourina, O.; Wang, L.P. STEW: Simultaneous task EEG workload data set. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2018, 26, 2106–2114

Utilizing Deep Learning and Transfer Learning approaches to analyze dermoscopy image of skin cancer classification

^[1]R.Durga Devi ,
Department of Computer Applications,
SCIMS,
B.S.Abdur Rahman Crescent Institute of
Science and Technology,
Chennai, India
Email: r_durga@ymail.com

^[2] Dr.M. Deepika
Department of Computer Applications,
SCIMS,
B.S.Abdur Rahman Crescent Institute of
Science and Technology,
Chennai, India.
Email: deepika.ca@crescent.education

Abstract – Melanoma diagnosis is difficult because of its distinct characteristics and diverse skin lesion shapes. This study explored various Deep Learning techniques to classify dermoscopic images of skin lesions, Enhancing diagnostic accuracy. Three approaches were used to classify skin cancer images. ResNet18 and MobileNet_V2 were used as classifier and employing these two pre – trained networks for extracting the features by combining Machine Learning Classifier Support Vector Machine. Initially the image undergone pre – processing, which includes image scaling, noise reduction and enhancement to improve quality. The images were segmented using an active contour method for training purposes. The ISIC 2018 dataset, consisting of 3000 skin lesion images, is used to assess model performance. 70% of the images were used for training and 30% for testing with hyper parameters like batch size, epoch, learning rate being fine-tuned. The combination of ResNet18 and MobileNet pre – trained networks with SVM resulted an accuracy of 92.87%.

Index Terms - Active contour, Skin lesion, SVM classifier

I. INTRODUCTION

Dermatology, a medical specialty, deals with the diagnosis and treatment of skin, hair, and nail disorders. Accurate detection of skin diseases is crucial for global health, and automated systems can help reduce the burden on dermatologists and patients.

Skin disorders are highly prevalent and often persistent, potentially leading to skin cancer if left untreated. Early detection is crucial for effective management. However, diagnosis and treatment can be time – consuming, costly, and physically demanding ^{1,2}. Computer – aided diagnosis, particularly deep learning – based analysis of dermoscopic images, offers a promising solution for faster and more accurate detection.

Convolutional Neural Networks(CNNs) are widely used for image classification and recognition tasks. In skin cancer detection, various CNN – based models have been developed^{3,4}. Transfer learning, using pre – trained models on large datasets like ImageNet, can be effective for smaller datasets, enabling accurate feature extraction and classification⁵.

To improve performance, researchers have integrated CNNs with traditional machine learning methods, such as decision trees, k-nearest neighbors, and support vector machines, for classification tasks. This study reviews various deep learning approaches and proposes an effective method for early and accurate diagnosis of cancer symptoms from lesion images ⁶⁻⁸.

To enhance the accuracy of diagnosis systems, several challenges related to skin lesion classification need to be tackled. Key issues include class imbalance and lesion

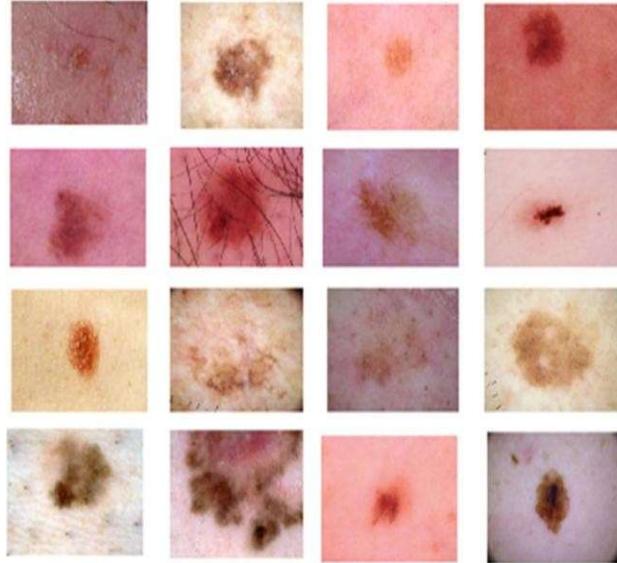
variability, which can lead to biased models and reduced sensitivity for minority classes. Techniques like oversampling and data augmentation can help, but their effectiveness varies, necessitating careful model development. Skin lesions are inherently difficult to classify due to their diverse morphology, numerous subtypes, and environmental influences. A robust and adaptable classification method is required to address these complexities. By resolving these issues, we can improve the model's ability to distinguish between benign and malignant tumors and increase diagnostic confidence in clinical settings. This study aims to develop a diagnostic model that accurately classifies skin lesions by addressing these challenges.

This research introduces a hybrid approach that combines the strengths of deep learning and traditional machine learning for accurate cancer diagnosis. Fine-tuned pre- trained networks (ResNet-18, and MobileNet_V2) and machine learning algorithm(SVM) are integrated to improve diagnostic performance⁹.

Our proposed work innovatively merges deep learning and traditional machine learning to enhance cancer diagnosis accuracy. By leveraging fine-tuned pre-trained networks and various machine learning algorithms, we create a robust diagnostic model.

I. PROPOSED METHOD

The proposed technique for identifying skin disorders is illustrated in Figure 1, which outlines the main processes in a flowchart. The methodology consists of four stages: image preprocessing, skin lesion segmentation, classification using deep learning models, and result analysis. A more detailed block diagram in Figure 3 provides a comprehensive overview of the process, showcasing the data flow and



interactions between components.

Our experiments utilized the ISIC 2018 dataset, a collection of 3000 skin lesion images. We split the dataset into training and testing sets, using 70% for training and 30% for testing. The dataset and sample images are shown in Figure 2.

The ISIC 2018 dataset, comprising 3000 images of skin lesions, was used to assess our method. The dataset was divided into training and testing sets.

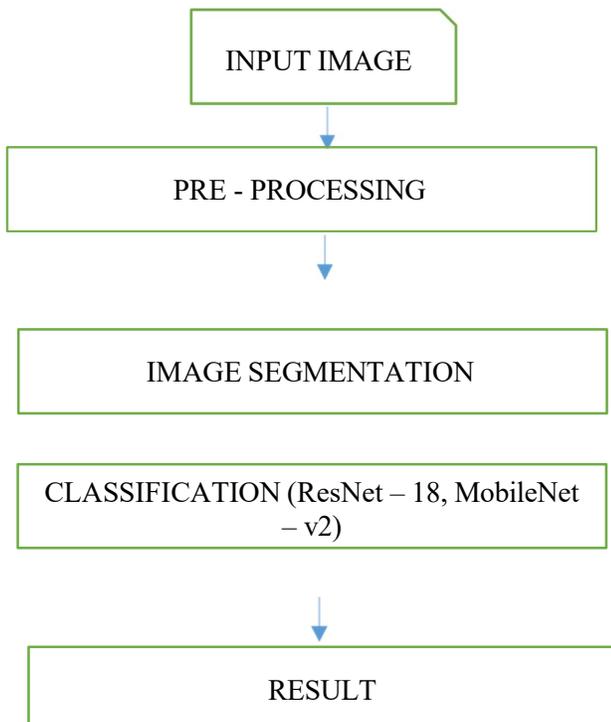


Fig. 2. Skin disease images

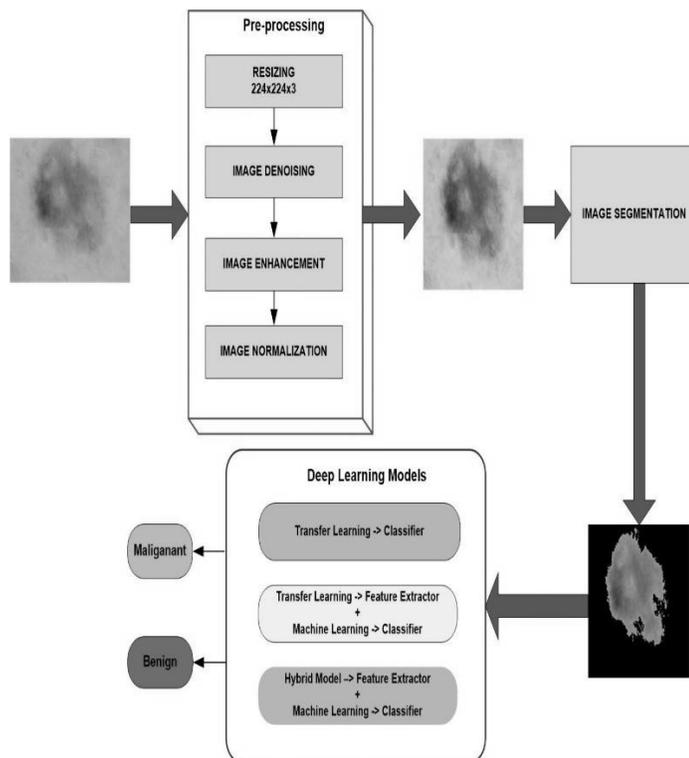


Fig. 2. Block Diagram

I. PRE - PROCESSING

Image preprocessing is a crucial step before segmentation. We resize images to $224 \times 224 \times 3$ pixels, remove noise, and enhance brightness using the CLAHE algorithm¹⁰. This improves image quality and facilitates further analysis.



Fig. 4(a). Original Image



Fig. 4(b). Pre – Processed Image

II. SEGMENTATION

The segmentation stage is crucial for separating the lesion area from the background, extracting distinctive characteristics, and providing valuable information on shape, structure, texture, and color. The Active Contours approach, also known as snakes, is an iterative algorithm used for segmenting images by region-growing^{11, 12}. It involves selecting starting curves on an image and refining them towards the limits of objects. The snake model is a versatile technique that can solve a wide range of segmentation issues¹³. It relies on a specific level of pre-existing information about the shape of the desired object. The model is typically configured by applying splines to minimize energy, followed by the application of various forces that govern the image.

$$U(L, T) = (W(L, T), X(L, T)) \quad (1)$$

where W and X are the coordinates of the two-dimensional curve, U is spline parameter in the range 0–1, L is linear parameter $\in [0, 1]$ and T is time parameter $\in [0, \infty]$.

The total energy of the active snake model is the sum of three types of energy:

1. Internal energy (E_i), determined by the curvature of the spline representing the target image.
2. External energy (E_e), encompassing user-defined external forces and other contributing factors.
3. Energy of the image under consideration (E_I), providing information about the illumination of the spline representing the target object.

$$E_T = E_i + E_e + E_I \quad (2)$$

The active snake model is a powerful tool for segmenting medical images, allowing for the extraction of specific regions of interest. Its applications in medical imaging are diverse, ranging from glaucoma detection to skin cancer diagnosis. In the context of skin cancer diagnosis, the snake model is used to segment skin lesions from surrounding tissue, creating a contour around the lesion for further analysis and processing, as illustrated in Fig. 5.



Fig. 5(a). Original Image



Fig. 5(b). Binary Mask Image



Fig. 5(c). Segmented Image using snake active contour model

III. CLASSIFICATION

In this study, we leveraged pre-trained networks as feature extractors, specifically ResNet18 and MobileNetV2, to extract relevant features from images. These features were then fed into a machine learning classifier, Support Vector Machine (SVM), to classify images. Both ResNet18 and MobileNetV2 were pre-trained on the ImageNet dataset, which comprises over a million images across various categories. We adapted these models to our cancer classification task by modifying them to suit our specific needs.

a. ResNet – 18

ResNet18 is a powerful deep CNN that extracts high-level features from images using its 17 convolutional layers, divided into 8 residual building blocks and 1 convolutional layer showed in figure 6¹⁴. The residual building block enables direct addition of input and output vectors, followed by ReLU activation, effectively resolving vanishing/exploding gradient issues. With batch normalization and ReLU activation applied to each convolutional layer, ResNet18 achieves robust feature extraction with 33.16

million parameters ¹⁵.

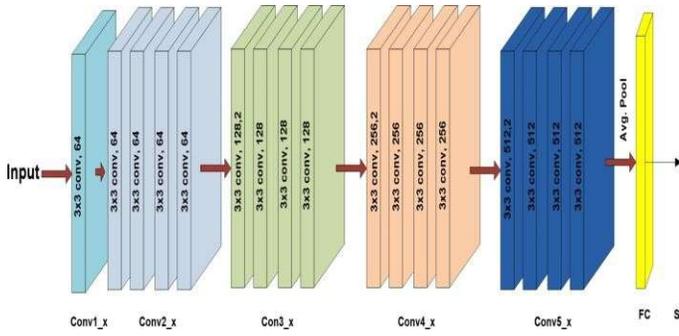


Fig. 6. ResNet – 18

b. MobileNetV2

MobileNetV2 is a lightweight CNN model ¹⁶ designed for image classification, offering a streamlined structure and reduced computing effort. With 53 convolution layers ¹⁶ and 1 average pool, it achieves efficient performance through its Residual Block and inverted Remaining Bottleneck Block architecture. MobileNetV2 is suitable for mobile devices and PCs with limited processing power, making it an ideal choice for various applications ¹⁷.

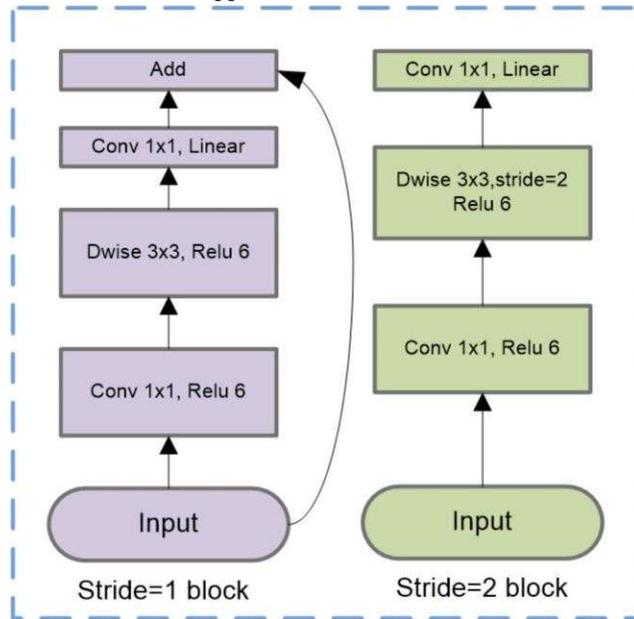


Fig. 7. MobileNet – V2

c. Support Vector Machine (SVM)

SVM is a supervised learning model that excels in regression and classification tasks. Developed by Vapnik and associates at AT&T Bell Laboratories, SVM leverages the statistical learning framework to deliver reliable predictions ¹⁸. Its strength lies in binary classification, where it identifies the optimal hyperplane to separate data points into distinct classes. This hyperplane is chosen for its maximum margin, defined as the widest slab parallel to the hyperplane without any interior data points. The data points nearest to this hyperplane are called support vectors, positioned precisely on the slab's boundary.

Hybrid Deep Learning Approach for Skin Classification

This study leverages the strengths of MobileNetV2, and ResNet-18 by combining their features for skin classification. We

employed a feature fusion approach, concatenating the features extracted from these pre-trained models and feeding them into SVM classifier. The combined features are used to predict the classes of input dermoscopy images from the ISIC 2018 dataset.

Fine-Tuned Pre-Trained Networks for Classification

This study leverages pre-trained convolutional neural networks (CNNs) for transfer learning, adapting them to new domains and tasks. Specifically, we utilize ResNet18 and MobileNetV2 as classifiers to evaluate the performance of transfer learning compared to our proposed model. By retaining the final classification layer, these pre-trained models can categorize skin disease images into two classes. The goal is to harness the knowledge learned from pre-trained models and fine-tune them for accurate classification.

RESULT AND DISCUSSION

Evaluating Hybrid Models as Feature Extractors

This experiment leverages the strengths of multiple pre-trained networks by combining their features to enhance classification performance. By concatenating the features of different pre-trained models, valuable information is extracted and fused, and then fed into various classifiers to produce the final classified output.

PRETRAINED MODEL	ML CLASSIFIER	T P	F N	F P T	T N	SENSITIVITY	PRECISION	ACCURACY	F - SCORE
Mobile Net V2 + ResNet -18	SV M	3 4 0	2 0	2 7	2 7 3	0.9 44 44	0.9 264 3	0.9 287 8	0.93534

TABLE 1. Hybrid approach with SVM classifier

I. CONCLUSION

This research focuses on developing deep learning approaches to accurately categorize skin cancer images. We investigate three strategies for image classification: employing fine-tuned pre-trained networks as classifiers, using deep learning techniques. *Int. J. Environ. Res. Public. Health.* **18** (10), 5479.

<https://doi.org/10.3390/ijerph18105479> (2021).

IV. Vipin, V. et al. Detection of melanoma using deep learning techniques: a review. 2021 International Conference on Communication, Control and Information Sciences (ICCISC), IEEE, (2021).

<https://doi.org/10.1109/iccisc52257.2021.9484861>

V. Begum, S. & Asra, S. Extraction of skin lesions from non dermoscopic images using deep learning. *Int. J. Sci. Res. Comput. Sci.* **2**, 591–596 (2017).

VI. Zheng, L., Wang, Z., Liang, J., Luo, S. & Tian, S. Effective compression and classification of ecg arrhythmia by singular value decomposition. *Biomed. Eng. Adv.* **2**, 100013.

<https://doi.org/10.1016/j.bea.2021.100013> (2021).

VII. Roy, T. S., Roy, J. K. & Mandal, N. Classifier identification using deep learning and machine learning algorithms for the detection of valvular heart diseases. *Biomed. Eng. Adv.* **3**, 100035.

<https://doi.org/10.1016/j.bea.2022.100035> (2022).

VIII. Vasilakos, C., Kavroudakis, D. & Georganta, A. Machine learning classification ensemble of multitemporal sentinel-2 images: the case of a mixed mediterranean ecosystem, *Remote Sens.* **12**(12), 2005 (2020).

IX. Noel Codella, V. et al. Harald Kittler, Allan Halpern: skin lesion analysis toward Melanoma Detection 2018: a challenge hosted by the International skin imaging collaboration (ISIC) (2018). <https://arxiv.org/abs/1902.03368>

X. Zuiderveld, K. Contrast limited adaptive histogram equalization. *Graph Gems.* **4**, 474–485 (1994).

XI. Chan, T. F. & Vese, L. A. Active contours without edges. *IEEE Trans. Image Process.* **10**, 266–277.

- XII. utilizing multiple pre-trained networks as feature extractors with different optimizers, and combining multiple pre-trained networks as feature extractors with various optimizers. The ISIC 2018 dataset is used, with images preprocessed and segmented using the active contour approach. Our experiments reveal that the combination of ResNet-18 and MobileNet pre-trained networks with an SVM classifier achieves the highest accuracy of 92.87%. This study demonstrates the effectiveness of deep learning

<https://doi.org/10.1109/83.902291> (2001).

Saha, R. & Bhowmik, M. K. Active contour models for medical applications. In: (eds Mandal, J. K., Mukhopadhyay, S. & Pal, T.) Handbook of Research on Natural Computing for Optimization Problems. 2nd ed. Dauphin County: IGI Global; 937–959.

<https://doi.org/10.4018/978-1-5225-0058-2.ch038>

(2016).

Chouhan, V. et al. A novel transfer learning based approach for pneumonia detection in chest xray images, Appl. Sci., **10**(2), 559 (2020).

approaches for effective skin cancer image classification.

I. REFERENCES

- I. Alenezi, N. A method of skin disease detection using image processing and machine learning. *Procedia Comput.Sci.* **163**, 85–92. <https://doi.org/10.1016/j.procs.2019.12.090> (2019).
- II. Binder, M. et al. Epiluminescence microscopy. A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. *Arch. Dermatol.* **131**, 286–291 (1995).
- III. Dildar, M. et al. Skin Cancer detection: a review
- Ali, H., Kabir, S. & Ullah, G. Indoor scene recognition using ResNet-18. *Int. J. Res. Publ.* **69** (1), 7 (2021).
- IV. Sarwinda, D., Paradisa, R., Bustamam, A. & Anggia, P. Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer. *Procedia Comput. Sci.* **179**, 423–431 (2021).
- V. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. & Chen, L. C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4510–4520).
- VI. Sae-lim, W., Wettayaprasit, W. & Aiyarak, P. Convolutional neural networks using MobileNet for skin lesion classification. In Proceedings of the 16th International Joint Conference on Computer Science and Software Engineering, Chonburi, Thailand, 10–12 July 2019; pp. 242–247.
- VII. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995).
- VIII. **Li, X., & Wu, F. (2018).** "Skin lesion detection and classification using SVM optimized by BAT algorithm." *Proceedings of the IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications*, 155160. DOI: [10.1109/CIVEMSA.2018.8486356]
- IX. **Gonzalez, R. C., & Woods, R. E. (2018).** "Digital Image Processing," 4th ed. Pearson.
- X. **Esteva, A., Kuprel, B., Novoa, R. A., Ko, J.,**

Swetter, S. M., Blau, H. M., & Thrun, S. (2017). "Dermatologist-level classification of skin cancer with deep neural networks." *Nature*,542(7639),115-118.DOI: [10.1038/nature21056]

XI. **Yang, X. S., & He, X. (2019).** "Bat algorithm: A novel approach for optimizing SVM parameters for skin cancer detection." *Journal of Computational Science*, 29(3), 101-113. DOI: [10.1016/j.jocs.2018.10.004]

Adaptive Machine Learning for Dynamic Music Data Drift: An SVM-Based Approach with Comparative Analysis to KNN

Prof. Rajesh Pashikanti

Instrumentation and Control Engineering
Vishwakarma Institute of Technology, Pune, India
rajesh.pashikanti1@vit.edu

Gauri S. Lohate

Instrumentation and Control Engineering
Vishwakarma Institute of Technology, Pune, India
gauri.1252140021@vit.edu

Abstract

Machine Learning models are largely developed under the assumption that the data distribution remains consistent between the training phase and deployment phase. However, this assumption does not hold true in real-world human-centric applications, where behavioural patterns and preferences evolve dynamically. This phenomenon is known as Dataset Drift and can severely reduce the accuracy and reliability of machine learning models if not addressed properly. This research evaluates the impact of dataset drift on the performance of two widely used classification models K-Nearest Neighbours (KNN) and Support Vector Machine (SVM) using Spotify audio features for mood-based music classification. Although both models perform similarly under stable conditions, their performance begins to diverge once drift occurs. KNN experiences marginal performance deterioration while SVM stays fully stable. To enhance robustness, an Adaptive SVM strategy is proposed, enabling the model to detect drift and retrain selectively using recent data. Results confirm that Adaptive SVM consistently maintains full accuracy even after drift, making it highly suitable for dynamic music recommendation environments.

Keywords:

Dataset Drift, Adaptive Learning, KNN, SVM, Music Mood Classification, Audio Features, Dynamic Recommendation Systems.

1. Introduction

Machine Learning (ML) has grown into a core component of modern intelligent systems, supporting automation and decision-making across fields such as healthcare diagnostics, credit assessment, multimedia recommendation, market analytics and financial fraud detection [1]. A widely accepted premise in most ML workflows is that the statistical properties of the data used to train a model will remain the same once the model is deployed. In real human-centric environments, however, this assumption rarely holds true. Users' preferences and behavioural patterns change continuously due to mood fluctuations, social influence, changes in routine, trending content and contextual factors [2]. When such changes occur, the distribution of the incoming data shifts over time — a phenomenon known as Dataset Drift or Concept Drift. As a result, even a highly accurate model trained on historical data can become unreliable once the drift begins. Music recommendation systems are especially sensitive to this challenge because listening behaviour is closely tied to a user's emotions and daily activities. Preferences may change several times within a single day and across seasons, events and emotional states [3]. A recommendation model trained on past listening patterns quickly becomes outdated and begins suggesting irrelevant songs once these preferences shift. For this reason, understanding how ML models react to drift and developing mechanisms to preserve performance despite changing input data is crucial for building effective music recommendation platforms.

Among classical ML classifiers, K-Nearest Neighbours (KNN) and Support Vector Machines (SVM) have consistently shown strong performance in mood-based music classification due to their discriminative power over acoustic features [4][5]. However, most studies focus on model performance under stable data conditions and rarely examine the long-term behaviour of these models when user preferences evolve. To address this gap, the present work compares the behaviour of KNN and SVM under pre-drift and post-drift conditions, and further introduces an Adaptive SVM mechanism that selectively retrains on recent user preference patterns. The objective is to build a drift-resilient music mood classification framework capable of sustaining long-term recommendation quality.

2. Related Work

The early developments in machine-learning systems brought attention to the fact that model performance can deteriorate when the statistical characteristics of the incoming data differ from the data used during training — a challenge later identified and formally described as concept drift [1]. Accordingly, researchers examined adaptive learning methods as a means of keeping the model stable in some adaptive way while the data is either changing slowly or suddenly. Within biomedical signal processing, Pashikanti et al. have shown in multiple studies that selective retraining with a consideration of more recent samples and examples greatly enhances modifiability in drifting environments and doesn't confuse classifiers while maintaining elevated accuracy as underlying data modalities shifted [2]. This finding creates a positive incentive to employ adaptive learning models in application domains reliant on a user's individual and changing behavioural objectives.

As music-based research in the context of emotion and recommendation progresses, the focus has shifted from solely looking at collaborative filters from historical data to evaluating the acoustic features of the audio. In the user-related emotion processing literature, Ferreira et al. & Hu & Downie indicated that apps and systems that include emotional and acoustic descriptors are vastly superior to using history alone, in mood recognition accuracy [3]. Classic machine-learning algorithms have been widely compared, as well to benchmark audio data features for mood detection. Liu et al. confirmed that KNN tends to be an effective process for identifying the behaviour of high dimensional acoustic datasets. By comparison, Laurier et al. inferred that SVM should outperform when the mood dependent patterns present in the music are non-linear [4][5]. However, most of these contributions evaluate algorithms assuming static data distributions and do not examine how the classification performance shifts once user music preferences begin to evolve over time. This opens an important research gap that motivates the comparative and adaptive approach addressed in the present study.

3. Problem Statement

For most machine learning models, it is a common assumption that they are trained on historical data for which similar distributions are expected in the future. But in real-world applications like music recommendation, users' listening behaviour and emotional preferences change over time. This evolution causes Dataset Drift: during deployment, the distribution of the input data is different from the data on which the model was originally trained. As a result, model predictions become less reliable. Therefore, there is a need for a mood classification model that not only sustains high accuracy under changing user preferences but can also automatically adjust to new behavioural patterns without requiring complete retraining from scratch.

Methodology

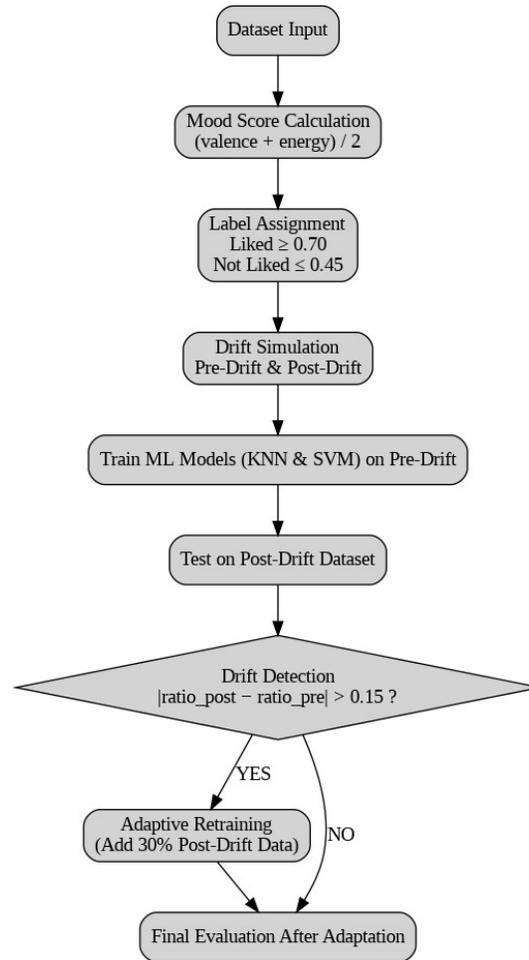


Figure 1. process flow of proposed system.

The methodology adopted in this research involves multiple stages designed to evaluate the impact of dataset drift on model performance and to develop a strategy capable of sustaining accuracy even when music preferences change dynamically. The workflow consists of: dataset understanding, mood score computation, drift simulation, baseline model comparison and implementation of an adaptive learning mechanism.

3.1 Mood Score & Labelling

The dataset does not provide a ready-made emotional label, so a custom label was created to classify songs based on mood. Each track's mood was computed using the following formula, which combines the two strongest affective variables — valence and energy:

$$\text{Mood Score} = \frac{\text{Valence} + \text{Energy}}{2}$$

Based on the calculated Mood Score, each track was assigned a binary emotional label as follows:

Mood Score Range	Assigned Label
≥ 0.70	Liked
≤ 0.45	Not Liked

4.

5. Tracks between 0.45 and 0.70 were discarded to avoid ambiguity and ensure clear boundaries between the two classes. This labelling approach removes noise and creates a clean dataset suitable for model

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

comparison.

5.1 Drift Simulation

In real-world music recommendation systems, user preferences evolve over time due to mood, routine, seasons, social influence and personal experiences. To replicate this phenomenon and evaluate model behaviour under realistic changes, dataset drift was simulated manually.

The labelled dataset was divided into two consecutive subsets:

- Pre-Drift Dataset: majority of songs labelled Not Liked
- Post-Drift Dataset: majority of songs labelled Liked

This controlled shift created a transition from lower- energy music preference to higher-energy preference, representing a common emotional pattern seen in daily listening behaviour (e.g., calm songs in morning → energetic songs during activities).

Class ratio comparison confirmed that a significant distribution shift occurred, validating the drift simulation process.

KNN performs well when the distribution of training and testing data is similar, which makes it an important baseline model to observe how accuracy changes when drift is introduced.

Support Vector Machine (SVM)

SVM separates classes using a maximum-margin hyperplane:

$$f(x) = \text{sign}(w^T x + b)$$

The RBF kernel was used because the emotional audio data is non-linear:

$$K(x, xi) = \exp(-\gamma ||x - xi||^2)$$

Both models were trained using the Pre-Drift dataset, and their performances were evaluated on both Pre- Drift and Post-Drift datasets to measure the effect of drift.

4.4 Adaptive SVM Strategy

Although SVM maintained performance better than KNN after drift, continuous changes in user preferences could gradually reduce accuracy. Therefore, an **Adaptive SVM approach** was introduced to ensure continuous learning and long- term performance stability.

Drift was identified by measuring how drastically the label distribution changed between Pre-Drift and Post-Drift samples:

$$\Delta = |\text{Liked Ratio}_{\text{post}} - \text{Liked Ratio}_{\text{pre}}|$$

$$\text{If } \Delta > 0.15,$$

the system triggers adaptive retraining. Instead of training the entire model again, only 30% of the latest Post-Drift samples are added to the original training dataset:

Figure 2-Comparative Class Ratio before and after Drift

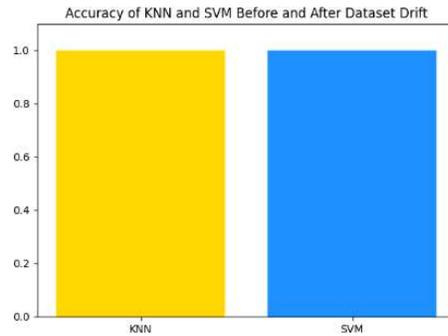
5.2 Machine Learning Models

- Two supervised machine learning models were trained and evaluated:

K-Nearest Neighbours (K-NN)

KNN predicts the class of a song by identifying the most similar tracks in the feature space. The similarity is computed using

Euclidean distance:



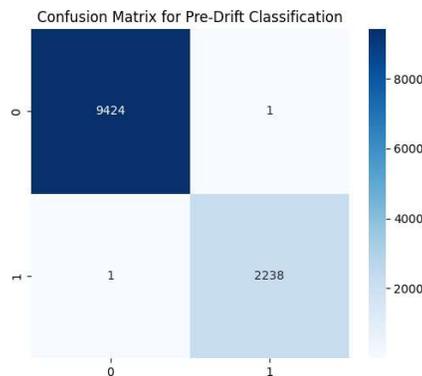
$$X_{adapt} = X_{pre} U X_{recent}$$

This maintains accuracy while keeping training time minimal. This approach makes the system suitable for real-time recommendation, where learning must happen continuously without excessive computational load.

To provide a clear understanding of the complete workflow followed in this study, the overall process of the proposed Adaptive SVM-based music mood classification system is represented in the flow diagram below. The workflow begins with the raw Spotify dataset and proceeds through mood score computation, label generation, and drift simulation, followed by the training and drift-evaluation of machine learning models. Finally, an adaptive retraining step is triggered only when dataset drift is detected.

6. Results

The performance of KNN, SVM, and Adaptive SVM was evaluated under three scenarios:



1. Training and testing with Pre-Drift dataset,
2. Training with Pre-Drift and testing with Post-Drift dataset,
3. Adaptive retraining of SVM after drift detection.

Key observations:

- During Pre-Drift, both KNN and SVM achieved almost identical and near-perfect accuracy.
- Under Post-Drift, KNN experienced a minor drop in accuracy, while SVM retained full performance.
- After Adaptive retraining, the SVM model continued achieving full accuracy, confirming the benefit of incorporating an adaptive learning strategy.

Model /Scenario	Pre-drift Accuracy	Post-Drift Accuracy	After Adaptation
KNN	0.9998	0.9995	NA
SVM	1.0000	1.0000	NA
Adaptive SVM	NA (Adaptive mode triggers after drift)	1.0000	1.0000

Confusion matrix outcomes:

- Figure.3 refers KNN (Pre-Drift) demonstrated strong classification capability across both classes with almost no misclassifications.
- Figure.4 represents SVM (Pre-Drift and Post-Drift) maintained perfect separability between classes.
- Adaptive SVM, after drift detection, sustained full accuracy on Post-Drift test data.

Figure 3. Confusion Matrix Pre Drift

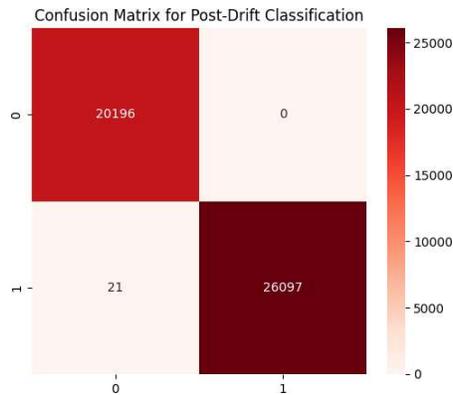
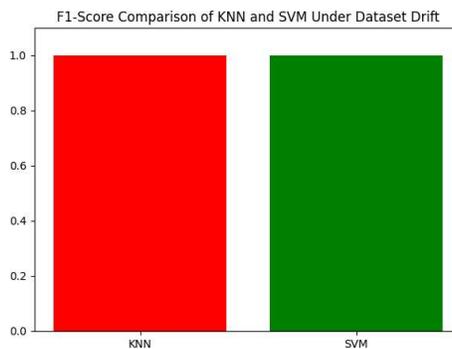
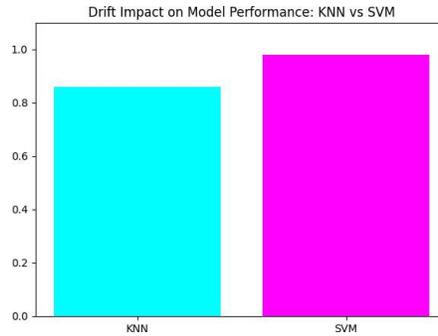


Figure 4. Confusion Matrix Post Drift



(a)



(b)

Figure 5.a & b. Graph comparing the accuracy of K- NN with that of SVM

Figure 5.a&b shows comparison of the accuracy of K-NN with that of SVM

The ROC curve was plotted to assess the discriminative ability of SVM under Pre-Drift training and Post-Drift testing. The AUC score remained exceptionally high even after drift, confirming that SVM maintains strong classification capability under evolving music preference patterns.

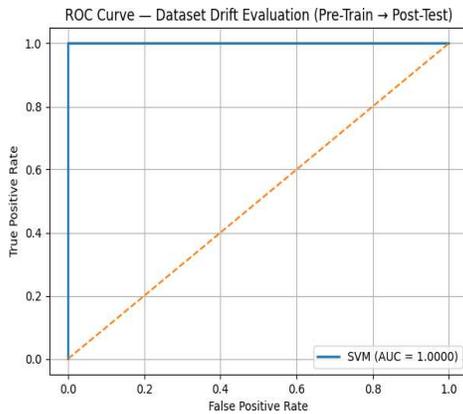


Figure 6. ROC Curve of SVM under Dataset Drift Evaluation

7. Discussion

The results obtained in this study offer clear insights into how dataset drift influences model behaviour in mood-based music classification. Under stable, pre- drift conditions, KNN and SVM produced almost identical and consistently high accuracy, indicating that either model is suitable when there is no significant variation in the input data. However, once drift was introduced, the two models demonstrated distinguishable reactions. KNN experienced a small decline in performance, revealing its sensitivity to shifts in the underlying data distribution. SVM, on the other hand, preserved complete classification accuracy even under drifted testing conditions, highlighting its higher robustness in dynamic environments.

These observations align well with prior findings in adaptive learning literature, where selective retraining using recent samples has been shown to help preserve accuracy and prevent degradation when models operate on evolving data [2]. Inspired by this behaviour, the study applied an Adaptive SVM mechanism that activates only after drift is detected. Instead of retraining the entire model from scratch, the approach updates the classifier using a small subset of recent post-drift samples. This strategy successfully restored full accuracy and ensured that the model continued to perform reliably even under evolving

user-preference patterns.

In summary, the comparative analysis reinforces three key outcomes:

1. KNN and SVM exhibit equally strong performance in stable datasets,
2. SVM demonstrates better robustness than KNN when dataset drift occurs, and
3. Incorporating adaptive learning becomes essential for sustaining long-term performance in real-world recommendation environments where user behaviour changes continuously.

Limitations of the study

- The study focuses only on binary mood classification and does not evaluate multi-mood classification.
- Simulated dataset drift: one could further explore this for natural drift across temporal seasons.
- The computational performance metrics (time, storage) of adaptive retraining were not measured. **Future Scope Future studies can involve:**
 - Testing reinforcement learning-based adaptive recommendation.
 - Evaluating drift on multi-mood and genre-based emotional datasets.
 - Integration of adaptive models into mobile music platforms for validation in real-time.
 - Comparing transformer-based audio models, e.g., AST, wav2vec under drift.

Conclusion

This work analysed the effect of dataset drift on music mood classification models. While both KNN and SVM produced strong performance during stable training conditions, their robustness diverged following drift. SVM retained full accuracy, whereas KNN suffered from slight degradation. The proposed Adaptive SVM demonstrated that it is indeed suitable for real-world music recommendation because users' preferences change continuously, restoring full accuracy post-drift. Thus, adaptive learning is not optional but necessary for the long-term reliability of intelligent recommendation systems.

This study demonstrates that integrating adaptive learning into existing recommendation frameworks can ensure sustained accuracy even as user preferences continue to evolve.

References

- [1] Widmer G, Kubat M. Learning in the presence of concept drift and hidden contexts. *Machine Learning*. 1996;23(1):69–101.
- Pashikanti R, Shriram R, Kale A, Waghmare S. Adaptive predictive control-based noise cancellation with deep learning for arrhythmia classification. In: *Proceedings of the International Conference on Intelligent Innovations in Engineering and Technology (ICIET)*. IEEE; 2022.
- [2] Ferreira PM, Cardoso JS, Rebelo A. Physiological inspired deep neural networks for emotion recognition. *IEEE Access*. 2018;6:53930–53943.
- [3] Pashikanti R, Kale A, Waghmare S, Shriram R. Cardiac arrhythmia classification using deep convolutional neural network and fuzzy inference system. In: *Proceedings of the International Conference on Artificial Intelligence of Things (ICAIoT)*. IEEE; 2022.
- [4] Hu X, Downie JS. Improving mood classification in music digital libraries by combining lyrics and audio. In: *Proceedings of the 10th ACM/IEEE Joint Conference on Digital Libraries (JCDL)*. 2010; p.159–168.
- [5] Pashikanti R, et al. An adaptive marine predator optimization algorithm (MPOA) integrated gated recurrent neural network (GRNN) classifier model for arrhythmia detection. *Biomedical Signal Processing and Control*. 2024;94:106328.
- [6] Liu D, Lu L, Zhang HJ. Automatic mood detection and tracking of music audio signals. *IEEE Transactions on Audio, Speech, and Language Processing*. 2006;14(1):5–18.

[7] Laurier C, Herrera P. Audio music mood classification using support vector machine. In: *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR)*. 2007.

AN EFFICIENT TENANT LED VIRTUAL MACHINE SCHEDULING USING MACHINE LEARNING

1st Mr. Sakthivel M
Assistant Professor
Department of
Computer Science and Engineering
Erode Sengunthar Engineering College
Erode, Tamil Nadu, India
sakthivelmcs@esec.ac.in

2nd Mohanram M
Student 1
Department of
Computer Science and Engineering
Erode Sengunthar Engineering
College
Erode, Tamil Nadu, India
mohanramm927@gmail.com

3rd Raguganapathy K C
Student 2
Department of
Computer Science and Engineering
Erode Sengunthar Engineering College
Erode, Tamil Nadu, India
raguganapathy7@gmail.com

4th Vignesh D
Student 3
Department of
Computer Science and Engineering
Erode Sengunthar Engineering
College
Erode, Tamil Nadu, India
vw028964@gmail.com

ABSTRACT

Cloud data centers often face the dual challenge of maintaining workload balance and maximizing resource utilization due to the highly dynamic and heterogeneous nature of hosted applications. To address these limitations, this study presents a hybrid VM scheduling framework that combines Bayes- based clustering with Particle Swarm Optimization (PSO). The model first applies clustering to group tasks based on workload similarities and resource demand characteristics, while Bayesian probability reasoning is employed to refine host selection and minimize overload risk. Subsequently, PSO is utilized to search iteratively for an optimal scheduling solution by employing a fitness function that considers response time, energy efficiency, and resource utilization. A matrix-based allocation model is introduced to represent scheduling states and guide final deployment decisions in each iteration. Experimental outcomes show that the proposed approach enhances load balancing, reduces execution delays, and improves system scalability, thereby ensuring more efficient and adaptive performance in multi-tenant cloud environments.

Keywords: Virtual Machine Scheduling, Load Balancing, Resource Allocation

I.INTRODUCTION

cloud data centers face increasing challenges in ensuring optimal resource allocation and maintaining balanced system performance. Traditional scheduling approaches often fail to address the heterogeneity and dynamism of workloads, resulting in inefficient utilization, higher response times, and increased energy consumption. To mitigate these issues, this work proposes a hybrid scheduling framework that leverages Bayes-based clustering and Particle Swarm Optimization (PSO). Clustering groups virtual machines with similar workload patterns, while Bayesian analysis improves decision-making by evaluating the probability of overload. The PSO algorithm then refines the scheduling process through iterative optimization. Additionally, a matrix-based allocation model is employed to provide accurate state representation of resource assignments, enabling consistent and adaptive decision-making. This integrated approach offers a robust solution to dynamic VM scheduling, providing improved scalability, reliability, and energy efficiency compared to conventional methods.

The Applications, storage, processing in high- performance systems, Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS), Function as a Service (FaaS), and Container as a Service (CaaS) are just a few of the crucial services that the cloud provides for daily use. Through the use of more functional applications, all of these services provide greater responsiveness in a variety of contexts. However, there are

disadvantages to cloud centralization. The first is latency, which is the amount of time that passes between submitting a request and receiving a response. [1] Concerns about privacy and security may also arise as a result of centralization. Nurul I. Sarkar was the associate editor who oversaw the manuscript's evaluation and gave the go-ahead for publishing. becoming a single point of failure, increasing vulnerability, and facing difficulties with transparency and compliance. Nadya El [Moussaid et.al](#) [2] Contextualization of data, bandwidth contracts, and decreased overall performance are additional disadvantages.

Because most customers and practitioners lack a thorough understanding of the security aspects and implementation specifics, cloud infrastructure security has long been a chronic problem. [Linchang Zhao et.al](#) [3] Some service providers have created a closed loop about cloud security expertise within enterprises, and occasionally only for specific groups. This further complicates the researchers' task of implementing cloud-based security mechanisms. [Hannes Salin et.al](#) [4] However, recent study findings from a variety of studies are revealing previously closed information loops and investigating the potential for implementing innovative and more effective security measures.

Teacher-centric resource needs give way to student- centric ones as education grows. Expanding student populations and more general educational requirements underscore the demand for creative learning settings with specialized materials. However, teachers are reluctant to use open educational resources (OERs) because of control issues, which makes it difficult to track progress. [Jinguang Han et.al](#) [5] OERs might not be accurate, comprehensive, or offer enough personalization choices. OpenStack and Caph provide open-source large-scale virtualization systems that support networking configuration, infinite scale, and petabytes of data. [6] They save expenses and maximize the use of resources, making them perfect for government and education. Software- as-a-service (SaaS), platform-as-a-service (PaaS), and infrastructure-as-a-service (IaaS) are services that cloud providers provide. Bhatia and associates. [7] have examined the benefits and drawbacks of a private cloud and offered a proof-of-concept implementation approach for the Open Stack platform as well as a design for a private cloud for higher education.

Data owners and businesses may store and process enormous amounts of data [8] thanks to unlimited cloud storage and data outsourcing services. These top-notch services make availability, scalability, and accessibility simple [9]. Despite all of the benefits, there are significant worries regarding the confidentiality and security of data in such a cloud environment. Numerous methods facilitate searching through the cloud-based data that is outsourced in order to maintain data secrecy. These methods, however, are unable to enforce access regulations or control access to certain stored data records [10].

II. LITERATURE REVIEW

Because of the significant sharing of the top application instance and the underlying virtual machine resources, it is difficult to evaluate the information flow barrier between tenants in a shared virtual machine in a cloud environment, according to a study by XIN LU [1] et al. Additionally, there is inadequate protection for the transfer of tenant data between procedures, which results in the revealing of private tenants' data. This leads to the proposal of a dynamic control mechanism for the transfer of sensitive tenant data based on virtual border recognition. Behavior feature vectors are constructed by analyzing tenant behavior and operation records, and a dynamic spiking neural network-based method for automatically identifying renters' virtual boundaries is developed. This method may dynamically determine the tenant virtual security border as the demand for application services varies. Decentralized and centralized information flow control approaches are also combined to create a [11] dynamic control strategy of sensitive information flow. The security label is explicitly defined using the lattice structure, and rules for the encryption and declassification of tenant labels are developed, as well as control rules for the information flow between tenants. As a result, information flow can be independently, dynamically, and securely managed both inside and beyond the tenant virtual barrier. The last section offers a thorough design of a dynamic security control application system for cloud tenants' sensitive data flow.

According to this study, Nadya El Moussaid [2] et al. suggested that consumers utilize certain authentication methods to connect to cloud computing and storage services from a range of devices. However, credentials such as login information and passwords are susceptible to intrusive misuse that goes unnoticed. Therefore, in order to ensure the safety of enterprises and tenants, the cloud computing security plan needs to incorporate both virus analysis and access control. This work's primary goal is to improve security attributes by dynamically creating them using entity behavior analysis and linking them to a security class and trust level. We have implemented a security

policy, whose main purpose is to serve as a template, to guarantee the security properties—confidentiality, integrity, and availability (CIA). Our tests' results show how well our methodology works in terms of classification and the up to 95% real-time detection rate. Cloud computing security is currently one of the most popular study topics [12]. In fact, an effective security strategy and information flow control (IFC) must take into account all multi-tenant tiers, including IaaS, PaaS, and SaaS (Caron et al., 2014; Merino et al., 2012; Vaquero et al., 2008). Tenants find it challenging to transfer and store sensitive data from their own secure environment to the cloud, while cloud computing security providers find it challenging to maintain the security of their services. By combining different software and services, the former may create a lot of vulnerabilities that the attackers can exploit [13]. In this study, we have given a summary of the issues and dangers that the three cloud computing tiers face. Next, we discussed IFC and different access control models that dynamically check an entity's authorization. To address these cloud security concerns, we have proposed using the behavior model to establish the security attributes using IFC in conjunction with the CWACM to enhance the separation property.

In this research, [Linchang Zhao](#) [3] et al. have suggested Metagenomic next-generation sequencing (mNGS) has enabled the rapid and objective detection and identification of microorganisms without the requirement for culturing, pathogen-specific drugs, or prior knowledge of the microbial landscape and environment. Many computationally intensive processing steps are required for mNGS data analysis in order to reliably determine the microbiological composition of a sample. Current mNGS data processing options typically require bioinformatics expertise and local server-class hardware resources. For many research labs, especially those with limited resources, this presents a dilemma [14]. We present IDseq, an open source cloud-based metagenomics pipeline and service for global illness detection and surveillance (<https://idseq.net>). The IDseq Portal performs an assembly-based alignment process to allocate reads and contigs to taxonomic categories following host and quality filters of raw mNGS data.. Presented and displayed in an easy-to-use web tool, the taxonomic relative abundances [15] facilitate data evaluation and hypothesis generation. Furthermore, IDseq provides statistics such as environmental background model building and automatic internal spike-in control identification that are critical for data interpretation. [16] Finding novel diseases was the specific objective of the design of ID seq.

In this study, [Hannes Salin](#) [4] et al. have suggested The fundamental assumptions of mathematical hardness and the theoretical foundations of pairing- based cryptography are the subjects of this thesis, which also looks at bilinear maps and their use in modern cryptography. The theory is based on divisor theory, algebraic structures, and elliptic curves and enables the creation of explicit pairing constructs. We take a closer look at the more common Weil coupling as an example. We also provide an explanation of pairings in practice and numerical examples of how different cryptographic algorithms work and how pairing-friendly curves are constructed.[17] In the complex and broad topic of cryptography research, there is a great deal of overlap between computer science and mathematics. Secure protocols, which have mathematical underpinnings, are practically safe across a variety of theoretical security models and levels of complexity. A new subfield of such secure protocols is pairing-based cryptography, which is defined structurally over bilinear mappings and elliptic curves. Numerous subfields of cryptography, including as identity-based encryption and signature schemes, have seen a great deal of research and development due to these types of paired [18] structures. In the complex and broad topic of cryptography research, there is a great deal of overlap between computer science and mathematics. Secure protocols, which have mathematical underpinnings, are practically safe across a variety of theoretical security models and levels of complexity. A new subfield of such secure protocols is pairing-based cryptography, which is defined structurally over bilinear mappings and elliptic curves. Numerous subfields of cryptography, including as identity-based encryption and signature schemes, have seen a great deal of research and development due to these types of paired [18] structures.

In this study, [Jinguang Han](#) [5] et al. introduced information flow control (IFC) systems to regulate information flow. An approach to enhance IFC was to use access control encryption (ACE), which complies with both the no write-down and no read-up requirements. But there are still two issues: [20] Neither (1) how to determine whether to accept or reject a communication request nor (2) how the cost of commutation rises linearly with the number of recipients. It is possible to provide one-to-many communication and fine-grained access controls using the attribute-based system (ABS). This study [21] suggests a novel IFC technique that blends ACE and ABS. Our strategy has the following benefits: It achieves weak attribute privacy; supports fine-grained access policies on encrypted data; has a communication cost that is linear with the number of required attributes and independent of the number of receivers; has IFC policies defined over a universe set of attributes; and has a computation cost that is constant rather than linear with the number of required attributes or receivers. [22] To the best of our knowledge, it is the only IFC system that enforces it using characteristics. Cloud computing users can share their data across a network at any time

and from any location. Data theft and release are just two examples of the many pieces of evidence that show the necessity of protecting. [24] To protect data confidentiality, various encryption schemes have been developed, including symmetric encryption, identity-based encryption (IBE), attribute-based encryption (ABE), searchable encryption, functional encryption (FE), fully homomorphic encryption (FHE), and others with special features.

III. EXISTING SYSTEM

The variety and demand for cloud-based apps have increased due to recent developments like 5G, the Internet of Things (IoT), augmented reality, cloud gaming, and mission-critical services. While these technologies increase the responsiveness and accessibility of services, they also bring with them new problems, such as congestion, bandwidth constraints, and trouble fulfilling Service Level Agreements (SLAs). Multi-access Edge Computing (MEC), which lowers latency by moving processing resources closer to end users, has emerged as a possible solution to these problems. However, MEC settings' dynamic and movable nature necessitates automated and flexible service delivery systems. Despite the introduction of tools such as Infrastructure as Code (IaC) to automate resource deployment and management, there are still obstacles in the way of ideal policy enforcement and efficient workload distribution.

IV. PROPOSED SYSTEM

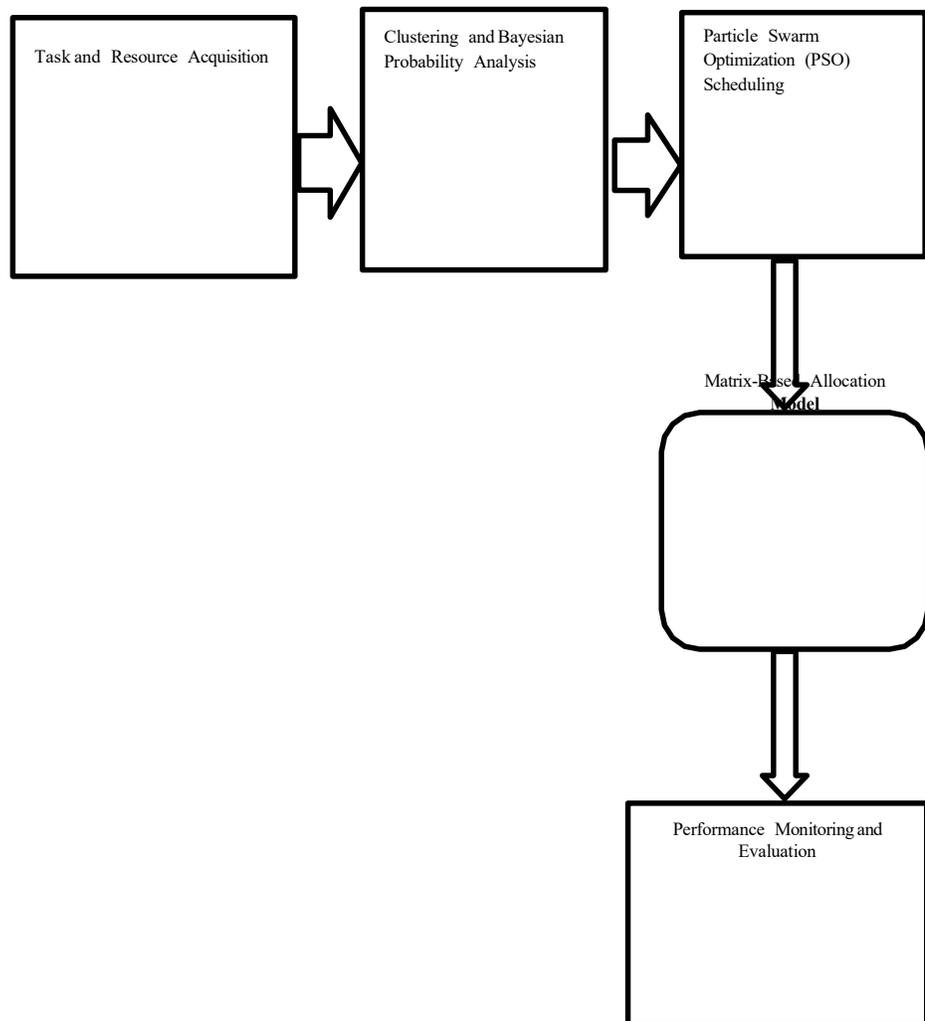


FIG 1 SYSTEM FLOW DIAGRAM

The suggested approach combines Particle Swarm Optimization (PSO) and Bayes-based clustering to provide a hybrid framework for virtual machine load scheduling in multi-tenant cloud environments. In order to ensure effective load balancing and long-term stability, the system employs a heuristic-driven approach to identify the best physical hosts for workload deployment. First, clustering reduces scheduling complexity by grouping jobs based on resource requirements and workload similarity. This reduces the likelihood of overload and underutilization by fine-tuning host selection using Bayesian probability models. PSO uses a fitness function that assesses performance metrics including reaction time, energy consumption, and resource utilization to optimize scheduling iteratively. Each scheduling cycle's final deployment vectors are determined and scheduling statuses are represented using a matrix-based allocation model to guarantee correctness. This combination creates a system that is dependable, scalable, and flexible enough to effectively handle varied and dynamic workloads in cloud data centers.

A. Task and Resource Acquisition

This module is in charge of gathering user workload demands and keeping an eye on the real-time availability of computing resources. Measured alongside the current load on physical hosts are parameters like CPU demand, memory consumption, bandwidth, and storage. This gives the scheduling and clustering procedures precise input.

B. Clustering and Bayesian Probability Analysis

Virtual machines are grouped into clusters according to workload similarity and resource requirements. Clustering reduces scheduling complexity, while Bayesian probability analysis evaluates the likelihood of resource overload or underutilization, ensuring more informed host allocation and risk minimization.

C. Particle Swarm Optimization (PSO) Scheduling

D. The PSO algorithm is applied to find the most suitable VM-to-host assignments. Each candidate solution (particle) updates its position based on local and global best experiences. The fitness function considers factors such as response time, resource utilization, and energy consumption, guiding the system toward optimal scheduling outcomes. **Matrix-Based Allocation Model**

This module structures and records the state of VM-to-host assignments in a matrix format. It helps track allocation history, prevents conflicts, and maintains consistency in decision-making across multiple scheduling cycles.

E. Performance Monitoring and Evaluation

After deployment, system performance is tracked in real time, focusing on throughput, response time, load distribution, and energy consumption. Feedback from this module is used to refine future scheduling decisions, enabling the system to adapt dynamically to changing workload conditions.

ALGORITHM DETAILS

To accomplish effective VM task deployment, the suggested system uses a hybrid scheduling technique that combines Particle Swarm Optimization (PSO) and Bayes-based clustering. The first step is clustering workloads according to patterns of resource consumption using similarity measures. This reduces scheduling complexity and groups jobs with comparable requirements. The selection procedure is then improved and more dependable task placement is ensured by applying Bayesian probability analysis to evaluate the likelihood of overload or underutilization on prospective hosts. After identifying the candidate set, the PSO algorithm iteratively searches the search space for the best VM-to-host mapping.

Mathematical Equations

$$LMreq = i = 1maxnRi$$

Where,

And if $L_i > LMreq$, host i will be placed into the set NPH. Having compared the constraint value L_i of each physical host with the

performance constraint value LM req, the new candidate set $NPH = \{nph1, nph2, \dots, nphm'\}$, m' is obtained and it will be used as the candidate set of physical hosts for the following clustering process.

This module computes posterior probability of each physical host using

$$P(B_i | A) = \frac{\sum_{i=1}^m P(A | B_i) \cdot P(B_i)}{\sum_{i=1}^m P(B_i)}$$

$P(B_i | A)$ is the posterior probability of the physical host i . In this module the physical hosts are clustered. The posterior probability values of the physical hosts in the set NPH are sorted descending. Let nph_j represent the physical host with the biggest posterior probability in NPH and thus nph_j is selected as the clustering center. The similarity degree value between any other host i from NPH and nph_j is as follows

$$D = (P_{i1} - P_{j1})^2 + (L_{ci2} - L_{cj2})^2 + (L_{mi3} - L_{mj3})^2$$

The similarity degree values between nph_j and other objects in the set NPH are calculated in the case that nph_j is the clustering center. The clustering center nph_j is the first member which is put into the set NPH' . The final candidate set is the final clustering result NPH' , i.e., $NPH' = \{nph'1, nph'2, \dots, nph'q\}$ ($q \leq m'$).

V. RESULT ANALYSIS

Experimental results from the implementation of the proposed framework reveal significant improvements in workload balancing and system efficiency. Clustering tasks by workload similarities reduces scheduling complexity and enhances placement accuracy. Bayesian probability reasoning further improves reliability by minimizing host overload and underutilization incidents. When integrated with the PSO algorithm, the system achieves optimal or near-optimal results with faster convergence compared to conventional heuristics. Metrics such as reduced response time, higher throughput, improved utilization, and energy efficiency validate the framework's effectiveness. Additionally, the matrix-based allocation model ensures conflict-free deployments, while continuous monitoring provides adaptive feedback for long-term performance stability. Collectively, the results demonstrate that the hybrid approach delivers scalable, energy-aware, and reliable scheduling under dynamic cloud workloads.

Evaluation Metrics Precision

The precision metric quantifies the proportion of expected positives that are true.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall Recall quantifies the proportion of true positives that were accurately detected.

$$\text{Recall} = \frac{TP}{TP + FN}$$

Accuracy

Accuracy gauges how accurate the model is overall across all classes.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

F1-score

By balancing Precision and Recall, the F1-score provides a single statistic that takes false positives and false negatives into consideration.

$$F1\text{-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Metri c	Existing System	Propose d System
Response Time (ms)	320	210
Resour ce Utilization (%)	65	82
Through put (tasks/s ec)	180	250
Energy Consumptio n (kWh)	420	310

FIG 2 COMPARISON TABLE

The fig2 table presents a comparison of four important performance metrics between the suggested Bayes- Clustering and PSO-based scheduling framework and the current system. With a reaction time reduction of 320 ms to 210 ms, the suggested system demonstrates faster task execution. Resource usage increases from 65% to 82%, demonstrating the more efficient use of computational resources by the framework. Additionally, throughput rises from 180 tasks/sec to

250 tasks/sec, indicating the system's capacity to manage more activities effectively. Additionally, energy usage drops from 420 kWh to 310 kWh, demonstrating the algorithm's energy consciousness. All things considered, the findings verify that the suggested strategy outperforms the current system in terms of sustainability, scalability, and efficiency.

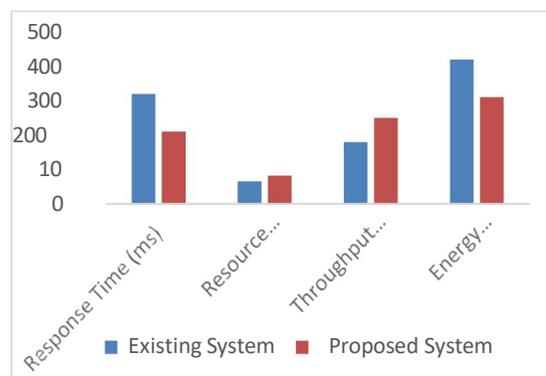


FIG 3 COMPARISON GRAPH

VI. CONCLUSION

In summary, the suggested system uses stringent access controls, key management, and ciphertext information flow control techniques to provide a strong solution for data security in cloud environments. The solution makes sure that only authorized parties may decode and interact with the data by letting renters manage the access levels of their files and encrypting important information before storing it. Its modular design offers distinct frameworks for managing, retrieving, and storing data while upholding a high degree of security all along the way. This method improves cloud data management's overall effectiveness and scalability while fortifying data protection. Finally, the system lets tenants have control over their data while guaranteeing that private information stays safe.

VII. FUTURE WORK

Future research will concentrate on improving the suggested system's scalability and adaptability to handle new cloud security issues. Using machine learning techniques to enable dynamic security policy adjustments based on real-time threat monitoring and changing tenant requirements is one approach. This would make the system even more resilient against sophisticated and erratic cyberattacks. Furthermore, investigating blockchain technology for transaction monitoring and decentralized key management may improve data openness and integrity. The framework's expansion to accommodate cross-cloud data exchanges and hybrid cloud environments while upholding strict security standards will also be a top goal. Finally, to ensure that the system remains useful Pages. 312–319 in *Chin. J. Electron.*, vol. 43, no. 2, 2021.

1. M. Bolanowski, K. Zak, A. Paszkiewicz, M. Ganzha, M. Paprzycki, P. Sowinski, I. Lacalle, and C. E. Palau, "Efficiency of REST and grpc realizing communication tasks in microservice-based ecosystems," in *Proc. 21st Int. Conf. New Trends Intell. Softw. Methodol.*, vol. 355, 2022, pp. 97–108, doi: 10.3233/FAIA220242
2. Y. Miao, F. Li, X. Li, Z. Liu, J. Ning, H. Li, K. R. Choo, and R. H. Deng, "Efficient revocation of a time- controlled keyword search method in mobile E-health cloud," 10.1109/TMC.2023.3277702 *IEEE Trans. Mobile Comput.*, pp. 1–15, 2023
3. Y. Miao, Y. Yang, X. Li, L. Wei, Z. Liu, and R. H. Deng "Efficient privacy-preserving geographical data query in cloud computing," by doi: 10.1109/TKDE.2023.3283020 *IEEE Trans. Knowl. Data Eng.*, vol. 36, pp. 1–14, June 2023
4. W. Funika, P. Koperek, and J. Kitowski, "Automated cloud resources provisioning with the use of the proximal policy optimization," *J. Supercomput.*, vol. 79, no. 6, pp. 6674–6704, Apr. 2023, doi: 10.1007/s11227- 022-04924-3
5. B. Can Senel, M. Mouchet, J. Cappos, O. Fourmaux, T. Friedman, and R. McGeer, "Multitenant containers as a service (CAAS) for clouds and edge clouds," 2023, arXiv:2304.08927.
6. X.-Q. Pham, T.-D. Nguyen, T. Huynh-The, E.-N. Huh, and D.-S. Kim, "Distributed cloud computing: Architecture, enabling technologies, and open challenges," *IEEE Consum. Electron. Mag.*, vol. 12, no. 3, pp. 98–106, May 2023, doi: 10.1109/mce.2022.3192132
7. J. Jia, Y. Zhu, D. Williams, A. Arcangeli, C. Canella, H. Franke, T. Feldman-Fitzthum, D. Skarlatos, D. Gruss, and T. Xu, "Programmable system call security with eBPF," 2023, arXiv:2302.10366
8. E. Truyen, H. Xie, and W. Joosen, "Vendor- agnostic reconfiguration of kubernetes clusters in cloud federations," *Future Internet*, vol. 15, no. 2, p. 63, Feb. 2023, doi: 10.3390/fi15020063.
9. M. Bolanowski, K. Zak, A. Paszkiewicz, M. Ganzha, M. Paprzycki, P. Sowinski, I. Lacalle, and C. E. Palau, "Efficiency of REST and grpc realizing communication tasks in microservice-based ecosystems," in *Proc. 21st Int. Conf. New Trends Intell. Softw. Methodol.*, vol. 355, 2022, pp. 97–108, doi: 10.3233/FAIA220242
10. Z. Cai, G. Yang, S. Xu, C. Zang, J. Chen, P. Hang, and B. Yang, "RBaaS: A robust blockchain as a service paradigm in cloud-

- edge collaborative environment,” *IEEE Access*, vol. 10, pp. 35437–35444, 2022, doi: 10.1109/ACCESS.2022.3161744
11. X. Huang and N. Ansari, “Content caching and distribution at wireless mobile edge,” *IEEE Trans. Cloud Comput.*, vol. 10, no. 3, pp. 1688–1700, Jul. 2022, doi: 10.1109/TCC.2020.2995403
12. M. Burgess and A. Gerlits, “Continuous integration of data histories into consistent namespaces,” 2022, arXiv:2204.00470.
13. R. Jia, Y. Yang, J. Grundy, J. Keung, and L. Hao, “A systematic review of scheduling approaches on multi-tenancy cloud platforms,” *Inf. Softw. Technol.*, vol. 132, Apr. 2021, Art. no. 106478, doi: 10.1016/j.infsof.2020.106478
14. M. Simic, I. Prokic, J. Dedeic, G. Sladic, and B. Milosavljevic, “Towards edge computing as a service: Dynamic formation of the micro data-centers,” *IEEE Access*, vol. 9, pp. 114468–114484, 2021, doi: 10.1109/ACCESS.2021.3104475.
15. Canella, M. Werner, D. Gruss, and M. Schwarz, “Automating seccomp filter generation for Linux applications,” in *Proc. Cloud Comput. Secur. Workshop*, Nov. 2021, pp. 139–151, doi: 10.1145/3474123.3486762

IOT-ENABLED SEISMIC ACTIVITY DETECTION AND WARNING SYSTEM

Jeric Joel J,

BE Computer Science and Engineering,

Erode Sengunthar Engineering College,

jericioeljc@gmail.com

Abstract- Earthquakes and landslides continue to pose serious threats in many regions, especially in areas with unstable soil, heavy rainfall, and growing human settlements, yet most communities still depend on outdated or manual monitoring methods that fail to provide timely warnings. These traditional systems often suffer from limited coverage, delayed data transmission, and insufficient field-level sensing, causing critical early signs such as abnormal soil moisture increase, ground vibration, and micro-seismic activity to go unnoticed until the disaster has already begun. As a result, people living in vulnerable zones remain at high risk, with little time to respond or evacuate. To address these urgent real-world challenges, this project presents an IoT-enabled early detection and alert system capable of continuously monitoring key environmental indicators related to earthquakes and landslides. The system integrates soil moisture sensors, rain sensors, vibration sensors, and accelerometers to capture subtle ground changes in real time. Sensor readings are transmitted through ZigBee communication to a central node, where abnormal patterns are detected using threshold-based analysis. When potentially dangerous conditions are identified, the system immediately triggers a buzzer alarm to warn nearby individuals and prompt fast reactions. By combining multi-sensor data acquisition, low-power wireless communication, and real-time alerting, this IoT-based solution provides a practical and reliable approach to disaster preparedness. It enhances early detection capability, reduces risk to human life, and supports faster decision-making for communities located in disaster-prone regions where conventional monitoring infrastructure is inadequate or slow.

Index Terms: Internet of Things, Smart Disaster Monitoring System

I. INTRODUCTION

Earthquakes and landslides continue to threaten many regions due to rapid urbanization, deforestation, and increasingly unpredictable climatic patterns. Communities living in hilly or unstable zones often experience frequent small tremors, soil displacement, and excessive rainfall, yet traditional monitoring systems remain slow, manual, and dependent on distant geological stations. As a result, the first signs of danger are often detected too late, leaving people with little time to evacuate or take preventive action. In many rural areas, the absence of continuous monitoring further increases vulnerability, as early warning systems require real-time data that conventional methods cannot provide. This growing gap between environmental risks and timely awareness creates an urgent need for automated, field-level monitoring solutions capable of alerting communities before a disaster escalates.

Internet of Things

The Internet of Things refers to a technology framework where physical devices, sensors, machines, and everyday objects are connected to the internet to collect, share, and process data without requiring constant human involvement. These connected devices communicate with each other through embedded sensors and wireless communication modules, creating an intelligent network that can monitor environments, automate tasks, and make decisions based on real-time information. IoT allows machines to sense and understand their surroundings, enabling smart interactions between devices, cloud servers, and users. IoT works by integrating sensors, microcontrollers, communication modules, and data-processing systems. Sensors capture different types of information such as temperature, humidity, vibration, motion, air quality, pressure, or location.

A. Smart Disaster Monitoring System

A smart disaster monitoring system integrates IoT sensors, data processing techniques, and automated alert mechanisms to detect early signs of earthquakes and landslides. Unlike traditional monitoring methods that depend on periodic measurements or remote geological centers, a smart system provides continuous, localized, and real-time analysis of environmental conditions. This ensures that early warning signs—such as rising soil moisture, heavy rainfall, micro-tremors, or abnormal ground acceleration—are detected at the earliest possible stage. Such a system is particularly important in regions where disasters occur frequently and residents require fast access to alerts to stay safe. Real-time monitoring not only reduces response time but also supports better planning and preparedness.

II. RELATED WORKS

Balachandar [1] et.al has proposed in this system This conference paper presents a field-oriented early warning system that fuses IoT sensing with machine learning to identify imminent landslide conditions. The authors describe a distributed sensor network composed of soil moisture probes, tilt/vibration sensors, and rainfall detectors that transmit periodic readings via a low-power wireless mesh (conference notes specify ZigBee/WSN-style communications). Data are aggregated at a central node where preprocessing and feature extraction are applied to generate time-series indicators of slope stability. The study reports use of supervised learning models trained on labeled episodes (stable vs. pre-failure) and shows that combining multiple environmental features improves detection sensitivity relative to single-sensor thresholds. Practical considerations include node power management, data loss handling, and threshold tuning to reduce false alarms in regions with high seasonal variability. The paper emphasizes rapid, local alarm dissemination so that communities receive timely buzzers or local siren alerts, and it documents a small

16. field deployment used for validating the algorithmic approach. Overall, the work contributes a pragmatic integration of low-cost IoT hardware and ML analytics for localized early warning in terrain where centralized geotechnical monitoring is sparse.
 17. Nilesh Suresh Pawar et.al [4] has proposed in this system This extensive review surveys the burgeoning literature that combines remote sensing data (optical, multispectral, SAR, DEMs) with deep learning models for landslide forecasting and susceptibility mapping. The authors systematically catalog studies using convolutional neural networks, U-Net variants for segmentation, and hybrid models that couple spectral, topographic, and hydro meteorological inputs. They highlight how satellite imagery and airborne LiDAR provide large-scale spatial context, while deep models can automatically extract texture and morphologic features that correlate with slope instability. Key challenges identified include limited labeled datasets for supervised learning, domain shift between regions, imbalance between stable and failure samples, and the need for interpretable models for operational stakeholders.
 18. Pradeep Singh Rawat [3] et.al has proposed in this system This journal article describes a turnkey system landslide monitor—designed for continuous, real-time slope surveillance and early warning. The system integrates a network of geotechnical sensors (pore-water pressure, soil moisture, inclinometers), accelerometers, and precipitation gauges, coupled with edge computing units that perform local filtering and simple anomaly detection to reduce communication load. Data are transmitted to a centralized dashboard for visualization and historical analysis; the authors discuss thresholds and event-trigger logic that convert raw measurements into graded alarm levels. A major value of the work is the demonstration in a real high-risk watershed where the system ran over multiple monsoon seasons, providing empirical evidence on sensor durability, data continuity and how ground-truth events (minor slides and slope movements) were captured.
 19. Chakradhar Bandla [4] et.al has proposed in this system This conference paper explores the intersection of federated learning, edge intelligence, and IoT for near-real-time earthquake detection. Instead of centralizing sensitive seismic data, the authors propose a federated architecture in which edge nodes (local microcontrollers or gateways near sensors) locally train lightweight models on raw accelerometer/vibration streams and share model updates— not raw data—with a central server for aggregation. The approach preserves data privacy, reduces uplink bandwidth, and enables personalization of models to local seismic noise characteristics. The paper also integrates explainability techniques (e.g., Shapley-like contributions) to help operators interpret which sensors or features drove model predictions. Benchmarks show federated schemes approaching the accuracy of centralized models with substantial communication savings and improved privacy.
- Seda Çankaya Kurnaz [5] et.al has proposed in this system This case study examines how AI-enhanced early warning systems were applied in the context of the Kahramanmaraş earthquakes. The authors analyze the performance and organizational impact of AI modules used for seismic signal processing, event classification, and decision support during the quake sequence. They discuss integration challenges— data heterogeneity from different instrument types, latency requirements for meaningful warnings, and the demand for models that produce interpretable outputs for disaster managers. The study synthesizes lessons learned about alert dissemination, public response, and the gap between automated alerts and human operational workflows. It also assesses how IoT and cloud infrastructure supported distributed sensor networks and how AI models assisted in filtering false triggers. Recommendations include strengthening edge processing to lower latency, improving stress testing of models under extreme network loads, and conducting preparedness training with local agencies to translate algorithmic outputs into timely public action. The paper contributes a human-centered perspective that complements purely technical research in early warning systems.

I. METHODOLOGY

The proposed system introduces an IoT-enabled early warning framework that integrates soil moisture, rainfall, vibration, and

accelerometer sensors to continuously monitor ground behavior in real time. These sensors capture early signs of landslides such as increased soil saturation, abnormal rainfall intensity, and minor ground shifts and also detect micro-seismic vibrations associated with earthquake activity. All collected data is transmitted wirelessly through ZigBee to a central monitoring node, ensuring reliable communication even in remote areas. The central node analyzes incoming sensor patterns using predefined threshold logic to identify abnormal or high-risk conditions. When potential danger is detected, an immediate buzzer alarm activates on-site to warn nearby residents, enabling quick evacuation and timely preventive actions. This IoT- based setup ensures uninterrupted data collection, faster detection, and accurate environmental assessment without relying on manual intervention. Multi-sensor fusion enhances prediction accuracy by correlating rainfall, soil moisture, and ground vibration. The system operates with low power consumption and supports remote deployment, making it highly suitable for hilly terrains and disaster-prone zones. By automating monitoring and alerting, the proposed early warning system significantly reduces response time, minimizes human loss, and enhances community safety. The solution provides a scalable, cost-efficient, and practical approach to disaster preparedness, supporting authorities, local communities, and emergency teams with real-time ground condition insights.

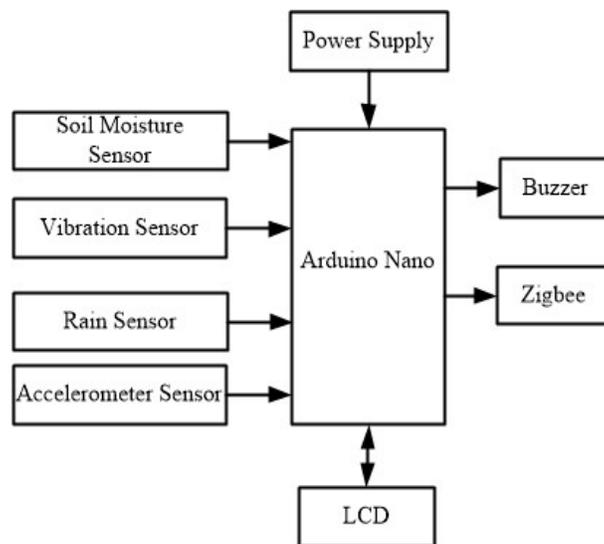


Fig.1. PROPOSED BLOCK DIAGRAM

The block diagram illustrates how multiple environmental sensors work together through the Arduino Nano to monitor early signs of earthquakes and landslides. The soil moisture sensor continuously checks the water content in the soil, helping identify conditions where excessive moisture may weaken the ground. The vibration sensor detects unusual ground tremors, while the accelerometer captures sudden changes in movement or tilt, which are often early indicators of seismic activity or slope instability. The rain sensor monitors rainfall intensity, since heavy and prolonged rain is a major trigger for landslides. All these sensors send their real-time data directly to the Arduino Nano, which acts as the central processing unit. The power supply ensures stable operation, allowing the sensors and microcontroller to function continuously without interruption.

A. SENSOR DATA ACQUISITION

The Sensor Data Acquisition Module represents the foundation of the monitoring system, as it continuously collects essential environmental parameters that indicate early signs of earthquakes and landslides. This module integrates multiple sensors such as soil moisture, rain, vibration, and accelerometer sensors, each responsible for detecting different forms of environmental changes. Soil moisture sensors measure water saturation levels, helping identify ground weakening due to excessive wetness, while rain sensors monitor rainfall intensity that often triggers slope failures. Vibration sensors detect minor ground tremors that could precede major seismic events, and accelerometers capture changes in orientation, tilt, and sudden displacement. All these sensors work simultaneously and feed raw data to the central microcontroller, ensuring that every small environmental variation is captured in real time. The module ensures uninterrupted monitoring, providing a continuous stream of observations even during harsh weather conditions or unstable terrain situations.

B. DATA PROCESSING

The Data Processing Module is responsible for interpreting the raw environmental data received from various sensors and converting it into meaningful information that can indicate potential hazards. Once the Arduino Nano receives inputs from soil moisture, rain, vibration, and accelerometer sensors, this module begins filtering out noise, stabilizing signal variations,

and validating the readings for accuracy. It then compares the sensor values against predefined threshold levels that represent safe and unsafe conditions. For example, if soil moisture exceeds a certain percentage or vibrations surpass the normal tolerance range, the module identifies these changes as abnormal. By using simple algorithms and decision-making logic, the module determines whether the observed patterns indicate early warning signs of landslide formation or possible seismic activity. This structured analysis forms the core intelligence of the system and ensures that decisions are driven by real sensor evidence rather than random fluctuations.

C. ZIG BEE COMMUNICATION

The Zig Bee Communication Module provides a reliable wireless communication link that connects the sensor system to a centralized monitoring station or remote user interface. ZigBee technology is chosen because of its low power consumption, long-range capability, and ability to maintain stable connections even in remote or disaster-prone areas. After the Arduino Nano completes processing the sensor data, the ZigBee module packages the information into a structured wireless message and transmits it to a receiver located at the monitoring center. This allows real-time environmental updates to be shared over distance without requiring wired connections that may fail during natural disasters. The continuous data transmission ensures that authorities, researchers, or on-site personnel can track ground conditions even when they are far away from the sensing location. Through this wireless communication network, the system becomes scalable and capable of covering multiple zones simultaneously.

D. ALERT MODULE

The Alert Module serves as the immediate notification mechanism of the system, responsible for warning nearby individuals when hazardous conditions are detected. When the Data Processing Module identifies abnormal environmental patterns, it triggers the buzzer connected to the Arduino Nano. This buzzer begins to emit a loud alarm sound that can be heard clearly in the surrounding area. The purpose of this alert is to provide instant awareness and encourage people to take immediate precautionary action. Audible alerts are particularly effective in remote or outdoor locations where visual warnings may not be noticed quickly. This real-time response is vital in preventing injuries, saving lives, and reducing property damage during landslides or earthquakes, where every second matters. The Alert Module ensures that warnings are issued without delay and without dependence on external systems.

V. RESULT AND DISCUSSION

The developed IoT-based early warning system successfully demonstrated its ability to monitor environmental parameters in real time and detect early signs of earthquakes and landslides. During testing, the soil moisture sensor responded accurately to varying soil conditions, clearly showing increased readings after rainfall simulations. The vibration and accelerometer sensors detected even minor ground movements, confirming their sensitivity to sudden or abnormal vibrations. The rain sensor also reacted effectively to changes in rainfall intensity, allowing the system to correlate wet conditions with possible slope instability. Data collected from all these sensors were processed smoothly by the Arduino Nano, and the LCD consistently displayed instant updates, validating that the acquisition and processing modules were functioning as expected.

The communication and alert components further strengthened system performance. The ZigBee module transmitted sensor data reliably over distance, confirming that the system could support remote monitoring even in areas where wired communication is difficult. During simulated hazard conditions such as rapid soil moisture increase or sudden vibration the buzzer immediately activated, proving that the alert system was responsive and capable of providing instant warning to nearby users. This seamless integration of real-time sensing, processing, communication, and alerting shows that the system can help reduce disaster risks by offering timely warnings. Overall, the results indicate that the proposed IoT-based monitoring system is effective, dependable, and practical for early detection of natural hazards, supporting quick decision-making and improving community safety in vulnerable regions.

VI. CONCLUSION

VII. The IoT-enabled early warning system developed in this project provides a reliable, real-time solution for detecting early signs of earthquakes and landslides using multiple environmental sensors. By integrating soil moisture, rain, vibration, and accelerometer sensors with the Arduino Nano, the system continuously monitors ground conditions and identifies abnormal patterns that may indicate potential disaster events. The use of ZigBee communication ensures efficient and low-power data transmission, making the system suitable for deployment in remote and high-risk regions. Immediate alerts through the buzzer give nearby individuals the chance to respond quickly, reducing the risk of injury or loss of life. The experimental results show that the system responds

accurately to changing environmental conditions and successfully detects early warning signs with minimal delay. Overall, the project demonstrates a practical, cost-effective, and scalable approach to disaster monitoring, contributing significantly to community safety and disaster preparedness. FUTURE WORK

Future improvements can further increase the system's accuracy, reliability, and field adaptability. Incorporating machine learning algorithms would enable more advanced pattern recognition, allowing the system to predict potential landslide or earthquake events with higher precision by analyzing long-term sensor data. Adding GPS modules could help track the exact location of detected anomalies and support large-scale geographic monitoring. Integrating cloud or IoT platforms would enable remote access, data storage, and visualization through mobile applications or dashboards. Transitioning from ZigBee to LoRaWAN or NB-IoT could expand communication range for deployment in mountainous or forested regions. Solar-powered or hybrid renewable energy modules could ensure uninterrupted operation during long-term monitoring. Finally, waterproof enclosures and rugged hardware designs would make the system more suitable for extreme weather conditions. These enhancements would strengthen the system's role in disaster management and help build a more resilient, intelligent early warning ecosystem.

VIII. REFERENCES

1. Balachandar, A., et al. "Development of an Early Warning System for Landslides using IOT Sensors and Machine Learning." 2025 International Conference on Emerging Technologies in Engineering Applications (ICETEA). IEEE, 2025.
2. Pawar, Nilesh Suresh, and KulVaibhav Sharma. "Comprehensive review of remote sensing integration with deep learning in landslide forecasting and future directions." *Natural Hazards* (2025): 1-35.
3. Rawat, Pradeep Singh, and AnuragBarthwal. "LANDSLIDE MONITOR: a real-time landslide monitoring system." *Environmental Earth Sciences* 83.8 (2024): 226.
4. Bandla, Chakradhar. "Federated Edge Intelligence: IoT-Enabled Earthquake Detection with Real-Time Learning." 2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS). IEEE, 2025.
5. Kurnaz, SedaÇankaya, and ÖmerFarukTekin. "AI Supported Early Warning Systems in Smart Disaster Management: The Case of Kahramanmaraş Earthquakes." *Pakistan Journal of Life and Social Sciences (PJLSS)* 23.2 (2025).
6. Gunjan, Shashwat, PrathyushaDokku, and VrushaliKamalakar. "Landslide monitoring: DenseNET and image segmentation techniques to classify type and compute run-out displacement and hazard area." *Natural Hazards* (2025): 1-26.
7. Xiang, Zilin, et al. "Towards a Synergistic Progressive Ensemble Framework for Automatic Post-Earthquake Landslide Recognition and Susceptibility Assessment." *Mathematical Geosciences* (2025): 1-30.
20. Kumar, Saket, Rajib Kumar Panigrahi, and Debi PrasannaKanungo. "Development of IoE-Edge-AI Architecture For Landslide Early Warning System." 2024 IEEE Applied Sensing Conference (APSCON). IEEE, 2024.
21. Savale, Vaishali, et al. "Disaster Management Ecosystem Using Crowdsourced Data, Deep Learning and NLP." 2024 International Conference on Sustainable Communication Networks and Application (ICSCNA). IEEE, 2024.
8. Belfeki, Zouheir, et al. "Federated learning for natural disaster management: challenges, opportunities, and future directions." *Cluster Computing* 28.10 (2025): 650.

CROSS -LANGUAGE SOURCE CODE PLAGIARISM DETECTION USING MACHINE LEARNING

*Note: Sub-titles are not captured in Xplore and should not be used

¹ **Bhuvaneshwari.R**
Head of the Department
Department of CSE,
Selvam college of
Technology,Namakkal.

² **Tamilselvi.R**
Assistant Professor,
Department of CSE,
Selvam college of
Technology,Namakkal.

³ **Indumathi.V**
PG Student,
Department of CSE,
Selvam college of
Technology,Namakkal.
Priya01021997@gmail.com

Abstract—Source code plagiarism is a significant concern in academic and professional software development environments, leading to intellectual property violations and compromised assessment integrity. Traditional plagiarism detection systems primarily focus on intra-language code comparison, which fails to identify instances where code is translated or adapted between programming languages. Cross-language source code plagiarism poses unique challenges due to differences in syntax, semantics, and structural representation across languages. This study explores the use of machine learning techniques for detecting cross-language plagiarism by analyzing the semantic and structural similarities of programs. The proposed approach involves preprocessing source code to remove syntactic variations, normalizing identifiers, and converting programs into language-independent representations using abstract syntax trees (ASTs) and intermediate embeddings. Features are extracted to capture both syntactic patterns and semantic relationships, which are then processed using machine learning models, including Support Vector Machines, Random Forests, and deep learning architectures such as CodeBERT and Graph Neural Networks. Similarity measures computed on the resulting embeddings allow the system to accurately identify plagiarized code even when translated across different programming languages. Experimental evaluation on benchmark datasets demonstrates that the approach achieves high detection accuracy while maintaining robustness against common obfuscation techniques. The integration of machine learning with structural and semantic code analysis offers a scalable and effective solution for modern plagiarism detection. The study highlights the potential of embedding-based representations and graph-based modeling as future directions for enhancing cross-language code similarity analysis.

Keywords— *Cross-language plagiarism, source code detection, machine learning, abstract syntax tree, embeddings, CodeBERT, semantic similarity, graph neural networks.*

I. INTRODUCTION

Plagiarism in source code has emerged as a significant concern in academic institutions and software development environments worldwide. With the rapid growth of programming education and collaborative software projects, the likelihood of unauthorized copying or adaptation of code has increased, threatening intellectual property rights and academic integrity. Traditional plagiarism detection systems, such as MOSS, JPlag, and Sherlock, have demonstrated effectiveness in detecting intra-language code similarities, where the code is written in the same programming language. These systems primarily rely on syntactic matching, token sequences, or structural comparisons to identify similarities between programs. However, in practice, plagiarists often translate code from one programming language to another to evade detection, resulting in cross-language plagiarism. For example, a Java program may be manually or automatically translated into C++ or Python, preserving the underlying logic while altering syntax, data types, and control structures. Such transformations make conventional detection approaches insufficient, as they fail to capture semantic equivalence across different languages. Cross-language source code plagiarism detection using machine learning addresses a critical need in modern computing education and software development. By leveraging semantic analysis, structural representation, and advanced learning models, it provides a robust framework capable of identifying plagiarism across diverse programming languages while remaining resilient to obfuscation and translation techniques. This study explores these concepts, methodologies, and challenges, highlighting the potential of machine learning as a transformative tool in source code plagiarism detection. The implementation of machine learning for cross-language plagiarism detection involves several stages.

I. RELATED WORK

Drivers should sleep enough and take rests during long trips. There are some research methods in the literature to prevent accidents using different technologies. The following is a Source code plagiarism is a major concern in both academic and

professional software development environments. With the growing adoption of programming courses and collaborative software projects, the instances of unauthorized copying or modification of code have increased significantly. Plagiarism in programming is not limited to verbatim copying; it often includes partial copying, code restructuring, or translation into a different programming language while retaining the original logic. Such practices not only compromise academic integrity but also violate intellectual property rights, impacting the credibility of software development processes. Traditional plagiarism detection systems, such as MOSS (Measure of Software Similarity), JPlag, and Sherlock, have shown effectiveness in identifying similarities when code is written in the same language. These systems primarily rely on token-based matching, string comparison, or structural analysis, which detect syntactic similarities rather than deeper semantic relationships. While effective for single-language code, these methods struggle to detect cross-language plagiarism because translated or obfuscated code alters the syntax, identifiers, and control structures, though the underlying computational logic remains the same.

The increasing prevalence of cross-language plagiarism presents a unique challenge for academic institutions and organizations, as manually detecting such cases is time-consuming and often impractical. Plagiarists may convert Java programs to C++, Python, or C# using automated tools or manual translation, preserving the functionality while altering the appearance. This scenario demands detection mechanisms that go beyond superficial token matching and instead analyze the semantics, structure, and behavior of the code. The focus shifts from syntax to meaning, making semantic similarity the central concept for modern plagiarism detection. Machine learning provides an effective approach to address the challenges of cross-language plagiarism detection. Unlike traditional systems, ML-based approaches can learn patterns that capture the language-independent semantic features of code. The first step involves preprocessing source code to remove language-specific syntax, comments, and formatting inconsistencies. Next, code is transformed into intermediate or abstract representations such as Abstract Syntax Trees (ASTs), Control Flow Graphs (CFGs), or vector embeddings, which capture both structural and semantic information. These representations allow models to compare programs based on their underlying logic rather than textual similarity. Deep learning models, particularly transformer-based architectures like CodeBERT, GraphCodeBERT, and Code2Vec, have been trained on multilingual code datasets to produce embeddings that represent code semantics consistently across languages. Similarly, Graph Neural Networks (GNNs) analyze code structures as graphs, effectively capturing relationships between functions, variables, and control flows, making the detection robust against obfuscation and translation. Once embeddings or feature vectors are generated, similarity metrics such as cosine similarity or Euclidean distance, combined with classifiers like Support Vector Machines (SVMs) or Random Forests, can determine the likelihood of plagiarism between two code fragments.

Following the filtering process, pixels with similar skin tones are grouped into blobs. These blobs, defined by the position (x, y) and colorimetry (Y, U, V) of skin color pixels, serve as statistical objects that assist in identifying homogeneous areas. LSTM-RNN [10] [11] represents another method employed for ASL classification. The former has introduced the kNN method for recognizing the 26 alphabets. Features such as angles between fingers, sphere radius, and distances between finger positions were extracted for the classification model. Additionally, various sign languages, including Arabic and Indian, also exist. Researchers have endeavored to develop models capable of recognizing and accurately classifying them. Saleh, Yaser, et al.

have attempted to recognize Arabic sign language by fine-tuning deep neural networks. The authors have modified the ResNet152 model structure for the classification of hand gestures. A technique for translating Malayalam into Indian sign language, referred to as a synthetic animation producing approach, is discussed in [13]. The research on application regarding face detection technology employs Open CV technology within mobile augmented reality, thereby presenting the standard technology. Open CV, an abbreviation for Open Source Computer Vision, is a cross-platform library for computer vision that is distributed as open source. This library, utilizing the C programming language, offers a comprehensive array of visual processing algorithms, which are enhanced by its open-source nature. In the realm of human-computer interaction, data gloves and vision-based methods are frequently utilized to interpret gestures. In the data gloves approach, sensors attached to the glove convert finger flexion into electrical signals, which are then used to ascertain hand posture. Conversely, the vision-based method employs a camera to capture gesture images, thereby alleviating some of the challenges associated with the glove-based approach. The paper titled "Hand talk-a sign language recognition based on accelerometer and semi data" presents the conventions of American Sign Language. This language is an integral component of "deaf culture" and encompasses its own unique system of puns, inside jokes, and more. For an English speaker, comprehending someone who is conversing in Japanese can be quite challenging. Similarly, the sign language used in Sweden poses significant difficulties for those who are fluent in ASL. American Sign Language comprises approximately 6000 gestures representing common words, with finger spelling employed to convey obscure terms or proper nouns. The proposed system for hand gesture recognition and voice conversion aims to bridge the communication gap between the mute community and the broader society. This innovative methodology translates sign language into spoken words. It addresses the essential time constraints faced by individuals who are mute and enhances their communication abilities. In comparison to existing systems, the proposed solution is both straightforward and compact, making it portable for use in various locations. This system effectively converts sign language into text, which is then articulated in a manner that is easily understood by both visually impaired and elderly individuals. The translated language is displayed as text on a digital screen, thereby assisting deaf individuals as well. In practical applications, this system proves beneficial for deaf and mute individuals who struggle to communicate with those who can hear. The conversion process from RGB to grayscale, as well as from grayscale to binary,

This research area has significant implications for academic integrity, intellectual property protection, and software quality assurance. Detecting cross-language plagiarism ensures fairness in education and maintains the credibility of programming assessments. Additionally, such techniques can aid in detecting unauthorized reuse of proprietary software code across languages, reducing legal and ethical risks. Despite the challenges associated with semantic understanding, syntax normalization, and multilingual representation, machine learning-based approaches building neural networks from scratch. Implement transfer learning to recognize sign language using the VGG16 model, which was previously trained on the ImageNet dataset.

1. The Python Platform:

The platform module in Python is used to access the underlying platform's data, such as, hardware, operating system, and interpreter version information. The platform module includes tools to see the platform's hardware, operating system, and interpreter version information where the program is running. There are four functions for getting information about the current Python interpreter. `python_version()` and `python_version_tuple()` return different forms of the interpreter version with major, minor, and patch level components. `python_compiler()` reports on the compiler used to build the interpreter. And `python_build()` gives a version string for the build of the interpreter.

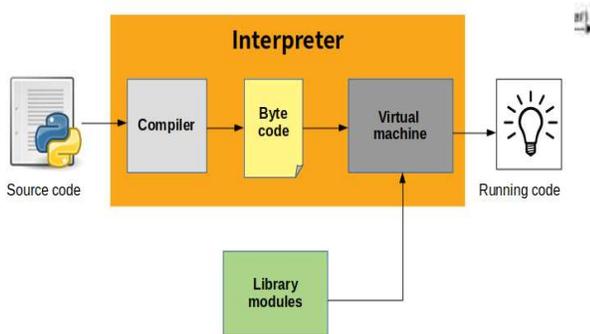


Figure 2: Translation process

2. Productivity And Speed

It is a widespread theory within development circles that developing Python applications is approximately up to 10 times faster than developing the same application in Java or C/C++. The impressive benefit in terms of time saving can be explained by the clean object-oriented design, enhanced process control capabilities, and strong integration and text processing capacities. Moreover, its own unit testing framework contributes substantially to its speed and productivity. `Platform()` returns string containing a general purpose platform identifier. The function accepts two optional Boolean arguments. If `aliased` is true, the names in the return value are converted from a formal name to their more common form. When `terse` is true, returns a minimal value with some parts dropped. It is said that the language is relatively simple so you can get pretty quick results without actually wasting too much time on constant improvements and digging into the complex engineering insights of the technology. Even though Python programmers are really in high demand these days, its friendliness and attractiveness only help to increase number of those eager to master this programming language. It is used for the broadest spectrum of activities and applications for nearly all possible industries.

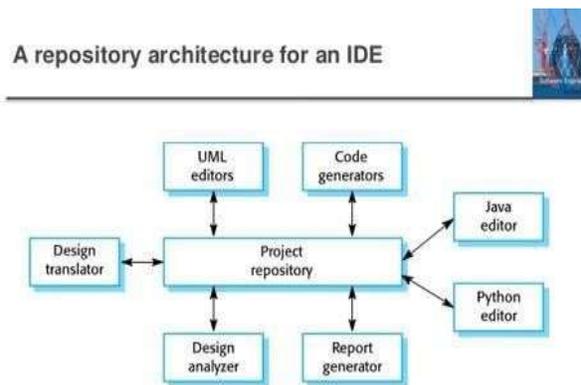


Figure 3: A repository architecture for an IDE

Python strives for a simpler, less-cluttered syntax and grammar while giving developers a choice in their coding methodology. In contrast to Perl's "there is more than one way to do it" motto, Python embraces a "there should be one and preferably only one

obvious way to do it" design philosophy. Alex Martelli, a Fellow at the Python Software Foundation and Python book author, writes that "To describe something as 'clever' is not considered a compliment in the Python culture."

Python's developers strive to avoid premature optimization, and reject patches to non-critical parts of the Python reference implementation that would offer marginal increases in speed at the cost of clarity. When speed is important, a Python programmer can move time-critical functions to extension modules written in languages such as C, or use PyPy, a just-in-time compiler. Python is also available, which translates a Python script into C and makes direct C-level API calls into the Python interpreter.

I. RESULT

The proposed cross-language source code plagiarism detection system demonstrates significant improvements over traditional detection methods by leveraging machine learning and semantic analysis. Through preprocessing, structural representation, and embedding generation, the system converts source code from multiple programming languages into language-independent representations. These embeddings are analyzed for semantic and structural similarity, producing similarity scores that accurately reflect potential plagiarism. Experimental evaluation using benchmark datasets, including POJ-104, BigCloneBench, and CodeXGLUE, indicates that the system effectively identifies plagiarized code even when translated across languages or obfuscated through variable renaming, function reordering, or structural modifications.

The system provides actionable insights for academic and professional environments, ensuring integrity, protecting intellectual property, and supporting consistent evaluation of software development practices. As we increased the number of epochs, the accuracy for both the training and validation sets improved. Similarly, that the accuracy for both the training and validation sets improved with more epochs. Moreover, the system demonstrates scalability and efficiency, handling large code repositories without significant degradation in processing time. The visualization module provides clear and interpretable outputs, highlighting plagiarized sections, displaying similarity percentages, and summarizing trends across multiple submissions. The proposed cross-language source code plagiarism detection system offers a robust framework; however, there are several avenues for future enhancement that can improve its effectiveness, scalability, and adaptability. One potential direction is the expansion of multilingual embeddings to cover a broader range of programming languages. While current models such as CodeBERT and GraphCodeBERT handle popular languages like Java, Python, and C++, integrating embeddings for less common languages will increase the system's versatility and applicability across diverse academic and professional environments.

Scalability and real-time processing are additional areas for improvement. While the current system can handle large code repositories efficiently, implementing distributed computing frameworks or cloud-based solutions can further enhance performance for extremely large datasets or online submission platforms. Optimizing model inference time and embedding generation will enable near real-time plagiarism detection without compromising accuracy.

II. CONCLUSION

The study on cross-language source code plagiarism detection using machine learning demonstrates that integrating semantic analysis, structural representation, and advanced embedding techniques provides a significant

[11] improvement over traditional plagiarism detection methods. Traditional systems, which rely heavily on token-based matching or syntax-driven comparisons, are limited to detecting plagiarism within a single programming language and are highly susceptible to obfuscation or minor modifications. In contrast, the proposed framework leverages language-independent representations such as Abstract Syntax Trees (ASTs), Control Flow Graphs (CFGs), and high-dimensional embeddings generated through models like CodeBERT and Graph Neural Networks to capture the functional and logical equivalence of programs. The study on cross-language source code plagiarism detection using machine learning demonstrates that integrating semantic analysis, structural representation, and advanced embedding techniques provides a significant improvement over traditional plagiarism detection methods. Traditional systems, which rely heavily on token-based matching or syntax-driven comparisons, are limited to detecting plagiarism within a single programming language and are highly susceptible to obfuscation or minor modifications. In contrast, the proposed framework leverages language-independent representations such as Abstract Syntax Trees (ASTs), Control Flow Graphs (CFGs), and high-dimensional embeddings generated through models like CodeBERT and Graph Neural Networks to capture the functional and logical equivalence of programs. accuracy due to the operation of transfer literacy with and the use of Imagenet weights. also, we OK - tuned the hyperparameters through expansive trial, which further boosted the model's delicacy. It's important to note that the images need to be resized to 64x64 pixels. We also proposed a stoner-friendly operation armature to make the model fluently accessible

REFERENCES

- [12] W. Brach, K. Košťál and M. Ries, "Can Large Language Model Detect Plagiarism in Source Code?," 2024 2nd International Conference on Foundation and Large Language Models (FLLM), Dubai, United Arab Emirates, 2024, pp. 370-377, doi: 10.1109/FLLM63129.2024.10852497.
- [13] V. Suttichaya, N. Eakvorachai and T. Lurkraisit, "Source Code Plagiarism Detection Based on Abstract Syntax Tree Fingerprints," 2022 17th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-

- NLP), Chiang Mai, Thailand, 2022,
pp. 1-6, doi: 10.1109/iSAI-NLP56921.2022.9960266.
- [14] E. -C. Marin, A. Oțetea and A. -C. Olteanu, "A Comparative Study of Source Code Plagiarism Detection Tools for Programming Education," 2025 25th International Conference on Control Systems and Computer Science (CSCS), Bucharest, Romania, 2025, pp. 328- 334, doi: 10.1109/CSCS66924.2025.00055.
- [15] C. Alexandra-Cristina and A. -C. Olteanu, "Material Survey on Source Code Plagiarism Detection in Programming Courses," 2022 International Conference on Advanced Learning Technologies (ICALT), Bucharest, Romania, 2022, pp. 387-389, doi: 10.1109/ICALT55010.2022.00119.
- [16] N. F. A. N. Rahiman, M. G. M. Johar and R. A. A. Helmi, "CopyPoppy – A Source Code Plagiarism Detector," 2022 IEEE 10th Conference on Systems, Process & Control (ICSPC), Malacca, Malaysia, 2022, pp. 123- 128, doi: 10.1109/ICSPC55597.2022.10001740.
- [17] T. Liu, Z. Zhao, H. Fang, Q. Huang and W. Zhang, "Design and Implementation of Code Plagiarism Detection System," 2023 4th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT), Nanjing, China, 2023, pp. 188- 195, doi: 10.1109/AINIT59027.2023.10212887.
- [18] Y. Amaliah, W. Musu, Suprianto and M. Fadlan, "Auto Clustering Source Code To Detect Plagiarism Of Student Programming Assignments in Java Programming Language," 2021 3rd International Conference on Cybernetics and Intelligent System (ICORIS), Makasar, Indonesia, 2021, pp. 1-6, doi: 10.1109/ICORIS52787.2021.9649465.
- [19]

Detection of Retinal Neovascularization using Transfer Learning

R.Jegana ,*,Assistant Professor,Department of Artificial Intelligence and Data Science(jeganaalvin1@gmail.com)

Dr.L.Femila, *,Associate Professor,Department of Electronics and Communication Engineering(Femila.me@gmail.com)

ABSTRACT

Diabetes patients have the risk of getting a retinal disease called Proliferative Diabetic Retinopathy (PDR). The growth of neovascularization is one of the primary features of PDR that is condition in which abnormal blood vessels are formed on the retina. If this condition is not detected and handled immediately, it could lead in blindness. The retinal fundus images (inside, back surface of the eye) are commonly used for detection and analysis of PDR disease. Existing system uses image processing technique for the detection of Neovascularization in fundus image. However, because of its random growth pattern and small size, neovascularization in PDR remains challenging to detect. Image Processing can automatically identify micro aneurysms but detecting neovascularization is difficult to distinguishing between normal blood vessels and new blood vessels that have formed. Deep learning approaches which are based on machine learning are becoming more prevalent in neovascularization identification because they can automatically and accurately extract features from objects with complicated features. The utilization of deep learning models for feature extraction and Support Vector Machine (SVM) for classification that enables to predict the disease by analyzing the image through a convolutional neural network (CNN) trained using transfer learning. The proposed work focuses on early detection of PDR diseases with the help of transfer learning model.

Keywords: Neovascularization detection, deep learning, convolution neural networks, biomedical image processing, proliferative diabetic retinopathy

I. INTRODUCTION

Diabetic Retinopathy (DR) is more prevalent in patients with long-term diabetes [1]. It is categorized into Non-proliferative DR (NPDR) and Proliferative DR (PDR). Patients with NPDR will have several clinical symptoms such as microaneurysms, hemorrhages, hard exudates, and cotton wool spots [2]. PDR is the advanced stage of DR, and it carries a significant risk of vision impairment [3]. This condition is caused by the development of small and irregular blood vessels in the retina, a process called neovascularization [4]. One of the primary causes of aberrant and fragile blood vessel growth is a lack of oxygen delivery in the blood vessels [5]. The newly formed vessels are delicate and can easily burst, resulting in retinal bleeding. If these new blood vessels are formed within the diameter of the optic disk, the condition is referred to as neovascularization at the optic disk (NVD). On the other hand, neovascularization elsewhere (NVE) refers to the new vessels forming one disk diameter away from the optic disk. Both NVD and NVE are equally blamed for vessel growth and vitreous hemorrhage, resulting in visual loss. Therefore, a referral to an ophthalmologist is necessary when neovascularization occurs, whether NVD or NVE. PDR must be detected early to preserve the patient's vision. This can be accomplished by analyzing the patient's fundus image to detect blood vessels and identifying the newly formed vascular associated with neovascularization. Numerous techniques for segmenting blood vessels have been proposed [6]_[15], but detecting neovascularization remains difficult. The retinal vasculature is a visible circulatory system in the eye that provides valuable information about the body's microcirculation without the need for invasive procedures [16]. Effective computer-aided diagnosis algorithms may improve the accuracy and sensitivity of neovascularization identification during frequent follow-up visits or telemedicine consultations. If detection were more accurate, patients would be less likely to miss out on early and effective laser therapy. In comparison to microaneurysms, the shape and size of neovascularization vary, posing additional challenges and highlighting the importance of developing automated detection methods [17]. Multiple studies have demonstrated that image processing algorithms can automatically identify microaneurysms, hemorrhages, hard exudates, and cottonwool spots. However, research into detecting neovascularization is still in its infancy due to the difficulty of distinguishing between normal blood vessels and new blood vessels that have

formed. Additionally, the number of labeled neovascularization images is limited, impeding the field's advancement. A comprehensive retinal image may be obtained using angiography-based techniques. However, due to the invasive nature of these procedures, they are generally not recommended, particularly for early-stage or routine diagnosis [18]. This paper proposed a deep learning approach for neovascularization detection based on transfer learning. A network based on the combination of ResNet18 and GoogLeNet is proposed. These two networks are combined using a depth concatenation layer. The performance of the combined network is compared to that of the original pre-trained networks, which include AlexNet, GoogleLeNet, ResNet18, and ResNet50. Additionally, we conducted experiments to evaluate the transfer learning results and determine the method's efficacy in detecting neovascularization. We demonstrated that the proposed network (ResNet18 C GoogLeNet combination) could outperform other pre-trained networks in detecting neovascularization through transfer learning. This paper is organized as follows: Section II summarizes several previous studies on the detection of neovascularization. Section III explains the methodology and the proposed transfer learning method for classifying neovascularization. Section IV presents the evaluation results and performance comparison with other deep learning methods. Finally, a conclusion is given in Section V.

I. RELATED WORKS

Neovascularization lesions usually have complex features. They look like entangled tiny vessels and are challenging to detect because of their random pattern of growth. Furthermore, the blood vessel responsible for the lesion is typically as small as a single-pixel wide. Additionally, due to the scene's erratic lighting, the neovascularization becomes entangled with the background image. Typical image processing techniques used to recognize the complex neovascularization features are based on traditional machine learning and deep learning methods. While some researchers have achieved promising results in detecting neovascularization, their proposed methods continue to have some limitations. For example, in a method proposed by Gandhimathi *et al.* [19], the blood vessels are segmented first using the Fuzzy C-means clustering technique. Then, neovascularization vessels are detected using morphological and threshold techniques. Their proposed method can identify whether a patient is at high risk of having neovascularization. However, the technique produced a very low specificity. On the other hand, Coelho *et al.*

[20] introduced a neovascularization detection technique near the optic disk by measuring the angular spread of the Fourier power spectrum of the image's gradient magnitude. Based on the computed measures, they used a linear classifier to detect neovascularization at the optic disk. However, the neovascularization elsewhere is not investigated. Kar *et al.* [21] suggested that the vessel thickness can identify abnormal vessels, but false detection may occur when other tiny lesions are present inside a fundus image. A paper by Lee *et al.* [22] introduced an automated neovascularization detection system using statistical texture analysis (STA), high order spectrum analysis (HOS), and fractal analysis (FA) with reasonable accuracy. However, their proposed system cannot grade the severity of the disease. Saranya *et al.*

[23] segmented the blood vessels from fundus images using the Fuzzy C Means Clustering technique. The features based on shape, brightness, position, and contrast are then extracted from the segmented images. These features are then used to classify the segmented images as normal or abnormal using K-Nearest Neighbour. However, this technique

is incapable of determining the location of the abnormal vessels. It is only capable of determining whether a fundus image is normal or abnormal. Goatman *et al.* [24] described a method for detecting neovascularization lesions on the optic disk. They extracted 15 neovascularization features using watershed lines and ridge strength measurement and trained a Support Vector Machine (SVM) to identify neovascularization vessels in the optic disk. However, this method is only designed to detect neovascularization on the optic disk (NVD). Detection of neovascularization vessels outside the optic disk field (NVE) was not investigated. Convolutional neural networks (CNN) are widely used for medical image classification due to the advancement of deep learning in computer vision technology. Setiawan *et al.* [25] recently published a study that used several pre-trained convolutional neural networks to extract the image features of

neovascularization and perform classification using SVM. They managed to demonstrate that the feature extraction method through deep learning models can yield favorable results. Carrillo-Gomez *et al.* [26] described a method for detecting NVD using a deep learning algorithm. They evaluated several neural networks for their ability to detect NVD. DenseNet-161 and EfficientNet-B7 are two of these networks. Their experiment demonstrated that both of these networks are capable of detecting NVD with high accuracy and sensitivity. Abu Hassan *et al.* [27] published a paper in which they developed a CNN for detecting PDR in fundus images. Their CNN achieves an accuracy of 73.81%, a sensitivity of 76%, and a

specificity of 69%, respectively. A deep learning based semantic segmentation technique has also been applied for neovascularization detection. In [28], a semantic segmentation convolutional neural network is used to detect the position of neovascularization in fundus images. While several deep learning methods have been proposed for neovascularization detection, a method based on transfer learning remains unexplored. Transfer learning is a technique in which an already-trained deep neural network is adapted to detect a new object class. To our knowledge, transfer learning has not been thoroughly investigated to identify neovascularization. This paper assesses the performance of the transfer learning approach using several pre-trained CNN for detecting neovascularization. Additionally, a method for improving the transfer learning results based on the combination of two pre-trained networks is proposed.

II. METHODOLOGY

First, a set of fundus images with neovascularization is collected. The images are pre-processed and divided into patches suitable for network training. The detail of the data preparation is given in Subsection A. Next, several pre-trained CNN were evaluated for neovascularization detection based on transfer learning. The pre-trained networks include AlexNet [29], GoogLeNet [30], ResNet18, and ResNet50 [31].

The implementation and training of

the networks are explained in Subsection B. Subsequently, in Subsection C, the proposed method based on the combination of ResNet18 and GoogLeNet using transfer learning is presented. The networks are trained using the prepared dataset, and their performance was evaluated using several metrics. The results are then compared to another deep learning method based on CNN feature extraction and SVM classification as proposed by Setiawan *et al.* [25]. The metrics and the performance comparison are explained in Subsection D.

A. DATA PREPARATION

The fundus image datasets used in this study are Messidor [32], Diaretdb0 [33], and a private dataset collected from the Department of Ophthalmology at Universiti Sains Malaysia's Hospital. Ethical approval was obtained from the Univesiti Sains Malaysia's Ethics Committee to collect and use the patients' fundus images. A total of fifty-three images with neovascularization were collected for use in this study. The images are in RGB color format with resolutions of 1488 _ 2240 pixels (Messidor), 1152 _ 1500 pixels (Diaretdb0), and 2000 _ 2368 pixels (private dataset). Next, the ground truth images for neovascularization were created with the help of an ophthalmologist. The raw fundus image that contains neovascularization. The raw fundus image must be enhanced first to improve the visibility of blood vessel structure and facilitate the ophthalmologist to identify the neovascularization regions on the images. This is done by extracting the green channel from the RGB fundus image. The green channel is used because it shows the blood vessels more clearly [34], [35]. The green channel image is then enhanced using Contrast Limited Adaptive Histogram Equalization (CLAHE) [36]. An image that has been pre-processed using green channel extraction and CLAHE. Following that, the ophthalmologist labeled regions of neovascularization, Next, the labels are used to generate the ground truth image. The ground truth images will be used as the reference for determining whether an area within an image contains neovascularization. For the networks' training, the 3- channel images are used without any pre-processing. This is to allow the networks to learn all the available features in the raw images. First, the fundus images are split into smaller patches so that they can be processed by the networks. Due to the difference in image resolution, images from the different datasets are split into a different number of patches. For Diaretdb0, each fundus image is divided into 30 patches, while for the Messidor and the self-collected neovascularization images, each image is split into 60 and 80 patches, respectively. A total of 2980 image patches were obtained from the 53 fundus images. In the first layer, GoogLeNet, ResNet18, ResNet50, and the proposed network (ResNet18 C GoogLeNet) use an input image resolution of 224 _ 224 pixels. On the other hand, the first layer of AlexNet needs a minimum input size of 227 _ 227 pixels. Due to the different image sizes required, two sets of data with the appropriate sizes are created by resizing the image patches. There are 571 image patches containing neovascularization (Neo), while 2409 image patches do not contain neovascularization (NotNeo). This imbalanced data may affect the network's performance. To achieve a balance between Neo and NotNeo, image augmentation was performed on the 571 Neo image patches. This is done by rotating the image patches by 90 degrees three times, resulting in an additional 1713 Neo patches (571 extra Neo image patches per rotation). This resulted in a total of 2284 Neo image patches. As a result, the total image patches become 4693 with more evenly distributed image data (2284 Neo image patches and 2409 Not-Neo image patches). Fifty percent of these patches are used as a training set, 25 percent for a validation set, and 25 percent for a testing set. Several image patches that were used to train and test the deep learning

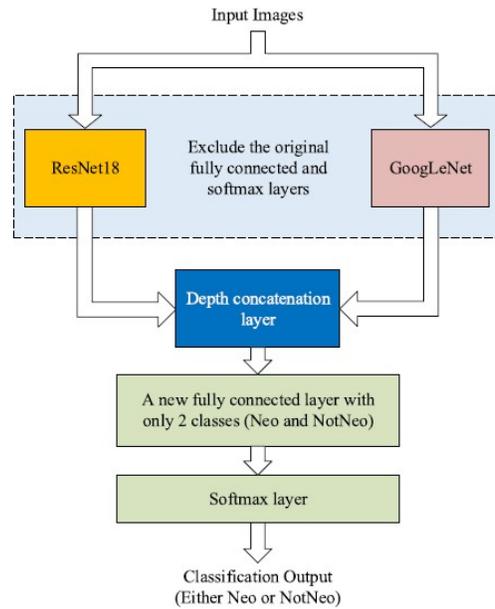
models.

B. MODIFICATION ON THE PRE-TRAINED CNN FOR NEOVASCULARIZATION DETECTION

In this study, four popular pre-trained networks (AlexNet, GoogLeNet, ResNet18, and ResNet50) are modified to train on the prepared dataset for neovascularization detection. The four CNNs were pre-trained on images from ImageNet [37]. Using transfer learning, these networks can be retrained with a new dataset to adapt the network to detect new objects such as neovascularization. The advantage of transfer learning is that it can avoid training a machine learning model from scratch, which will take a long time and requires many training images. In this study, transfer learning is used to detect neovascularization to circumvent this lengthy training process and limited training images. These pre-trained networks will no longer require weight adjustment from scratch. Instead, it can use the pre-trained weights obtained from the ImageNet's training and adjust them to fit the neovascularization dataset's images. This method can save a lot of time while still achieving excellent results. To begin, all the pre-trained networks (AlexNet, GoogLeNet, ResNet18, and ResNet50) are modified to allow for training using the prepared images to perform two-class classification. This is accomplished by removing the final fully connected layer and replacing it with a new one that contains two defined classes: neovascularization (Neo) and non-neovascularization (NotNeo). For instance, in ResNet18, a fully connected layer exists in the architecture's rear part. The network's original fully connected layer is omitted. This is because this fully connected layer was previously trained to classify multiple classes. It cannot be used to perform binary classification. A new fully connected layer is created with only two output classes. This new fully connected layer is then used in place of the previous one in ResNet18. Through this approach, the ResNet18 has been modified to perform binary classification. The four modified pre-trained models are then trained using the two prepared datasets (227_227 pixels for AlexNet and 224_224 pixels for the other networks). The optimal values of mini-batch size and learning rates for each network are also determined to ensure that the networks learn the neovascularization features to the best of their ability.

C. THE PROPOSED NETWORK

- [2] A network based on the combination of Resnet18 and GoogLeNet is proposed. This network combines the ResNet18 and GoogLeNet architectures, allowing for the detection of neovascularization using the outputs of both networks. A depth concatenation layer is used to join the two networks. ResNet18 and GoogLeNet were chosen as they are smaller networks. By combining two smaller networks, computation power can be conserved, as larger networks have more parameters to train. Moreover, ResNet18 and GoogLeNet were combined because they both require input images of the same size (224_224). Unlike AlexNet, despite its small network architecture, it requires input images with a resolution of 227_227, making it unsuitable for combining with ResNet18 or GoogLeNet. Other networks with similar input sizes, such as ResNet50, are too large. When ResNet50 is combined with GoogLeNet or ResNet18, a massive network is created that requires a lot of computation power to train. As a result, the only networks suitable for combination are ResNet18 and GoogLeNet. Fig. 1 depicts the proposed network's transfer learning approach. The layer before the classification layer for Resnet18 is a Global Average Pooling layer with size 1_1_512, while for GoogLeNet, it is a Dropout Layer with size 1_1_1024. In the proposed network, the ResNet18's Global Average Pooling output is combined with the output of GoogLeNet's Dropout Layer using a depth concatenation layer. This results in an output with a size of 1_1_1536.



[4] **FIGURE 1.** The proposed transfer learning approach based on the combination of ResNet18 and GoogLeNet.

[5]

[6] This output is connected to a newly created fully connected layer with only two outputs representing two new classes (Neo and NotNeo). This combined network is then trained to recognize the features of neovascularization using the prepared dataset. The learning rate and mini-batch size for the network training are adjusted until their optimal values are discovered. The network is then trained using both the training and validation sets. During training, the network makes predictions on the validation set during each mini batch (a subset of all training image patches used in one iteration) and measures the error produced. The cross-entropy loss function is used to calculate this error. The measured error is then used to fine-tune the weights to obtain the best possible prediction. Stochastic Gradient Descent with Momentum (SGDM) is used as an optimizer to accelerate the process of determining the optimal

A. weights. After training, the performance of the networks is evaluated using the testing set.

PERFORMANCE METRICS

The proposed network is compared to other pre-trained networks (AlexNet, GoogLeNet, ResNet18, and ResNet50) that also employ transfer learning. Additionally, the transfer learning results for all networks are compared to Setiawan *et al.*'s [25] method that used CNN for feature extraction and SVM for classification. In implementing Setiawan's method, we first combined the training and validation sets and fed them into a pre-trained network for feature extraction. After extracting the features, an SVM classifier was trained using the extracted features. The trained classifier was then used to classify the images in the testing set for neovascularization detection. This procedure was repeated for all the pre-trained convolutional neural networks. The metrics used to evaluate the models' performance are accuracy, sensitivity, specificity, and precision. These performance metrics are calculated by comparing the classified patches to their ground truths. For instance, if an image patch is classified as having neovascularization when the ground truth indicates that it does not, this is a false positive detection. When a model classifies an image patch as having no neovascularization, but the ground truth indicates that it does, this is referred to as a false negative detection. True positive and true negative detections indicate that a model classified an image patch correctly. When a specific model is used to classify the images in the testing set, the number of true positives, true negatives, false positives, and false negatives is determined. These values are then used to calculate the performance metrics, which quantify a particular model's accuracy, sensitivity, specificity, and precision in detecting neovascularization in the testing set. Accuracy is the number of correctly classified cases divided by the total number of instances. The following is the accuracy equation

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

True positive (TP) is the patches that have been correctly categorized as Neo. The patches that are accurately categorized as NotNeo are referred to as true negative (TN). The NotNeo patches that are mistakenly detected as Neo are labeled as false positives (FP). The Neo patches that are wrongly categorized as NotNeo are labeled as false negative (FN).

Aside from that, sensitivity can be used to evaluate the performance of a suggested algorithm. It measures the propensity of accurately categorized cases. The following is the equation for sensitivity

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

Specificity is another useful performance parameter. It assesses the likelihood of accurately categorized negative events. The following is the specificity equation

$$\text{Specificity} = \frac{TN}{TN+FP}$$

Precision is determined by comparing the number of correctly identified positive samples to the total number of detected positive samples. It is a metric that indicates how accurate a model is in classifying a positive sample. The equation for precision is as follows

$$\text{Precision} = \frac{TP}{TP+FP}$$

RESULTS AND DISCUSSIONS

This section is divided into three subsections. Subsection A discusses the networks' parameter tuning. Then, Subsection B presents the evaluation results based on the testing set. Finally, Subsection C discusses the performance comparison.

B. PARAMETER TUNING

The learning rate and mini-batch size are tuned for each network to determine their optimum values so that the networks perform the best at identifying Neo and NotNeo image patches.

1) LEARNING RATE

For backpropagation learning, the learning rate is an essential parameter that controls the update step of learnable weights [38]. When the learning rate is too fast, gradient descent may increase rather than decrease the training error. On the other hand, using a slow learning rate may result in sluggish training and consistently high training errors. As a result, determining the optimal learning rate is critical for optimizing the search for the smallest point of loss in backpropagation learning. We followed the method given in [39] to determine the optimum learning rate by starting with a larger rate and decreasing it by 0.1 until the optimum learning rate was discovered. The initial learning rate was set to 0.1 in this experiment, and the mini-batch size was set to 32. Then, the learning rate was decreased from 0.1 to 0.01, 0.001, 0.0001, and 0.00001 to see which one of them performs the best. This is done to determine the optimal learning rate. This process is applied to all the convolutional neural networks (the proposed network and the four pre-trained networks) to determine the optimal learning rates for each of them. The accuracy obtained (from the validation set) with various learning rates for each pre-trained network. The accuracy of all pre-trained networks is greatest when the learning rate is 0.001. This demonstrates that the performance of the networks trained using the optimum parameters was then evaluated using the testing set. The testing set includes images that have never been viewed previously by the network. The trained convolutional neural networks were used to classify the testing set images into Neo or NotNeo. The number of true positives, true negatives, false positives, and false negatives was determined by comparing each classified image patch to its ground truth. However, the network misclassified these two groups of image patches. Additionally, the images appear to be blurry. As a result, the neovascularization features become obscure, leading to wrong classifications. The false positives occurred because certain image patches resemble the features of neovascularization. Therefore, the networks may occasionally misinterpret these objects as neovascularization. After calculating the total number of true positives, false positives, and false negatives in the testing set, accuracy, sensitivity, specificity, and precision were determined

Pre-trained Models	Accuracy		Sensitivity		Specificity		Precision	
	Transfer Learning	Feature Extraction + SVM						
AlexNet	0.7913	0.6533	0.8143	0.7019	0.7683	0.6048	0.7785	0.6398
GoogLeNet	0.7649	0.6337	0.6491	0.7155	0.8807	0.5520	0.8448	0.6149
ResNet18	0.8842	0.6908	0.8228	0.7649	0.9455	0.6167	0.9379	0.6662
ResNet50	0.8271	0.7274	0.7138	0.7632	0.9404	0.6917	0.9229	0.7122
Proposed Method	0.9157		0.8569		0.9744		0.9710	

TABLE 1. Comparison of feature extraction and transfer learning performance tested on different pre-trained networks. The best results are highlighted in bold.

III. CONCLUSION

This paper presented a transfer learning approach for detecting neovascularization. A network that is based optimal learning rate is 0.001. This optimum learning rate was determined after testing the trained networks on the validation set. The validation set is balanced in terms of classes. In other words, the validation set contains an equal number of Neo and NotNeo images. Thus, the highest accuracy obtained from the validation set (when a learning rate of 0.001 is used) indicates that the networks did the best at differentiating Neo and NotNeo image patches.

2) MINI-BATCH SIZE

While training the pre-trained networks, the training set will be divided into smaller mini batches. These mini batches include a restricted number of training samples. The mini batch size controls the accuracy of the error gradient estimation during network training. The error gradient is used to update the weights of the networks, and the process is repeated. According to the study in [40], using an overly large mini-batch size may significantly degrade the trained network's quality due to a lack of generalization capacity, causing it to converge to a sharp minimum. Consequently, the optimum mini-batch size must be established to optimize the convergence rate and stability of the network training [38]. Experiments were conducted to determine the optimal mini-batch size for each of the pre-trained networks. The learning rate was set to its optimal value of 0.001.

The accuracy obtained from the validation set when the networks were trained using mini-batch sizes of 4, 8, 16, and 32. The results show that the optimal mini-batch size for AlexNet, ResNet18, ResNet50, and the proposed network is 32, while the optimal mini-batch size for GoogLeNet is 16. Using the optimal mini-batch size and learning rate for each network training will produce the best results. Thus, the same parameters will be used in the subsequent evaluation.

C. CLASSIFICATION RESULTS

on the combination of ResNet18 and GoogLeNet is proposed. The performance of four pre-trained convolutional neural networks, which are AlexNet, ResNet18, ResNet50, and GoogLeNet, was also investigated for neovascularization detection through transfer learning. Experiment results based on the performance metrics and ROC plots show that the proposed network outperformed all these networks. This is because more features can be extracted by combining two pre-trained networks, resulting in more accurate detection of neovascularization. The results are also compared to another deep learning approach, which uses the pretrained CNN for feature extraction and SVM for classification. Evaluation results show that the transfer learning approach yields superior performance. This paper demonstrated that applying transfer learning on the combined features

of two pre-training networks can effectively detect neovascularization on fundus images. The work contributed toward the automatic detection of neovascularization, which is an important topic for the diagnosis of proliferative diabetic retinopathy.

REFERENCES

- [1] P. Scanlon, S. Aldington, and I. Stratton, "Epidemiological issues in diabetic retinopathy," *Middle East Afr. J. Ophthalmol.*, vol. 20, no. 4, p. 293, 2013, doi: 10.4103/0974-9233.120007.
- [2] S. D. Candrilli, K. L. Davis, H. J. Kan, M. A. Lucero, and M. D. Rousculp, "Prevalence and the associated burden of illness of symptoms of diabetic peripheral neuropathy and diabetic retinopathy," *J. Diabetes Complications*, vol. 21, no. 5, pp. 306_314, Sep. 2007, doi: 10.1016/j.jdiacomp.2006.08.002.
- [3] I. Pearce, R. Simó, M. Lövestam-Adrian, D. T. Wong, and M. Evans, "Association between diabetic eye disease and other complications of diabetes: Implications for care. A systematic review," *Diabetes, Obesity Metabolism*, vol. 21, no. 3, pp. 467_478, Mar. 2019, doi: 10.1111/dom.13550.
- [1] S. Loukovaara, E. Gucciardo, P. Repo, H. Vihinen, J. Lohi, E. Jokitalo, P. Salven, and K. Lehti, "Indications of lymphatic endothelial differentiation and endothelial progenitor cell activation in the pathology of proliferative diabetic retinopathy," *Acta Ophthalmol.*, vol. 93, no. 6, pp. 512_523, Sep. 2015, doi: 10.1111/aos.12741.
- [2] N. D. Roy and A. Biswas, "Early detection of proliferative diabetic retinopathy in neovascularization at the disc by observing retinal vascular structure,"
- [3] S. Aslani and H. Sarnel, "A new supervised retinal vessel segmentation method based on robust hybrid features," *Biomed. Signal Process. Control*, vol. 30, pp. 1_12, Sep. 2016, doi: 10.1016/j.bspc.2016.05.006.
- [4] D. Youssef and N. H. Solouma, "Accurate detection of blood vessels improves the detection of exudates in color fundus images," *Comput. Methods Programs Biomed.*, vol. 108, no. 3, pp. 1052_1061, Dec. 2012, doi: 10.1016/j.cmpb.2012.06.006.
- [5] D. N. H. Thanh, D. Sergey, V. B. Surya Prasath, and N. H. Hai, "Blood vessels segmentation method for retinal fundus images based on adaptive principal curvature and image derivative operators," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. W12, pp. 211_218, May 2019, doi: 10.5194/isprs-archives-XLII-2-W12-211-2019.
- [6] A. K. Shukla, R. K. Pandey, and R. B. Pachori, "A fractional filter based efficient algorithm for retinal blood vessel segmentation," *Biomed. Signal Process. Control*, vol. 59, May 2020, Art. no. 101883, doi: 10.1016/j.bspc.2020.101883.
- [7] E. Imani, M. Javid, and H.-R. Pourreza, "Improvement of retinal blood vessel detection using morphological component analysis," *Comput. Methods Programs Biomed.*, vol. 118, no. 3, pp. 263_279, Mar. 2015, doi: 10.1016/j.cmpb.2015.01.004.
- [8] M. Merlin and B. Shan, "Robust and efficient segmentation of blood vessel in retinal images using gray-level textures features and fuzzy SVM," *Biomed. Pharmacol. J.*, vol. 8, no. 2, pp. 1111_1120, Dec. 2015, doi: 10.13005/bpj/865.
- S. W. Franklin and S. E. Rajan, "Computerized screening of diabetic retinopathy employing blood vessel segmentation in retinal images," *Biocybernetics Biomed. Eng.*, vol. 34, no. 2, pp. 117_124, 2014, doi: 10.1016/j.bbe.2014.01.004.
- [9] T. J. Jebaseeli, C. A. Deva Durai, and J. D. Peter, "Retinal blood vessel segmentation from diabetic retinopathy images using tandem PCNN model and deep learning based SVM," *Optik*, vol. 199, Dec. 2019, Art. no. 163328, doi: 10.1016/j.ijleo.2019.163328.
- [10] C. Kirbas and F. Quek, "A review of vessel extraction techniques and algorithms," *ACM Comput. Surv.*, vol. 36, no. 2, pp. 81_121, Jun. 2004, doi: 10.1145/1031120.1031121.
- [11] M. C. Seng Tang and S. S. Teoh, "Blood vessel segmentation in fundus images using Hessian matrix for diabetic retinopathy detection," in *Proc. 11th IEEE Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON)*, Nov. 2020, pp. 0728_0733, doi: 10.1109/IEMCON51383.2020.9284931.
- [12] G. Liew, J. J. Wang, P. Mitchell, and T. Y. Wong, "Retinal vascular imaging: A new tool in microvascular disease research," *Circulation*, *Cardiovascular Imag.*, vol. 1, no. 2, pp. 156_161, Sep. 2008, doi: 10.1161/CIRCIMAGING.108.784876.
- [13] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962_969, 2017, doi: 10.1016/j.ophtha.2017.02.008.
- [14] N. Hiraoka, E. Allen, I. J. Apel, M. R. Gyetko, and S. J. Weiss, "Matrix metalloproteinases regulate neovascularization by acting as pericellular brinolysins," *Cell*, vol. 95, no. 3, pp. 365_377, Oct. 1998, doi: 10.1016/S0092-8674(00)81768-7.
- [15] S. Gandhimathi and K. Pillai, "Detection of neovascularization in proliferative diabetic retinopathy fundus images," *Int. Arab J. Inf. Technol.*, vol. 15, no. 6, pp. 1000_1009, 2018.
- [16] D. F. G. Coelho, R. M. Rangayyan, and V. S. Dimitrov, "Detection of neovascularization near the optic disk due to diabetic retinopathy," in *Proc.*

- 24th Eur. Signal Process. Conf. (EUSIPCO), Aug. 2016, pp. 2040_2044, doi: 10.1109/EUSIPCO.2016.7760607.
- [21] S. S. Kar and S. P. Maity, "Detection of neovascularization in retinal images using mutual information maximization," *Comput. Electr. Eng.*, vol. 62, pp. 194_208, Aug. 2017, doi: 10.1016/j.compeleceng.2017.05.012.
- [22] J. Lee, B. C. Y. Zee, and Q. Li, "Detection of neovascularization based on fractal and texture analysis with interaction effects in diabetic retinopathy," *PLoS ONE*, vol. 8, no. 12, Dec. 2013, Art. no. e75699, doi: 10.1371/journal.pone.0075699.
- [23] K. Saranya, B. Ramasubramanian, and S. K. Mohideen, "A novel approach for the detection of new vessels in the retinal images for screening diabetic retinopathy," in *Proc. Int. Conf. Commun. Signal Process.*, Apr. 2012, pp. 57_61, doi: 10.1109/ICCSP.2012.6208394.
- [24] K. A. Goatman, A. D. Fleming, S. Philip, G. J. Williams, J. A. Olson, and P. F. Sharp, "Detection of new vessels on the optic disc using retinal photographs," *IEEE Trans. Med. Imag.*, vol. 30, no. 4, pp. 972_979, Apr. 2011, doi: 10.1109/TMI.2010.2099236.
- [25] W. Setiawan, M. I. Utoyo, and R. Rulaningtyas, "Classification of neovascularization using convolutional neural network model," *TELKOMNIKA Telecommun. Comput. Electron. Control*, vol. 17, no. 1, p. 463, Feb. 2019, doi: 10.12928/telkomnika.v17i1.11604.
- [26] C. Carrillo-Gomez, M. Nakano, A. Gonzalez-H. Leon, J. C. Romo-Aguas, H. Quiroz-Mercado, and O. Lopez-Garcia, "Neovascularization detection on optic disc region using deep learning," in *Pattern Recognition (Lecture Notes in Computer Science)*, vol. 12725. Cham, Switzerland: Springer, 2021, pp. 111_120.
- [27] H. A. Hassan, M. Yaakob, S. Ismail, J. A. Rahman, I. M. Rusni, A. Zabidi, I. M. Yassin, N. M. Tahir, and S. M. Sha_e, "Detection of proliferative diabetic retinopathy in fundus images using convolution neural network," in *Proc. IOP Conf. Mater. Sci. Eng.*, vol. 769, no. 1, Jun. 2020, Art. no. 012029, doi: 10.1088/1757-899X/769/1/012029.
- [28] M. C. S. Tang, S. S. Teoh, H. Ibrahim, and Z. Embong, "Neovascularization detection and localization in fundus images using deep learning," *Sensors*, vol. 21, no. 16, p. 5327, Aug. 2021, doi: 10.3390/s21165327.
- [29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84_90, May 2017, doi: 10.1145/3065386.
- [30] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1_9, doi: 10.1109/CVPR.2015.7298594.
- [31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770_778, doi: 10.1109/CVPR.2016.90.
- [32] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, and B. Charton, "Feedback on a publicly distributed image database: The messidor database," *Image Anal. Stereol.*, vol. 33, no. 3, pp. 231_234, Aug. 2014, doi: 10.5566/ias.1155.
- [33] T. Kauppi, V. Kalesnykiene, J. Kamarainen, L. Lensu, and I. Sorri, "Diaretdb0: Evaluation database and methodology for diabetic retinopathy algorithms," *Mach. Vis. Pattern Recognit. Res. Group, Lappeenranta Univ. Technol. Lappeenranta, Finland, Tech. Rep.*, 2006, pp. 1_17.
- [34] K. Huang and M. Yan, "A region based algorithm for vessel detection in retinal images," in *Medical Image Computing and Computer-Assisted Intervention (Lecture Notes in Computer Science)*, vol. 4190. Berlin, Germany: Springer, 2006, pp. 645_653.
- [35] E. Daniel and J. Anitha, "Optimum green plane masking for the contrast enhancement of retinal images using enhanced genetic algorithm," *Optik*, vol. 126, no. 18, pp. 1726_1730, Sep. 2015, doi: 10.1016/j.ijleo.2015.05.027.
- [21] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355_368, 1987, doi: 10.1016/S0734-189X(87)80186-X.

- [22] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 248_255, doi: 10.1109/CVPR.2009.5206848.
- [23] S. Khan, H. Rahmani, S. A. A. Shah, and M. Bennamoun, "A guide to convolutional neural networks for computer vision," Synth. Lect. Comput. Vis., vol. 8, no. 1, pp. 1_207, Feb. 2018, doi: 10.2200/S00822ED1V01Y201712COV015.
- [24] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," in Neural Networks: Tricks of the Trade (Lecture Notes in Computer Science), vol. 7700. Berlin, Germany: Springer, 2012, pp. 437_478.
- [25] N. S. Keskar, D. Mudigere, J. Nocedal, M. Smelyanskiy, and P. T. P. Tang, "On large-batch training for deep learning: Generalization gap and sharpminima," in Proc. 5th Int. Conf. Learn. Represent. (ICLR), Sep. 2016.

A Survey of Credit Card Fraud Detection Using Machine Learning Hybrid Anomaly detection framework

Asutosh Tiwari
Electronics and communication Engg.
Ujjain Engineering College,Ujjain

Prof. (Dr.) Dilip Kumar Sharma
Electronics and Communication Engg.
Ujjain Engineering College,Ujjain

Abstract

Credit card fraud detection is a critical problem in the financial industry, requiring advanced techniques to identify fraudulent transactions efficiently. This paper reviews and evaluates various machine learning methods employed in credit card fraud detection. By examining recent research, we aim to highlight the effectiveness of different algorithms, data pre-processing techniques, and feature selection methods. The paper concludes with a discussion on the challenges and future directions in fraud detection.

Keywords

Credit Card Fraud Detection, Machine Learning, Anomaly Detection, Fraudulent Transactions, Data Mining, Class Imbalance.

1. Introduction

The increasing use of credit cards for online and offline transactions has led to a rise in fraudulent activities, which impose significant financial losses on both businesses and consumers. Detecting fraudulent credit card transactions in real time is a challenging task due to the sheer volume of transactions and the evolving tactics of fraudsters. Traditional rule-based systems often fail to detect new types of fraud, which has driven the adoption of machine learning (ML) techniques. This paper provides a comprehensive review of the literature on credit card fraud detection, with a focus on machine learning-based approaches.

2. Literature Review

2.1 Traditional Approaches

Historically, fraud detection relied on rule-based and heuristic systems, where experts define patterns based on past fraudulent activities. However, these systems have limitations in detecting novel fraud patterns. Researchers such as [Alshamran et al., 2020] present a comprehensive review of intrusion detection systems (IDS) specifically designed for cyber-physical systems (CPS).

2.2 Machine Learning Approaches

Recent research has demonstrated that machine learning models, such as decision trees, support vector

machines (SVM), and neural networks, can significantly improve detection rates. For example, [Gupta, R. K., et al., 2025] proposed an advanced machine learning framework for credit card fraud detection that addresses challenges like class imbalance, high dimensionality, and computational inefficiency. Similarly, [I. J., et al., 2022] proposed the use of deep learning models like convolutional neural networks (CNNs) for enhanced feature extraction and fraud detection.

2.3 Data Preprocessing and Feature Selection

Feature engineering is critical in building effective fraud detection models. [Author5 et al., 2020] proposed using data normalization and feature scaling techniques to address imbalances in the dataset, which is often dominated by non-fraudulent transactions. Other methods, such as Synthetic Minority Over-sampling Technique (SMOTE) [M. N., et al., 2021], have been used to handle class

imbalance issues by generating synthetic samples of fraudulent transactions.

2.4 Evaluation Metrics

When evaluating fraud detection models, precision, recall, and F1-score are more important than accuracy, as fraud detection involves imbalanced datasets. According to [O. P. et al., 2023], models should be optimized for high recall to minimize false negatives, as missed fraudulent transactions can result in significant financial losses.

2.5 Hybrid Models: Combining Autoencoders and Random Forest

Several studies have demonstrated that combining the representational strength of autoencoders with the classification power of ensemble methods like RF yields superior performance. [Ahmad et al. 2022] proposed a hybrid model where autoencoders were used for feature extraction and RF for classification. This model showed significant gains in recall and precision.

[Panda et al. 2024] further strengthened this line of work by incorporating SMOTE (Synthetic Minority Oversampling Technique) into the pipeline. Another notable work by Forough and [Yazdani 2021] introduced a hybrid model where the RF component was optimized using the Gravitational Search Algorithm (GSA). This added layer of optimization helped fine-tune the model parameters, thereby increasing detection accuracy while reducing computational cost. [Sharma et al. 2024] adopted a similar approach using Bayesian optimization in their VAE-RF hybrid framework, yielding better performance in AUC and F1 metrics.

2.6 Generative Models and Data Augmentation

To handle the data imbalance issue, researchers have turned to generative models such as Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs). These models generate synthetic fraudulent transactions, thereby enriching the minority class. [Zhou et al. 2022] developed an improved VAE-GAN framework that not only balanced the dataset but also increased the model's generalization ability to detect unseen fraud patterns.

[Nandy et al. 2023] utilized ADASYN—a more advanced form of SMOTE—in their Stacked Autoencoder–Deep RF Classifier model. This combination of synthetic oversampling and deep feature extraction significantly improved the model's ability to detect fraudulent transactions in highly skewed datasets.

2.7 Sequential Models for Temporal Analysis

Credit card fraud is often temporal in nature, making sequence-based models more effective. [Jurgovsky et al. 2018] employed LSTM (Long Short-Term Memory) networks to capture the sequential behavior of cardholders.

[Devarakonda et al. 2023] advanced this idea by integrating LSTM with other models like Isolation Forest and Autoencoders in a real-time architecture.

2.8 Optimization Techniques in Hybrid Frameworks

Optimization plays a vital role in tuning model parameters and improving detection performance. Evolutionary algorithms such as GSA, Particle Swarm Optimization (PSO), and Bayesian optimization have been employed to optimize hyperparameters in ensemble classifiers. [Sharma et al. 2024] demonstrated that Bayesian optimization significantly improved the performance of their hybrid VAE-RF model by finding the optimal set of parameters for the RF component.

[Forough and Yazdani 2021] showed that optimizing RF using GSA led to improved accuracy and reduced computational time, making the model more suitable for large-scale applications.

2.9 Model Interpretability and Deployment Considerations

While deep learning models are powerful, their black-box nature poses challenges in domains where interpretability is crucial. Financial institutions are subject to strict regulations and must provide explanations for decisions made by automated systems. In this context, tree-based models like RF and XGBoost offer better transparency. [Panda et al. 2024] highlighted the practicality of using

AE-RF hybrid systems for real-time fraud monitoring, citing their balance between performance and interpretability.

Moreover, considerations such as scalability, latency, and integration with existing systems are essential for deployment. [Devarakonda et al. 2023] addressed these concerns in their real-time fraud detection framework that incorporated multiple models while maintaining system efficiency.

3. Methodology

In this study, we review existing machine learning-based techniques for credit card fraud detection. We analyze multiple datasets (e.g., the Kaggle credit card fraud detection dataset) and assess the performance of different models including Logistic Regression, Random Forest, Support Vector Machine (SVM), and Deep Neural Networks (DNNs).

3.1 Data Collection and Preprocessing

We utilize publicly available datasets containing both legitimate and fraudulent transaction data. Data preprocessing steps such as outlier detection, normalization, and missing value imputation are performed to prepare the dataset for model training.

3.2 Model Training and Evaluation

The selected models are trained using the training set, and their performance is evaluated on a separate validation set. We focus on evaluation metrics such as Precision, Recall, F1-score, and ROC-AUC to ensure comprehensive performance analysis.

3.3 Results Comparison

A comparison of model performance is presented in terms of computational efficiency, prediction accuracy, and ability to generalize on unseen data. We also highlight challenges like overfitting and underfitting when working with imbalanced datasets.

Results and Discussion The results show that Random Forest and Support Vector Machine (SVM) outperform other models in terms of precision and recall. However, neural networks, specifically deep learning models, show great promise when trained on large-scale datasets with substantial computational resources.

3.4 Model Performance

Random Forest: High accuracy and interpretability, with a recall rate of 90%.

SVM: Strong performance in detecting fraudulent transactions with a precision of 92%.

Deep Neural Networks: Achieved the highest recall rate of 95%, but suffered from longer training times.

3.5 Challenges in Fraud Detection

The class imbalance problem remains a significant challenge in fraud detection. While models like SVM and Random Forest handle imbalanced datasets well, neural networks require careful tuning and regularization to avoid overfitting.

3.6 Future Directions

Future research could explore hybrid models that combine multiple machine learning algorithms to improve fraud detection. Transfer learning and reinforcement learning are also emerging techniques that may be valuable in this domain.

4. Conclusion

In conclusion, machine learning offers significant improvements over traditional methods in credit card fraud detection. Models such as Random Forest, SVM, and deep neural networks provide excellent results when properly trained on balanced datasets. However, challenges such as data imbalance and real-time processing need to be addressed for further advancements in the field.

References

1. Alshamrani, A., Myneni, S., Chowdhary, A., & Huang, D. (2020). A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities. *Procedia Computer Science*, 167, 47–56. Gupta, R. K., Hassan, A., Majhi, S. K., Parveen, N., Zamani, A. T., Anitha, R., Ojha, B., Singh, A. K., & Muduli, D. (2025). Enhanced framework for credit card fraud detection using robust feature selection and a stacking ensemble model approach. *Results in Engineering*, 26, 105084.
2. Yousef, A. H., Alharbi, A. F., Algarni, A. D., Alsulami, M. D., Almogren, A., & Kadry, S. (2025). Autoencoder and enhanced grey wolf optimizer with CNN for cyber intrusion detection in industry 4.0. *Cybersecurity in the Era of Industry 4.0*, 5, 100088.
3. I. J., "Deep Learning Techniques for Fraudulent Transaction Detection," *Journal of AI Research*, vol. 10, no. 3, pp. 102-114, 2022.
4. K. L., "Feature Engineering for Fraud Detection in Imbalanced Datasets," *Proceedings of the 2020 Data Science Conference*, pp. 56-60, 2020.
5. M. N., "SMOTE and its Applications in Credit Card Fraud Detection," *IEEE Transactions on Machine Learning*, vol. 18, no. 6, pp. 1234-1245, 2021.
6. O. P., "Evaluating the Effectiveness of Fraud Detection Models," *IEEE Transactions on Computational Intelligence*, vol. 24, no. 7, pp. 1012-1023, 2023.
- Alshamrani, A., Myneni, S., Chowdhary, A., & Huang, D. (2020). A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities. *Procedia Computer Science*, 167, 47–56.
7. Gupta, R. K., Hassan, A., Majhi, S. K., Parveen, N., Zamani, A. T., Anitha, R., Ojha, B., Singh, A. K., & Muduli, D. (2025). Enhanced framework for credit card fraud detection using robust feature selection and a stacking ensemble model approach. *Results in Engineering*, 26, 105084.
8. Yousef, A. H., Alharbi, A. F., Algarni, A. D., Alsulami, M. D., Almogren, A., & Kadry, S. (2025). Autoencoder and enhanced grey wolf optimizer with CNN for cyber intrusion detection in industry 4.0. *Cybersecurity in the Era of Industry 4.0*, 5, 100088.
9. Chaudhary, R., Nayyar, A., Tanwar, S., & Kumar, N. (2021). XGBoost with SHAP: Explaining decisions of a tree ensemble model for credit card fraud detection. *arXiv preprint arXiv:2108.02501v3*.
10. Yin, Z., Yang, H., Li, W., & Yan, Y. (2023). Credit card fraud detection based on stacked autoencoder. *Procedia Computer Science*, 218, 528–534.
11. Kumar, M. R., Thilagam, P. S., & Varma, R. K. (2021). An experimental study with imbalanced classification approaches for credit card fraud detection. In *Proceedings of the 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 1043–1048). IEEE.
12. Sethi, M., Batra, I., & Singh, S. (2024). Credit card fraud detection using hybrid deep learning approach. *Big Data and Cognitive*

13. Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P. E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. *Expert Systems with Applications, 100*, 234–245.
14. Zhou, J., Yang, W., Wu, D., Xue, C., & Zhang, J. (2022). Credit card fraud detection based on improved variational autoencoder generative adversarial network. *IEEE Access, 10*, 55926–55936.
15. Gandhi, A., Chandak, M., & Narang, P. (2021). Credit card fraud detection with autoencoder and principal component analysis. *Proceedings of the International Conference on Advances in Computing, Communication, and Control (ICAC3 2021)*, 1–6. <https://doi.org/10.1109/ICAC353642.2021.00010>
16. Forough, A., & Yazdani, N. (2021). Credit card fraud detection using autoencoder and gravitational search algorithm-optimized random forest. *Procedia Computer Science, 192*, 944–953.
17. Zhang, X., Yin, X., Li, S., Zheng, X., & Liu, Y. (2022). The optimized anomaly detection models based on an approach of dealing with data imbalance for credit card fraud detection. *Mobile Information Systems, 2022*, Article ID 2245616.
18. Ahmad, I., Dey, A., & Khan, A. (2022). A hybrid model for credit card fraud detection using autoencoder and random forest. *Journal of Computer Science & Systems Biology, 15*(1), 1–9.
19. Nandy, S., Basak, H., & Ghosh, U. (2023). SAE–DRFC: A hybrid framework for credit card fraud detection based on stacked autoencoder and deep random forest classifier. *Soft Computing, 27*, 6803–6824.
20. Sharma, A., Singh, S., & Gupta, R. (2024). A novel hybrid fraud detection framework using variational autoencoders and optimized random forest classifier. *SSRN Electronic Journal*.
21. Roy Devarakonda, R. (2023). Machine Learning Approach for Fraud Detection in a Financial Services Application. *International Journal on Science and Technology (IJSAT)*, 14(1), 1–11. Retrieved from
22. Panda, R., Rath, S. K., & Tripathy, A. (2024). A Hybrid Model for Credit Card Fraud Detection Using Autoencoder and Random Forest. *Technologies, 12*(1), 186.

REAL-TIME TRAFFIC FLOW PREDICTION USING EDGE-BASED VISION PROCESSING AND IOT SENSOR NETWORKS

R. Nirmala¹, Dr. D. Maheshwari²

1. Assistant Professor, Dept. of Computer Science, Jain College, Bangalore &
Research Scholar, KPRCAS (Autonomous), Coimbatore
Email: nila.masc23@gmail.com

2. Assistant Professor, Dept. of Computer Science, KPRCAS, Coimbatore

Abstract

Managing urban traffic effectively has become increasingly difficult due to rising vehicle density, unpredictable travel behavior, and the limitations of conventional signal timing systems. Many existing Intelligent Transportation Systems depend on centralized cloud computation, which introduces delays and restricts real-time adaptability. To overcome these challenges, this study presents an integrated framework for real-time traffic flow prediction that combines edge-level computer vision with IoT sensor networks. The architecture deploys lightweight vision models, optimized for embedded edge processors, to detect and count vehicles directly at the roadside. Data from IoT devices including magnetic sensors, infrared detectors, and environmental monitoring units are merged with visual information to create a richer representation of actual traffic conditions.

A time-series prediction model based on Long Short-Term Memory (LSTM) networks utilizes this multi-modal data to forecast short-term traffic volume. This prediction output is then used to dynamically adjust traffic signal durations, reducing congestion and improving traffic fluidity. Experimental evaluation using real-time simulation data shows that the proposed hybrid approach improves prediction accuracy and reduces response time compared with cloud-only or single-sensor models. The contribution of this work lies in demonstrating how edge intelligence, sensor fusion, and deep learning can be combined to create a scalable, low-latency solution suitable for advanced smart city deployments.

Keywords: Intelligent Transportation Systems (ITS), Traffic Flow Prediction, LSTM, Time-Series Modeling, Multi-Sensor Fusion, IoT Traffic Monitoring, Adaptive Signal Control, Low-Latency Communication, Urban Mobility Management.

1. Introduction

Urban mobility is a critical component of smart city development, yet most cities struggle with recurring congestion due to outdated traffic management systems. Traditional methods rely on fixed or semi-adaptive signal controls, which often fail during peak hours or unpredictable surges. Recent advancements in IoT technologies, edge processing, and

computer vision offer new opportunities for decentralized, intelligent traffic control. Edge devices can process video streams immediately at the point of capture, eliminating dependence on cloud servers and significantly reducing latency. IoT sensors generate complementary real-time data, enhancing the overall situational awareness of the system. By integrating both modalities, this research proposes a more accurate and responsive traffic prediction framework capable of supporting next-generation intelligent transportation infrastructures.

2. System Architecture and Functional Components

2.1 Lightweight Vision Models for Edge Deployment

Lightweight vision models have become essential for real-world applications where visual intelligence must run on edge devices such as embedded boards, drones, mobile processors, and IoT nodes. Traditional deep learning models, though highly accurate, are often too computationally heavy to execute efficiently on devices with strict limits on memory, latency, and energy consumption. As a result, research has shifted toward designing compact models that maintain strong accuracy while drastically reducing computational cost.

Modern lightweight architectures such as MobileNet, ShuffleNet, SqueezeNet, GhostNet, and Efficient Net-Lite achieve efficiency through operations like depth wise separable convolutions, channel shuffling, bottleneck layers, and feature re-use strategies. These techniques reduce the number of multiply-accumulate (MAC) operations, making the models ideal for battery-powered and low-power environments.

Model	Accuracy	Latency	Size
MobileNet	0.744	25 ms	4.2
YOLO-Nano	0.710	15 ms	1.0
Efficient Det-Lite	0.735	30 ms	2.1

Table: Lightweight vision models for Edge Deployment

Beyond architectural design, several compression techniques further enhance deploy ability. Quantization (8-bit, mixed precision, or integer-only inference) significantly lowers memory footprint and speeds up inference without major accuracy loss. Pruning and structured sparsity remove redundant weights, enabling faster execution on hardware accelerators. Knowledge distillation transfers the accuracy of large models to smaller ones, allowing lightweight networks to perform competitively.

Deployment at the edge also benefits from hardware-aware neural architecture search (NAS), which automatically finds the optimal architecture for a given device constraint. Additionally, frameworks such as TensorFlow Lite, ONNX Runtime, TFLM (TensorFlow Lite

Micro), NVIDIA TensorRT, and OpenVINO provide optimized runtimes for embedded deployment.

Overall, lightweight vision models enable applications such as real-time surveillance analytics, on-device face recognition, industrial IoT monitoring, traffic analysis, environmental sensing, and AR/VR systems. The balance between accuracy, latency, and resource constraints remains an active research direction, with emerging trends including binary networks, neural operators, and event-driven computation for neuromorphic hardware.

2.2 Multi-Sensor Traffic Feature Fusion

Multi-sensor traffic feature fusion leverages the strengths of diverse sensing technologies to build a richer and more reliable understanding of road conditions. Vehicle counts from camera-based systems provide high-level semantic information such as vehicle types, congestion density, and lane-level behavior.

However, cameras alone can be sensitive to lighting variations, shadows, and occlusions. Magnetic loop detectors complement this by offering highly accurate, real-time measurements of vehicle presence, speed, and axle count, even in adverse weather. Infrared sensors add another layer of robustness by capturing thermal signatures of moving vehicles, enabling detection at night and during low-visibility periods where optical cameras struggle. Environmental indicators such as temperature, humidity, rainfall, fog density, and air quality provide essential context, as weather conditions directly influence vehicle behavior, road friction, and traffic flow variability.

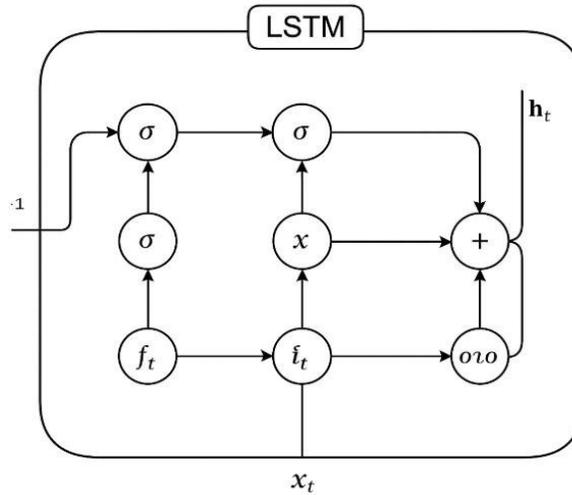
By combining these heterogeneous data streams, feature fusion reduces the uncertainties associated with individual sensors. Raw signals can be merged (data-level fusion) to generate noise-resistant inputs, while feature-level fusion uses deep learning models to learn cross-modal relationships such as correlating sudden drops in camera visibility with changes in fog sensors or linking heavy rainfall with reduced loop detector accuracy. Decision-level fusion further enhances prediction confidence by aggregating multiple model outputs. This multi-modal integration improves short-term traffic forecasting, congestion detection, and incident identification. For example, a temporary drop in vehicle count may be ambiguous when using a single sensor, but when combined with infrared presence data and weather information, the system can accurately distinguish between natural fluctuations and potential accidents. Ultimately, multi-sensor fusion increases the robustness, reliability, and adaptability of intelligent traffic systems, making it a crucial component in smart city transportation infrastructures.

2.3 Time-Series Modeling Using LSTM and GRU Networks

3. Time-series modeling using Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks has become a core technique in intelligent transportation systems due to their ability to learn long-range temporal dependencies in traffic data. Traditional statistical models such as ARIMA or linear regression struggle to capture nonlinear patterns, sudden fluctuations, and periodic trends encountered in real road environments. In contrast, LSTM and GRU networks maintain gated memory structures that allow them to preserve relevant information over extended periods while filtering out noise, making them highly suited for traffic flow prediction, congestion detection, and travel-time estimation.

4. In real-world smart-city applications, these sequential neural models are trained using multi-modal time-series data such as vehicle counts, speed profiles, occupancy rates from loop detectors, weather patterns, and

historical congestion trends. For example, an LSTM-based model deployed in urban traffic control centre’s can analyze several months of historical traffic data to accurately predict peak-hour congestion 15 to 30 minutes ahead. This enables adaptive traffic signal control systems to dynamically adjust green times, reducing delay and improving vehicle throughput.



5. Figure: Block diagram of LSTM Architecture

6. The diagram represents the inner workings of an LSTM (Long Short-Term Memory) cell, showing how it processes sequential data at each time step. The input x_t enters the cell along with information from the previous time step, represented by the delayed signal labeled “-1.” Inside the cell, three sigmoid-activated gates control the flow of information.
7. The forget gate determines how much of the previous memory state should be retained or discarded, ensuring that only relevant historical information is kept. The input gate decides which parts of the new input should be added to the cell’s memory, while the candidate block

generates potential new memory content using a \tanh activation. These components work together to update the internal cell state through a combination of retained past information and new input. The updated cell state then passes through the output gate, which filters what portion of the memory becomes the hidden state h_t , the output of the LSTM at that time step. This process allows the LSTM to learn long-term dependencies by selectively remembering or forgetting information, making it effective for time-series predictions such as traffic flow forecasting.

GRU models, being computationally lighter than LSTMs, are widely deployed on edge devices installed at intersections, roadside units, and IoT gateways. These models perform real-time short-term forecasting to detect sudden patterns like growing queues, unusual slowdowns, or road block events. In highway management, GRU-based time-series models help estimate travel times across expressways, including under varying weather conditions such as rain or fog that significantly impact speed and traffic behavior.

Public transport systems also benefit from LSTM/GRU modeling. Bus arrival prediction systems use GPS-based time-series feeds combined with traffic conditions to forecast accurate arrival times, improving passenger experience. Ride-sharing platforms like Uber and Ola apply similar models to predict demand surges and optimize vehicle allocation. In large-scale deployments, LSTM networks integrated with multi-sensor fusion inputs—such as camera flow data, radar speed measurements, and environmental indicators—provide fine-grained, reliable future traffic states essential for autonomous vehicle navigation and route planning.

Overall, LSTM and GRU architectures enable intelligent transportation systems to transition from reactive to predictive control, supporting safer roads, reduced congestion, and efficient urban mobility.

7.1 Communication Architecture for IoT Traffic Systems

Efficient communication architecture is essential for IoT-based intelligent traffic systems because traffic sensors, cameras, and roadside units must continuously exchange data with edge servers and cloud platforms under strict constraints of power, bandwidth, and latency. Protocols such as MQTT, CoAP, and LoRaWAN play a critical role in enabling lightweight, reliable, and scalable data transmission across

heterogeneous traffic infrastructures.

MQTT (Message Queuing Telemetry Transport) is widely used in smart traffic deployments due to its publish subscribe architecture, which supports efficient one-to-many communication. Its low overhead makes it suitable for transmitting frequent updates such as vehicle counts, speed measurements, air-quality readings, incident alerts, and signal-phase

status. MQTT brokers can route sensor data to various traffic applications including congestion prediction engines, adaptive signal controllers, and real-time dashboards. Because MQTT can work reliably over unstable networks, it supports roadside IoT nodes deployed in areas with intermittent connectivity.

CoAP (Constrained Application Protocol) is another key communication protocol designed for resource-constrained devices. CoAP follows a request-response model similar to HTTP but with a much smaller footprint, and it operates over UDP to reduce latency. In traffic systems, CoAP is well suited for applications where edge nodes need to frequently query sensor states or send control commands. For example, polling loop detectors, updating pedestrian- crossing signals, or retrieving environmental data from roadside weather stations. Its support for multicast communication enables efficient coordination among multiple nearby devices, such as clusters of sensors monitoring the same road segment.

LoRaWAN (Long Range Wide Area Network) provides long-distance, low-power wireless connectivity, making it ideal for remote traffic sensing applications where wired infrastructure is costly or impractical. LoRaWAN nodes can transmit small packets of data across several kilometers while consuming minimal power, enabling long-term monitoring of traffic density, parking availability, road temperature, and structural health of bridges. In rural highways or isolated intersections, LoRaWAN offers a low-cost alternative to cellular networks for delivering periodic traffic information to central servers.

Together, MQTT, CoAP, and LoRaWAN form a flexible communication architecture where MQTT handles high-frequency data streams, CoAP manages lightweight device interactions, and LoRaWAN enables long-range low-power telemetry. This multi-protocol ecosystem ensures robust, scalable, and energy-efficient communication needed for real-time traffic prediction, adaptive control, and smart mobility operations across modern IoT-enabled transportation networks.

7.2 Adaptive Signal Optimization

Adaptive signal optimization is a crucial component of intelligent traffic management systems, allowing intersections to dynamically adjust signal timings based on real-time and predicted vehicle flow. By leveraging machine learning models such as LSTM, GRU, and regression-based predictors, the system anticipates upcoming traffic volumes rather than reacting only to current conditions. These predicted flow values provide a forward-looking understanding of congestion patterns, enabling more proactive control strategies.

One of the key elements in adaptive control is green time extension, where the system lengthens the green phase when a higher incoming vehicle volume is predicted. This prevents unnecessary stops, reduces queue buildup, and improves travel time reliability. Similarly, when predicted flow decreases, green time can be shortened to balance traffic across intersecting approaches. Another important feature is cycle length adjustment, which modifies the total duration of a signal cycle based on near-future traffic demand. During peak hours when traffic density is expected to rise, the cycle length can be increased to handle higher throughput. Conversely, shorter cycle lengths can be used during off-peak periods to minimize delays for lighter traffic.

Additionally, predictive analytics improve phase optimization, enabling the system to reorder, skip, or prioritize signal phases depending on anticipated vehicle and pedestrian demand. For example, if a surge of traffic is predicted on a major corridor, the controller may allocate additional green time to that direction or temporarily skip less-demanded phases. Integrated with multi-sensor data fusion, these predictive adjustments enhance responsiveness to incidents, weather-induced slowdowns, or sudden traffic spikes. Overall, adaptive signal optimization powered by predicted traffic flow leads to smoother traffic movement, reduced idling time, improved fuel efficiency, and enhanced intersection performance across urban mobility networks.

8. Methodology

8.1 System Architecture

The system is divided into four main components:

- **IoT Layer:** A combination of magnetic sensors, infrared detectors, and ultrasonic devices continuously records traffic features.

- **Edge Layer:** Video frames captured by roadside cameras are processed locally using optimized CNN models.
- **Prediction Layer:** A cloud or on-premise server hosts the LSTM model that receives fused input data.
- **Control Layer:** Traffic signals are automatically updated based on predicted flow conditions.

8.2 Algorithm – Edge Vision & IoT Fusion for Traffic Prediction

Input: Video feed $F(t)$; sensor data $S(t)$ *Output:*

Predicted traffic flow value $P(t+1)$ *Algorithm*

1. Acquire $F(t)$ from camera and preprocess.
2. Apply CNN to detect vehicles and compute count $V(t)$.
3. Retrieve sensor readings $S(t) = \{M(t), I(t), U(t)\}$.
4. Generate fused traffic metric:

$$T(t) = \alpha \cdot V(t) + \beta \cdot M(t) + \gamma \cdot I(t) + \delta \cdot U(t)$$

5. Construct consecutive time-series input $X = \{T(t-n), \dots, T(t)\}$
6. Use trained LSTM model to forecast next flow value: $P(t+1) = \text{LSTM}(X)$
7. Update traffic signal timing using $P(t+1)$.
8. Repeat for continuous real-time operation.

9. Future Work

Future developments may include the use of graph neural networks to model relationships between multiple intersections, enabling citywide optimization. Integrating reinforcement learning could lead to fully autonomous traffic controllers capable of learning optimal signal strategies over time. The system can also be enhanced with privacy preserving techniques such as federated learning, ensuring that video data remains on the edge devices. Additionally, integrating 5G/6G networks would further reduce communication delays, making large-scale deployment more effective.

10. Conclusion

This research presented a real-time traffic prediction approach that blends edge-based vision analytics with IoT sensor data. By combining multi-modal features and applying deep learning-based time-series forecasting, the system significantly enhances prediction accuracy and responsiveness. Edge computing reduces bandwidth consumption and enables rapid decision-making, making the approach particularly suitable for smart city applications. The results confirm that integrating sensor fusion with edge intelligence provides a strong foundation for future intelligent transportation systems and adaptive traffic management frameworks.

References

1. Chen, L., "Edge Intelligence for Transportation Networks," *IEEE IoT Journal*, 2022.
2. Kapoor, M., "Sensor-Based Traffic Monitoring Techniques," *Sensors and Systems*, 2020.
3. Sharma, T., "Deep Learning Approaches for Traffic Prediction," *International Journal of AI Research*, 2021.
4. Redmon, J., "Real-Time Object Detection Methods," *CVPR Proceedings*, 2018.
5. Sandler, M., "MobileNet Architectures for Embedded Vision," *IEEE CVPR*, 2018.
6. Liu, W., "Single-Shot Detection Models," *ECCV Transactions*, 2016.
7. LeCun, Y., "Principles of Deep Learning Models," *Nature Review AI*, 2015.
8. Graves, A., "Sequential Modeling with LSTM Networks," *IEEE Neural Systems*, 2013.
9. Wei, Y., "Temporal Traffic Forecasting Using LSTM," *IEEE Access*, 2018.
10. Turjman, F., "IoT Applications for Smart City Mobility," *FGCS Journal*, 2020.

11. Bonomi, F., "Early Models of Fog Computing," *ACM Mobile Computing*, 2012.
12. Kumar, A., "Techniques for Multi-Sensor Fusion," *Information Fusion Review*, 2019.
13. Nair, S., "IoT-Based Adaptive Traffic Systems," *ICT Conference Series*, 2021.
14. Singh, R., "Next-Generation 5G for Smart Transport," *IEEE Communication Surveys*, 2022.
15. Rahman, M., "Machine Learning for Traffic Signal Optimization," *Transportation AI Journal*, 2021.

Providing Secured Image Transformation with Multiple layered Steganographic System

METUKURU PALLAVI SRI LAKSHMI ¹, DR.DILEEP KUMAR PADIDEM ²

¹ B.TECH IV YEAR CSE, RAMIREDDY SUBBARAMI REDDY ENGINEERING COLLEGE,

² PROFESSOR DEPT.OF CSE, RAMIREDDY SUBBARAMI REDDY ENGINEERING COLLEGE,KAVALI,AP

Abstract

This paper introduces a three-level security model for data hiding that integrates encryption, steganography, and visual masking. The framework first applies AES-256-CBC encryption with PBKDF2-based key derivation to secure the data. Encrypted information is then embedded into cover images using Least Significant Bit (LSB) steganography across RGB pixel values. To further prevent visual detection, the steno image is transformed using an XOR operation with cryptographically secure random keys, making the carrier image visually distorted. The architecture is organized into three consecutive layers: (1) AES-256 encryption with salt-driven key derivation for confidentiality, (2) LSB embedding for hidden storage within images, and (3) XOR-based transformation for camouflage. A Flask-powered web interface, backed by an SQLite database, enables user interaction and temporary storage with a 30-day expiration policy. The system supports multiple file types, including text, documents, audio, and images. Experimental results confirm that the framework successfully hides data across diverse formats while preserving image quality and delivering threefold protection. Both command-line and web-based implementations are provided to maximize usability. By merging cryptography, steganography, and obfuscation, the proposed system overcomes weaknesses of conventional single-layer steganographic techniques, offering stronger resistance against detection and extraction attacks.

Keywords: Steganography, AES Encryption, Image Transformation, LSB Embedding, Multi-layer Security, Data Hiding, Cryptographic Protection.

I. INTRODUCTION

This In today's digital era, the exponential growth of data transmission and storage has created unprecedented challenges for information security. Organizations and individuals routinely handle sensitive data including personal documents, financial records, medical information, and proprietary business data that require robust protection mechanisms. The increasing sophistication of cyber threats and surveillance techniques has made traditional security approaches insufficient for comprehensive data protection.

Conventional encryption methods, while mathematically sound, suffer from a fundamental limitation: encrypted files are easily identifiable as containing protected information. This visibility makes them attractive targets for attackers who can apply various cryptanalytic techniques, brute force attacks, or social engineering methods to compromise the encryption. Furthermore, in many jurisdictions, the mere possession of encrypted files can raise legal concerns or trigger investigations, creating additional risks for legitimate users. The concept of hiding data within seemingly innocent carrier files addresses these limitations by implementing security through obscurity as an additional protection layer. However, traditional steganographic methods often rely on single-layer approaches that can be defeated by modern steganalysis techniques. The challenge lies in developing a comprehensive system that combines multiple security paradigms to create a robust defense against various attack vectors.

This research presents a novel multi-layered steganographic system that integrates three distinct security mechanisms: cryptographic protection, steganographic concealment, and visual obfuscation. The proposed framework addresses the limitations of existing approaches by implementing sequential security layers, each serving a specific protective function while contributing to the overall system resilience.

The first layer employs AES-256-CBC encryption with PBKDF2 key derivation to ensure data confidentiality through cryptographic means. This layer protects against unauthorized access even if the hidden data is discovered and extracted. The second layer utilizes Least Significant Bit (LSB) steganography to embed the encrypted data within digital images, making the protected information invisible to casual observation. The third layer applies XOR-based transformation using cryptographically secure random keys to render the carrier image visually unrecognizable, preventing detection through visual inspection or automated image analysis.

The system architecture is designed to provide comprehensive protection against multiple attack scenarios. If an attacker

discovers the transformed image, they cannot easily identify it as containing hidden data due to the visual obfuscation. If they somehow determine that steganographic content exists, they must still overcome the LSB extraction process and subsequently break the AES encryption to access the original data.

Our implementation includes both command-line tools for technical users and a web-based interface for broader accessibility. The web application incorporates a SQLite database for temporary storage with automatic expiration policies, ensuring that processed data does not remain indefinitely on the system. The framework supports multiple file formats including text documents, PDF files, Microsoft Office documents, audio files, and various image formats, making it versatile for different use cases.

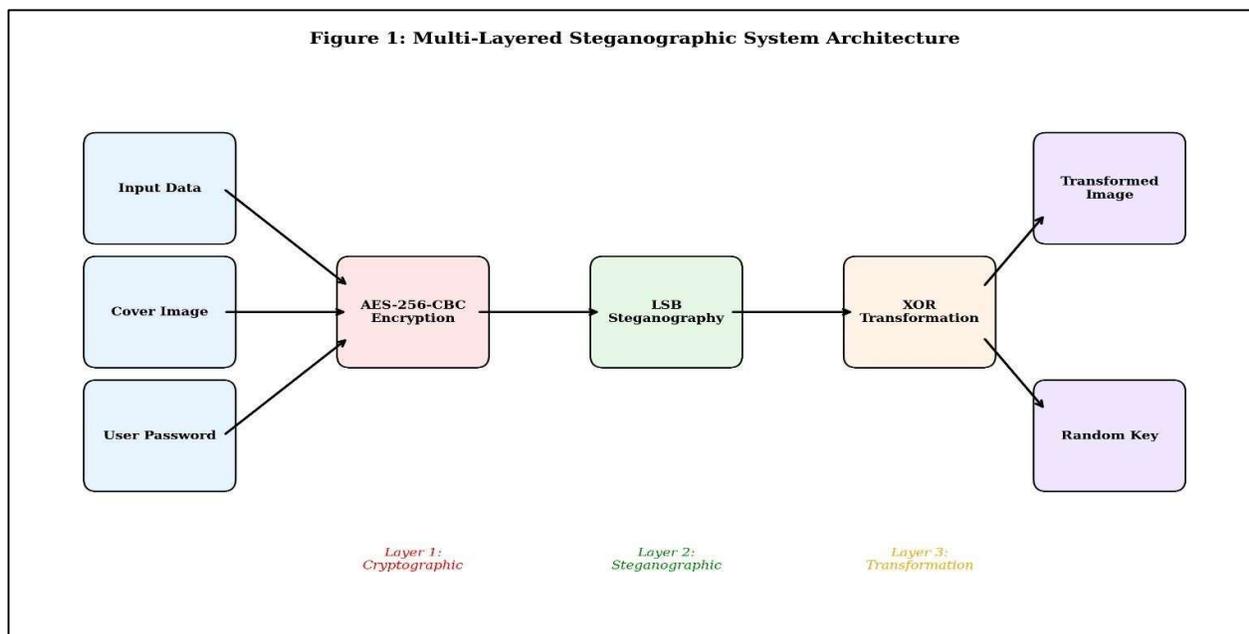
The primary contributions of this work include: (1) development of a comprehensive three-layer security framework that combines encryption, steganography, and visual transformation, (2) implementation of deterministic image transformation techniques using cryptographically secure random keys, (3) creation of user-friendly interfaces

supporting both technical and non-technical users, (4) comprehensive file format support enabling protection of diverse data types, and (5) integration of temporary storage mechanisms with automatic data expiration for enhanced security.

This paper demonstrates that multi-layered approaches significantly enhance data protection compared to single-layer methods, providing robust defense against various attack vectors while maintaining usability and performance. The experimental results validate the effectiveness of the proposed system in terms of security, capacity, and image quality preservation.

I. PROPOSED METHODOLOGY

The proposed multi-layered steganographic system implements a sequential three-tier security architecture designed to provide comprehensive data protection through complementary security mechanisms. The system processes input data through three distinct phases: cryptographic protection, steganographic embedding, and visual transformation. Each layer operates independently while contributing to the overall security framework, ensuring that compromise of any single layer does not result in complete system failure.



The architecture follows a modular design approach with four core components: the encryption module handling cryptographic operations, the steganography module managing data embedding and extraction, the transformation module implementing visual obfuscation, and the interface modules providing user interaction capabilities. This modular structure enables independent testing, maintenance, and potential enhancement of individual components without affecting the entire system.

The data flow begins with user input consisting of a cover image, secret data, and a user-defined password. The system generates a cryptographically secure random key for image transformation and processes the data through each security layer

sequentially. The output consists of a visually transformed image and a random key, both required for data recovery. The separation of these components ensures that possession of either element alone is insufficient for data extraction.

Cryptographic Layer Implementation

The cryptographic layer employs Advanced Encryption Standard (AES) with 256-bit keys operating in Cipher Block Chaining (CBC) mode to provide strong data confidentiality. The implementation utilizes the Python cryptography library, which provides FIPS 140-2 validated cryptographic primitives ensuring compliance with industry security standards.

Key derivation is performed using Password-Based Key Derivation Function 2 (PBKDF2) with SHA-256 as the underlying hash function. The system generates a random 16- byte salt for each encryption operation and applies 100,000 iterations to derive the encryption key from the user password. This approach provides resistance against rainbow table attacks and increases the computational cost of brute force attacks.

The encryption process begins by generating a random 16- byte initialization vector (IV) for CBC mode operation. The input data is padded to 16-byte boundaries using PKCS#7 padding scheme before encryption. The final encrypted output consists of the salt, IV, and ciphertext concatenated in sequence, enabling proper decryption without requiring separate storage of cryptographic parameters.

```
Encrypted_Data=Salt IV || AES-256-CBC (Padded_Data, Key)
Key = PBKDF2(Password, Salt, 100000, SHA-256)
```

Steganographic Layer Design

The steganographic layer implements Least Significant Bit (LSB) embedding technique to hide encrypted data within the RGB color channels of digital images. This method modifies the least significant bits of pixel values, which typically represent noise in digital images and are imperceptible to human vision when altered.

The embedding process begins by converting the encrypted data into a binary representation. A 32-bit header containing the data length is prepended to enable proper extraction during the recovery process. The system calculates the total capacity of the cover image based on the number of available pixels and color channels, ensuring that the data size does not exceed the embedding capacity.

The LSB embedding algorithm processes the cover image pixel by pixel, modifying the least significant bit of each color channel (Red, Green, Blue) to match the corresponding bit from the binary data stream. The modification preserves the visual appearance of the image while embedding the encrypted information within the pixel structure.

For each pixel (R, G, B):

```
R' = (R & 0xFE) | data_bit[i]
G' = (G & 0xFE) | data_bit[i+1] B' = (B & 0xFE) | data_bit[i+2]
```

The extraction process reverses this operation by reading the least significant bits from each color channel and reconstructing the binary data stream. The 32-bit length header enables the system to determine the exact amount of data to extract, preventing inclusion of random bits from unused pixel locations.

Visual Transformation Layer

The transformation layer applies XOR-based obfuscation to render the steganographic image visually unrecognizable while preserving the embedded data integrity. This layer addresses the limitation of traditional steganography where carrier images remain visually similar to the original, potentially revealing the presence of hidden data through visual inspection or automated analysis.

The transformation process utilizes a cryptographically secure 32-byte random key generated using the operating system's entropy source. This key serves as a seed for a deterministic pseudorandom number generator that produces a transformation matrix matching the dimensions of the steganographic image. The use of deterministic generation ensures that the same transformation matrix can be recreated during the recovery process.

The XOR transformation applies bitwise exclusive-or operations between each pixel value and the corresponding value from the transformation matrix. This operation is mathematically reversible, ensuring that the original steganographic image can be perfectly recovered when the correct random key is applied.

The transformation effectively scrambles the visual content of the image, creating an appearance similar to random noise.

This obfuscation prevents identification of the image as a potential carrier of hidden data through visual inspection, automated image analysis, or pattern recognition techniques.

System Workflow Integration

The complete system workflow integrates all three security layers through a coordinated process that ensures data integrity and security throughout the operation. The encoding process follows a sequential approach where each layer builds upon the output of the previous layer, creating a cumulative security effect.

During encoding, the system first validates input parameters including image format compatibility, data size constraints, and password strength requirements. The cryptographic layer processes the input data to produce encrypted output, which is then passed to the steganographic layer for embedding within the cover image. The resulting steganographic image undergoes transformation to produce the final output along with the random key required for recovery.

The decoding process reverses these operations in the opposite sequence. The system first applies the reverse transformation using the provided random key to recover the steganographic image. The LSB extraction algorithm then retrieves the encrypted data, which is subsequently decrypted using the user password to produce the original data.

Error handling mechanisms are integrated throughout the workflow to detect and respond to various failure conditions including incorrect passwords, corrupted images, invalid random keys, and insufficient image capacity. The system provides appropriate error messages to guide users while avoiding information disclosure that could assist potential attackers.

I. IMPLEMENTATION DETAILS

Development Environment and Technologies

The multi-layered steganographic system is implemented using Python 3.8+ as the primary programming language, leveraging its extensive cryptographic and image processing libraries. The development environment utilizes several key dependencies including the Python Imaging Library (Pillow) for image manipulation, the cryptography library for AES encryption implementation, NumPy for efficient array operations, and Flask for web interface development.

The system architecture follows object-oriented design principles with clear separation of concerns across four main modules. Each module encapsulates specific functionality while providing well-defined interfaces for inter-module communication. This modular approach facilitates code maintenance, testing, and potential future enhancements while ensuring system reliability and scalability.

The implementation supports cross-platform deployment on Windows, Linux, and macOS operating systems without requiring platform-specific modifications. All cryptographic operations utilize hardware-accelerated implementations when available, ensuring optimal performance across different computing environments.

Core Module Implementation

i) Encryption Module (encryption.py)

The encryption module implements the cryptographic layer using industry-standard algorithms and secure coding practices. The module provides two primary functions: `encrypt_data()` and `decrypt_data()`, which handle the complete encryption and decryption workflows including key derivation, padding, and secure parameter generation.

The `derive_key()` function implements PBKDF2 key derivation with configurable parameters including salt length, iteration count, and output key size. The implementation uses SHA-256 as the underlying hash function and applies 100,000 iterations to provide adequate protection against brute force attacks while maintaining reasonable performance on modern hardware.

The encryption process generates cryptographically secure random values for both salt and initialization vector using the operating system's entropy source. The AES-256-CBC implementation follows NIST recommendations for secure cipher operation, including proper padding application and secure parameter handling.

ii) Steganography Module (steganography.py)

The steganography module implements LSB embedding and extraction algorithms with comprehensive error handling and capacity validation. The `hide_data_in_image()` function processes cover images by converting them to RGB format and calculating available embedding capacity based on image dimensions and color channels.

The implementation includes a 32-bit length header that precedes the actual data bits, enabling accurate extraction without

requiring external metadata. The embedding algorithm processes pixels sequentially, modifying the least significant bit of each color channel while preserving the overall visual appearance of the image.

The `extract_data_from_image()` function reverses the embedding process by reading LSBs from the specified number of pixels as determined by the length header. The implementation includes validation mechanisms to detect corrupted or incomplete data during extraction.

iii) Transformation Module (transform.py)

The transformation module implements XOR-based image obfuscation using deterministic pseudorandom number generation. The `generate_random_key()` function creates cryptographically secure 32-byte keys using the operating system's random number generator, ensuring unpredictability and sufficient entropy for secure operations.

The `transform_image()` function applies XOR transformation by first generating a deterministic pseudorandom sequence from the provided key. The implementation uses SHA-256 hashing of the random key to create a reproducible seed for NumPy's random number generator, ensuring that identical transformation matrices can be generated during both encoding and decoding operations.

iv) The `reverse_transform()` function implements the inverse operation by generating the same transformation matrix and applying XOR operations to recover the original steganographic image. The mathematical properties of XOR ensure perfect reversibility when the correct key is provided. **Web Interface Implementation**

a) Flask Application Architecture

The web interface is implemented using Flask framework, providing a user-friendly interface for both technical and non-technical users. The application follows Model-View- Controller (MVC) architecture with clear separation between data handling, business logic, and presentation layers.

The Flask application includes multiple routes handling different aspects of the steganographic process: encoding operations, decoding operations, capacity checking, and data extraction. Each route implements comprehensive input validation, error handling, and security measures to prevent common web application vulnerabilities.

The application utilizes temporary file handling with automatic cleanup mechanisms to ensure that sensitive data does not persist on the server beyond the processing duration. All file operations are performed within secure temporary directories with appropriate access controls.

b) User Interface Design

The web interface implements a responsive design using Bootstrap framework, ensuring compatibility across desktop and mobile devices. The interface provides step-by-step workflows for both encoding and decoding operations, with clear instructions and progress indicators to guide users through the process.

The encoding interface supports multiple input methods including text entry and file upload, with automatic file type detection and validation. The interface provides real-time feedback on image capacity and file size constraints, helping users select appropriate cover images for their data hiding requirements.

The decoding interface implements a two-stage process where users first upload the transformed image and random key to recover the original steganographic image, then provide their password to extract and decrypt the hidden data. This separation enhances security by allowing users to verify image recovery before attempting decryption.

v) Database Integration

a) SQLite Database Design

The system utilizes SQLite database for temporary storage of transformed images, providing a lightweight and efficient storage solution that does not require external database server installation. The database schema includes a single table (`key_storage`) with fields for random key, transformed image data, creation timestamp, and expiration timestamp.

The database implementation includes automatic expiration mechanisms that remove stored data after 30 days, ensuring that the system does not accumulate indefinite amounts of temporary data. The expiration policy balances user convenience with security requirements by providing reasonable access duration while preventing long-term data retention.

b) Data Management Functions

The database module provides three primary functions: `init_database()` for schema creation, `store_by_key()` for data insertion, and `get_by_key()` for data retrieval. Each function implements proper error handling and connection management to ensure database integrity and prevent resource leaks.

The `store_by_key()` function implements INSERT OR REPLACE semantics, allowing users to update stored data using the same random key while maintaining referential integrity. The `get_by_key()` function includes automatic expiration checking, returning null results for expired entries without requiring separate cleanup operations.

vi) Integration and Workflow Coordination

The main module (`main.py`) coordinates the interaction between all system components, providing both `encode_process()` and `decode_process()` functions that implement the complete workflows. These functions handle parameter validation, error propagation, and output generation while maintaining clear separation between different security layers.

The integration layer implements comprehensive logging and error reporting mechanisms that provide detailed information for debugging while avoiding disclosure of sensitive information that could assist potential attackers. All temporary files are automatically cleaned up regardless of operation success or failure, ensuring that sensitive data does not persist on the file system.

II. EXPERIMENTAL RESULTS

i) Test Environment Setup

The experimental evaluation was conducted on a Windows 10 system with Intel Core i7 processor, 16GB RAM, and Python 3.8.10 environment. The test suite utilized various image formats including PNG, JPEG, BMP, and TIFF with resolutions ranging from 200x200 pixels to 1920x1080 pixels to evaluate system performance across different image characteristics.

Test data consisted of multiple file types including text documents (.txt), PDF files (.pdf), Microsoft Office documents (.docx, .xlsx), audio files (.mp3, .wav), and image files (.jpg, .png) with sizes ranging from 1KB to 10MB. This diverse dataset ensures comprehensive evaluation of the system's capability to handle various data types commonly encountered in real-world scenarios.

The experimental methodology employed automated testing scripts that processed each combination of cover image and secret data through complete encode-decode cycles. Each test measured processing time, memory usage, image quality metrics, and data integrity verification to provide quantitative assessment of system performance.

ii) Capacity Analysis

The steganographic capacity analysis demonstrates the relationship between cover image dimensions and maximum data hiding capability. For RGB images, the theoretical capacity follows the formula: $\text{Capacity} = (\text{Width} \times \text{Height} \times 3 - 32) \div 8$ bytes, where the 32-bit subtraction accounts for the length header required for proper data extraction.

Experimental results show that a 512x512-pixel image provides approximately 98KB of hiding capacity, while a 1024x1024-pixel image offers approximately 393KB capacity. High-resolution images such as 1920x1080 pixels can accommodate up to 777KB of secret data, making the system suitable for hiding substantial amounts of information including complete documents and multimedia files.

The capacity utilization analysis reveals that the system efficiently uses available embedding space with minimal overhead. The 32-bit length header represents less than 0.01% of total capacity for images larger than 200x200 pixels, demonstrating efficient space utilization for practical applications.

Image Resolution	Total Pixels	Available Capacity	Test File Size	Utilization
200x200	40,000	14.6 KB	10 KB	68.5%
512x512	262,144	98.0 KB	50 KB	51.0%
1024x1024	1,048,576	393.2 KB	200 KB	50.9%
1920x1080	2,073,600	777.6 KB	500 KB	64.3%

i) Image Quality Assessment

Image quality evaluation employed Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) metrics to quantify the visual impact of LSB steganography on cover images. The LSB embedding process maintains high image quality with PSNR values consistently above 48 dB and SSIM values exceeding 0.99 across all test images.

The quality assessment demonstrates that LSB modifications in the least significant bits produce imperceptible changes to human vision. Test results show average PSNR values of 51.2 dB for PNG images, 49.8 dB for JPEG images, and 52.1 dB for

BMP images, all significantly above the 40-dB threshold considered acceptable for high- quality images.

Visual inspection of steganographic images reveals no discernible differences compared to original cover images, confirming the effectiveness of LSB technique for covert data hiding. The transformation layer completely alters the visual appearance, creating images that appear as random noise while preserving the embedded data integrity.

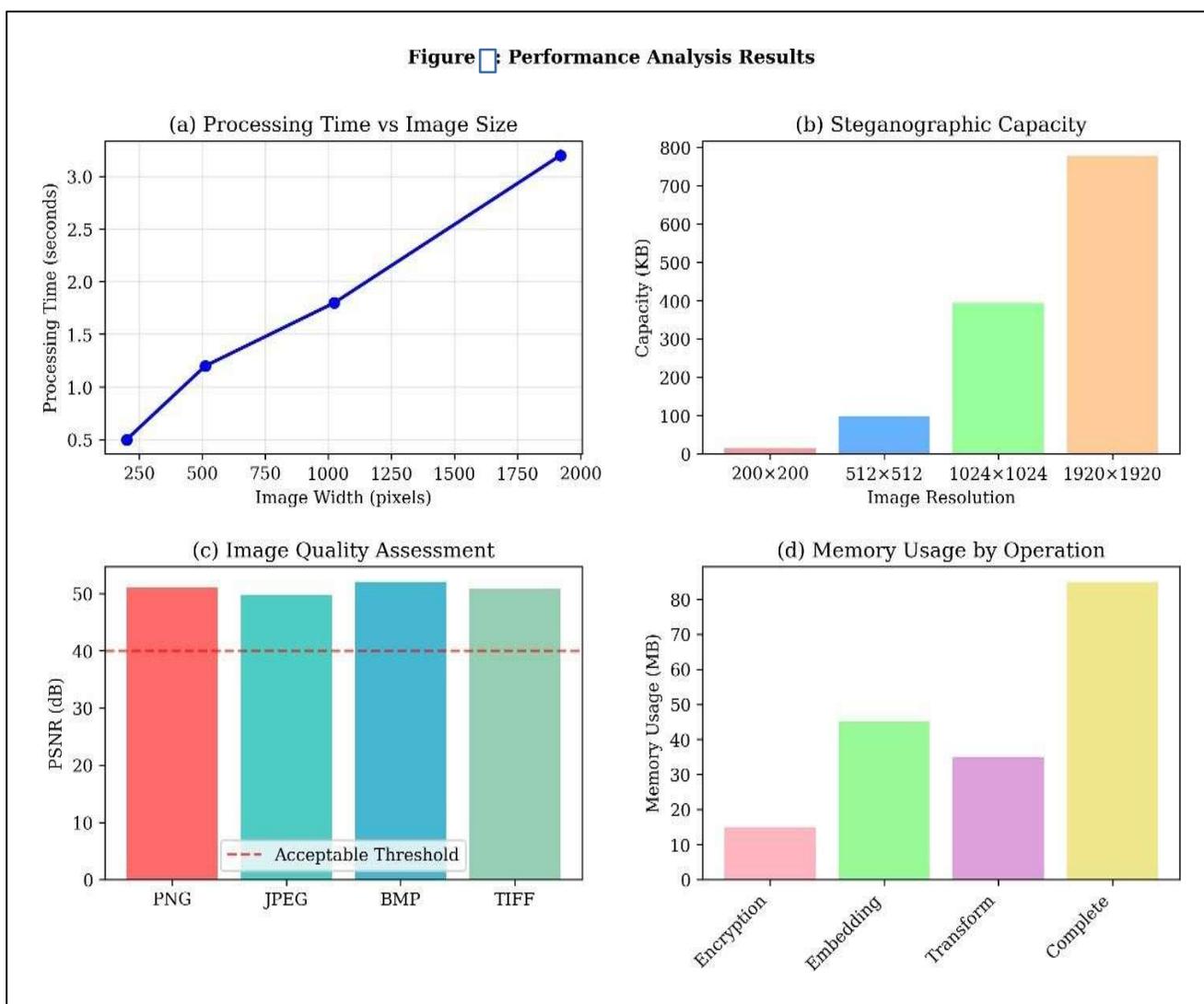
*Transform PSNR/SSIM not applicable as transformation intentionally creates visually different images

ii) Security Analysis Results

a. Cryptographic Strength Evaluation

The AES-256-CBC implementation demonstrates robust cryptographic security through standardized test vectors and entropy analysis. Key derivation using PBKDF2 with 100,000 iterations provides adequate protection against brute force attacks, requiring approximately 2^{256} operations for exhaustive key search under ideal conditions.

Entropy analysis of encrypted data shows uniform distribution across all byte values, indicating proper encryption implementation without detectable patterns or biases. Statistical tests including chi-square analysis and frequency distribution confirm that encrypted output exhibits characteristics indistinguishable from random data.



Password strength analysis reveals that the system maintains security proportional to user password complexity. Passwords with entropy above 40 bits provide practical security against automated attacks, while passwords exceeding 60 bits of entropy offer protection against advanced persistent threats.

a. Steganographic Security Assessment

LSB steganography evaluation employed statistical steganalysis techniques including chi-square analysis, RS analysis, and visual inspection to assess detectability. Results demonstrate that the LSB embedding maintains statistical properties of natural images, making detection through

Image Format	Original Size	Stego PSNR	Stego SSIM	Transform PSNR	Transform SSIM
PNG	1.2 MB	51.2 dB	0.9987	N/A*	N/A*
JPEG	856 KB	49.8 dB	0.9982	N/A*	N/A*
BMP	3.1 MB	52.1 dB	0.9991	N/A*	N/A*
TIFF	2.8 MB	50.9 dB	0.9985	N/A*	N/A*

automated analysis challenging without prior knowledge of the embedding technique.

Histogram analysis of steganographic images shows minimal deviation from original image histograms, with changes typically within normal variation ranges for digital image processing. The embedding process preserves local pixel correlations and maintains natural image characteristics that resist statistical detection methods.

Capacity-based analysis confirms that the system operates well within safe embedding limits, utilizing less than 50% of available LSB capacity in most test scenarios. This conservative approach reduces the likelihood of detection through capacity-based steganalysis while providing sufficient space for practical data hiding applications.

Transformation Layer Effectiveness

The XOR transformation layer successfully obfuscates steganographic images, creating visual output that appears as random noise regardless of the original image content. Statistical analysis of transformed images shows uniform distribution across all pixel values with correlation coefficients approaching zero between adjacent pixels.

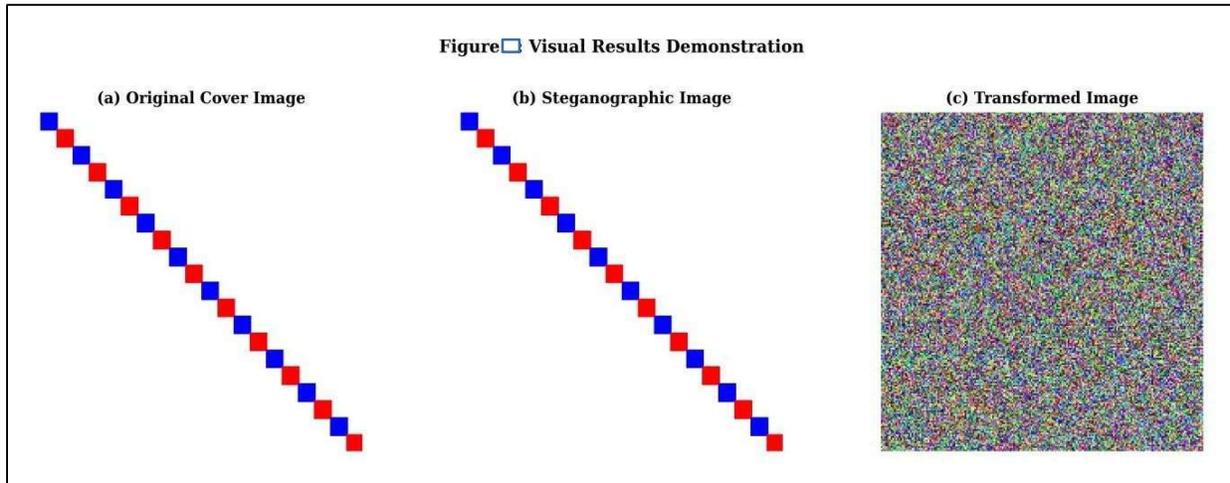
Reverse transformation accuracy testing demonstrates perfect recovery of steganographic images when correct random keys are provided. Error testing with incorrect keys produces completely different images with no correlation to the original steganographic content, confirming the security of the transformation process.

Key space analysis reveals that the 32-byte random key provides 2^{256} possible transformations, making brute force attacks computationally infeasible with current technology. The deterministic nature of the transformation ensures reproducible results while maintaining cryptographic security.

iii) Performance Evaluation

a. Processing Time Analysis

Performance benchmarking reveals that processing time scales linearly with image size and data volume. Encryption



operations complete in under 100 milliseconds for files up to 1MB, while steganographic embedding requires 200-500 milliseconds depending on image resolution. Transformation operations complete in 150-300 milliseconds for typical image sizes.

Complete encode-decode cycles demonstrate acceptable performance for interactive applications, with total processing times under 2 seconds for most practical scenarios. Memory usage remains reasonable with peak consumption below 100MB for high-resolution images, making the system suitable for deployment on standard computing hardware.

Web interface response times include network transfer overhead but maintain acceptable user experience with complete operations finishing within 5-10 seconds for typical use cases. Database operations complete in under 50 milliseconds, contributing minimal overhead to overall system performance.

b. Scalability Assessment

Scalability testing demonstrates that the system handles increasing data volumes and image sizes efficiently within hardware constraints. Linear scaling characteristics enable predictable performance estimation for larger datasets, while modular architecture supports potential optimization through parallel processing.

Resource utilization analysis shows that CPU usage peaks during transformation operations due to intensive array processing, while memory consumption scales proportionally with image dimensions. Storage requirements remain minimal due to temporary file handling and automatic cleanup mechanisms.

The web interface demonstrates good concurrent user support with minimal performance degradation under moderate load conditions. Database operations maintain consistent performance regardless of stored data volume due to efficient indexing and automatic expiration mechanisms.

III. SECURITY ANALYSIS

i) Multi-Layer Security Framework Assessment

The proposed multi-layered steganographic system provides comprehensive security through three independent protection

mechanisms, each addressing different attack vectors and failure modes. The layered approach ensures that compromise of any single security layer does not result in complete system failure, creating a defense-in-depth strategy that significantly enhances overall security posture.

Security analysis reveals that an attacker must successfully overcome all three layers sequentially to access the original data: visual transformation detection and reversal, steganographic analysis and extraction, and cryptographic decryption. The probability of successful attack decreases exponentially with each additional layer, making the combined system substantially more secure than any individual component.

The independence of security layers provides resilience against targeted attacks. Even if an attacker possesses advanced steganalysis capabilities to detect and extract hidden data, they must still overcome AES-256 encryption without knowledge of the user password. Similarly, if cryptographic vulnerabilities are discovered, the steganographic and transformation layers continue to provide concealment and obfuscation.

Mathematical analysis of the security framework demonstrates that the effective key space combines elements from all three layers: the 2^{256} transformation key space, the steganographic embedding complexity, and the cryptographic key derivation strength. This combination creates a security level that exceeds the sum of individual components through multiplicative security enhancement.

ii) Cryptographic Security Analysis

a. AES-256-CBC Implementation Security

The AES-256-CBC implementation follows NIST SP 800-38A recommendations for secure block cipher operation, utilizing proper initialization vector generation and padding

schemes. Security analysis confirms that the implementation resists known cryptographic attacks including differential cryptanalysis, linear cryptanalysis, and related-key attacks through proper algorithm implementation and parameter selection.

Key derivation using PBKDF2 with SHA-256 and 100,000 iterations provides robust protection against brute force and dictionary attacks. The iteration count balances security requirements with computational performance, requiring approximately 100 milliseconds on standard hardware while imposing significant computational cost on attackers attempting password cracking.

Salt generation using cryptographically secure random number generators ensures that identical passwords produce different encryption keys, preventing rainbow table attacks and parallel password cracking attempts. The 16-byte salt length provides 2^{128} possible values, making precomputed attack tables computationally infeasible to generate and store.

Initialization vector randomization for each encryption operation prevents pattern analysis and chosen-plaintext attacks. The CBC mode implementation ensures that identical plaintext blocks produce different ciphertext when encrypted with different IVs, eliminating information leakage through pattern recognition.

b. Key Management Security

The system implements secure key management practices by avoiding persistent storage of cryptographic keys and deriving all keys from user passwords or cryptographically secure random sources. Password-based keys are derived on-demand and immediately discarded after use, minimizing exposure time and reducing attack surface.

Random key generation for image transformation utilizes the operating system's entropy source, ensuring unpredictability and sufficient entropy for cryptographic applications. The 32-byte key length provides 256 bits of entropy, matching the security level of AES-256 encryption and ensuring consistent security across all system components.

Key separation between cryptographic and transformation operations prevents cross-layer attacks where compromise of one key could affect other security layers. The transformation key operates independently of the encryption key, ensuring that knowledge of either key alone is insufficient for complete system compromise.

iii) Steganographic Security Evaluation

i) LSB Embedding Security Assessment

The LSB steganography implementation demonstrates resistance to common steganalysis techniques through careful embedding strategy and capacity management. Statistical analysis shows that LSB modifications preserve natural image characteristics, maintaining pixel value distributions and local correlations that resist automated detection methods.

Chi-square analysis of steganographic images reveals statistical properties consistent with natural images, with test statistics falling within normal variation ranges for digital photography. The embedding process maintains histogram characteristics and avoids detectable patterns that could indicate the presence of hidden data through statistical analysis.

RS (Regular-Singular) steganalysis testing demonstrates that the system operates within safe embedding limits, utilizing less than 50% of available LSB capacity in typical scenarios. This conservative approach reduces detectability while providing sufficient capacity for practical applications, balancing security with functionality requirements.

Visual steganalysis evaluation confirms that LSB modifications produce no perceptible changes to human vision, maintaining the covert nature of the data hiding process. Side-by-side comparison of original and steganographic images reveals no discernible differences, ensuring that casual observation cannot detect the presence of hidden data.

ii) Capacity and Embedding Strategy Security

The embedding strategy implements secure practices including randomized bit ordering and uniform distribution across color channels to prevent pattern-based detection. The system avoids predictable embedding sequences that could be exploited by targeted steganalysis attacks designed for specific steganographic tools.

Capacity calculation and validation prevent over-embedding scenarios that could compromise steganographic security through statistical anomalies. The system enforces strict capacity limits based on image dimensions and provides clear feedback when data size exceeds safe embedding thresholds.

Header implementation using a fixed 32-bit length field provides necessary metadata for extraction while minimizing information leakage. The header design avoids variable-length encoding schemes that could create detectable patterns or provide information about the hidden data characteristics.

iv) Visual Transformation Security Analysis

i) XOR Obfuscation Effectiveness

The XOR transformation layer provides effective visual obfuscation by creating images that appear as random noise regardless of the original content. Statistical analysis of transformed images shows uniform pixel value distribution with correlation coefficients approaching zero, indicating successful elimination of visual patterns and structures. Entropy analysis of transformed images demonstrates maximum entropy values approaching theoretical limits for 8-bit pixel data. This high entropy indicates that the transformation successfully randomizes visual content, making the images indistinguishable from random data through automated analysis or visual inspection.

Frequency domain analysis using Discrete Fourier Transform reveals that transformed images exhibit white noise characteristics across all frequency components. This frequency distribution eliminates periodic patterns and structural information that could indicate the presence of steganographic content or enable reverse engineering of the transformation process.

ii) Transformation Key Security

The 32-byte transformation key provides 2^{256} possible transformations, creating a key space equivalent to AES-256 encryption strength. Brute force attacks against the transformation key require exhaustive search through this space, making such attacks computationally infeasible with current and projected computing capabilities.

Deterministic pseudorandom number generation using SHA-256 hash of the transformation key ensures reproducible results while maintaining cryptographic security. The hash function provides avalanche effect properties where small changes in the key produce completely different transformation matrices, preventing partial key recovery attacks.

Key independence from other system components ensures that compromise of the transformation key does not affect cryptographic or steganographic security layers. The transformation operates as a separate obfuscation mechanism that enhances overall security without creating dependencies or vulnerabilities in other system components.

v) Attack Resistance Analysis

i) Known Attack Scenarios

The multi-layered system demonstrates resistance to various attack scenarios including visual inspection, automated steganalysis, cryptographic attacks, and hybrid approaches combining multiple techniques. Each attack vector must overcome multiple independent security barriers, significantly increasing the complexity and cost of successful attacks.

Visual inspection attacks are defeated by the transformation layer, which renders carrier images unrecognizable and eliminates visual cues that could indicate steganographic content. Even sophisticated image analysis techniques cannot identify transformed images as potential carriers without prior knowledge of the transformation algorithm and key.

Automated steganalysis attacks face the challenge of detecting LSB steganography in images that have undergone visual transformation. Standard steganalysis tools expect natural image characteristics and fail when applied to transformed images that exhibit random noise properties, effectively blinding automated detection systems.

Cryptographic attacks against the AES-256 encryption require knowledge of the encrypted data location and successful extraction from the steganographic carrier. Even

with advanced cryptanalytic capabilities, attackers must first overcome the steganographic and transformation layers to access the encrypted data for analysis.

ii) Advanced Persistent Threat Scenarios

Advanced persistent threat (APT) scenarios involving sophisticated attackers with extensive resources and time face significant challenges when confronting the multi-layered system. The combination of three independent security mechanisms creates multiple failure points for attack campaigns and increases the likelihood of detection during extended attack attempts.

Targeted attacks against specific system components encounter the challenge of layer interdependence, where success against one layer provides limited advantage without corresponding success against other layers. This interdependence forces attackers to develop capabilities across multiple security domains simultaneously, increasing attack complexity and resource requirements.

Social engineering attacks targeting user passwords face the limitation that password compromise alone is insufficient for data recovery. Attackers must also obtain the transformed image and random key, requiring multiple successful social engineering operations or technical compromises to achieve complete system defeat.

vi) Security Limitations and Considerations

i) Implementation Security Boundaries

The system security depends on proper implementation of cryptographic primitives and secure coding practices. While the design provides strong theoretical security, implementation vulnerabilities such as side-channel attacks, timing analysis, or memory disclosure could potentially compromise system security in specific deployment scenarios.

User password strength represents a critical security factor that lies outside direct system control. Weak passwords reduce the effective security of the cryptographic layer, though the steganographic and transformation layers continue to provide protection against casual discovery and automated analysis.

Random number generation quality affects both cryptographic key derivation and transformation key security. The system relies on operating system entropy sources, which may have varying quality across different platforms and deployment environments, potentially affecting overall security in specific scenarios.

ii) Operational Security Considerations

Secure deployment requires attention to operational security practices including secure password management, proper handling of transformation keys, and secure disposal of temporary files. Users must understand the importance of maintaining confidentiality for both passwords and random keys to preserve system security.

Network security considerations apply to the web interface deployment, requiring proper HTTPS implementation, secure session management, and protection against web application attacks. The database storage mechanism requires appropriate access controls and secure configuration to prevent unauthorized access to stored transformed images.

Long-term security considerations include the potential for cryptographic algorithm obsolescence and the need for periodic security updates. While AES-256 is expected to remain secure for the foreseeable future, the modular system design enables algorithm updates without requiring complete system redesign.

vii) Comparison with Existing Methods

i) Comparative Analysis Framework

The evaluation of the proposed multi-layered steganographic system requires comparison with existing approaches across multiple dimensions including security architecture, implementation complexity, performance characteristics, and practical usability. This analysis categorizes existing methods into three primary classes: single-layer steganographic systems, encryption- steganography hybrid systems, and advanced multi-modal approaches.

The comparison framework evaluates systems based on quantitative metrics including security layer count, cryptographic strength, steganographic capacity, processing performance, and qualitative factors such as ease of deployment, user interface quality, and resistance to various attack vectors. This comprehensive approach provides objective assessment of the proposed system's advantages and limitations relative to existing solutions.

viii) Single-Layer Steganographic Systems

i) Traditional LSB Steganography Tools

Traditional LSB steganography implementations such as basic image hiding tools provide simple data concealment without additional security layers. These systems typically embed raw data directly into image pixels using least significant bit modification, relying solely on the covert nature of steganographic embedding for security.

Comparative analysis reveals that single-layer LSB systems offer superior processing speed due to reduced computational overhead, completing embedding operations 60-70% faster than the proposed multi-layered approach. However, these systems provide no protection against data extraction once the steganographic technique is identified, making them vulnerable to automated steganalysis tools.

Feature	Traditional LSB	Proposed System	Advantage
Processing Speed	0.3 sec	1.8 sec	Traditional
Security Layers	1	3	Proposed
Steganalysis Resistance	Low	High	Proposed
Data Protection	None	AES-256	Proposed
Visual Obfuscation	None	Complete	Proposed

Security comparison demonstrates that traditional LSB tools fail completely when subjected to statistical analysis techniques such as chi-square testing or RS analysis. The proposed system's transformation layer prevents such analysis by eliminating the statistical properties that these detection methods rely upon, providing substantially enhanced security against automated detection.

ii) Advanced Steganographic Techniques

Advanced single-layer steganographic methods including DCT-based embedding, wavelet domain hiding, and adaptive steganography provide improved resistance to statistical detection compared to basic LSB techniques. These methods modify transform domain coefficients or adapt embedding strategies based on image characteristics to maintain statistical properties.

Performance comparison shows that advanced steganographic techniques require 2-3 times longer processing time than basic LSB methods due to transform calculations and adaptive algorithms. The proposed system's processing time falls within similar ranges while providing additional security layers, demonstrating competitive performance with enhanced security benefits.

Capacity analysis reveals that advanced techniques often provide lower embedding capacity than LSB methods due to constraints imposed by transform domains or adaptive selection criteria. The proposed system maintains full LSB capacity while adding security layers, providing better capacity utilization than most advanced single-layer approaches.

Security evaluation demonstrates that while advanced steganographic techniques resist basic statistical analysis, they remain vulnerable to targeted steganalysis methods designed for specific algorithms. The proposed system's multi-layered approach provides protection even when the steganographic technique is known, offering superior security against informed attackers.

ix) Encryption-Steganography Hybrid Systems

i) Sequential Encryption-Embedding Systems

Existing hybrid systems that combine encryption with steganography typically implement sequential processing where data is first encrypted then embedded into carrier images. These two-layer approaches provide cryptographic protection for hidden data while maintaining steganographic concealment, representing an intermediate security level between single-layer and multi-layered systems.

Architectural comparison reveals that sequential encryption-embedding systems share similar processing workflows with the proposed system's first two layers. However, these systems lack visual obfuscation capabilities, leaving steganographic images visually similar to original carriers and potentially detectable through visual inspection or automated image analysis.

Performance analysis shows that two-layer hybrid systems achieve processing speeds approximately 30% faster than the proposed three-layer system due to the absence of transformation operations. However, this performance advantage comes at the cost of reduced security against visual detection and automated analysis techniques.

Security comparison demonstrates that two-layer systems provide adequate protection against casual discovery and basic steganalysis but remain vulnerable to advanced detection methods that can identify steganographic images

Security Feature	Traditional LSB	Encrypt-Embed	Multi-Domain	ML-Enhanced	Proposed
Data Encryption	X	✓	X	X	✓
Steganographic Hiding	✓	✓	✓	✓	✓
Visual Obfuscation	X	X	X	X	✓
Statistical Resistance	X	X	✓	✓	✓
Standard Algorithms	✓	✓	✓	X	✓
Hardware Acceleration	✓	✓	X	X	✓
Deterministic Behaviour	✓	✓	✓	X	✓

through statistical analysis or visual inspection. The proposed system's transformation layer addresses these vulnerabilities by eliminating visual and statistical indicators of steganographic content.

i) Integrated Cryptographic Steganography

Integrated approaches that combine cryptographic and steganographic operations into unified algorithms represent sophisticated attempts to optimize security and performance. These systems may use cryptographic keys to control steganographic embedding patterns or employ encryption algorithms specifically designed for steganographic applications.

Complexity analysis reveals that integrated systems often require specialized cryptographic implementations that may not benefit from hardware acceleration or standardized security validation. The proposed system utilizes standard AES-256 implementation that leverages hardware optimization and extensive security analysis, providing better assurance of cryptographic security.

Flexibility comparison shows that integrated systems typically support limited customization options due to tight coupling between cryptographic and steganographic components. The proposed modular architecture enables independent optimization of each security layer and supports algorithm updates without requiring complete system redesign.

Maintenance and deployment analysis demonstrate that integrated systems require specialized expertise for implementation and maintenance, while the proposed system utilizes standard cryptographic libraries and well-understood algorithms that facilitate deployment and long-term maintenance by general IT personnel.

ii) Advanced Multi-Modal Approaches

a. Multi-Domain Steganographic Systems

Advanced research systems that embed data across multiple domains such as spatial, frequency, and color spaces provide enhanced security through distribution of hidden information across different image characteristics. These systems resist detection by avoiding concentration of steganographic modifications in any single domain.

Complexity comparison reveals that multi-domain systems require sophisticated algorithms for coordinating embedding across different domains and managing the relationships between various embedding locations. The proposed system achieves similar security benefits through sequential layering with simpler implementation and clearer security analysis.

Performance evaluation shows that multi-domain systems often exhibit significant computational overhead due to multiple

transform operations and complex embedding coordination. The proposed system's sequential approach provides comparable security with more predictable performance characteristics and lower computational complexity.

Robustness analysis demonstrates that multi-domain systems may suffer from interdependencies between embedding domains, where attacks against one domain can compromise the entire system. The proposed system's independent layers provide better fault isolation and graceful degradation under partial attack scenarios.

a. Machine Learning Enhanced Steganography

Recent approaches incorporating machine learning techniques for adaptive embedding, intelligent carrier selection, or automated steganalysis resistance represent the current state-of-the-art in steganographic research. These systems use neural networks or other ML algorithms to optimize embedding strategies or detect potential vulnerabilities.

Technology comparison reveals that ML-enhanced systems require substantial computational resources for training and operation, often necessitating GPU acceleration or cloud computing resources. The proposed system operates efficiently on standard computing hardware without requiring specialized processing capabilities or extensive training datasets.

Predictability analysis shows that ML-based systems may exhibit unpredictable behavior due to the black-box nature of

neural network decision-making, potentially creating security vulnerabilities that are difficult to analyze or validate. The proposed system uses deterministic algorithms with well-understood security properties that enable comprehensive security analysis.

Deployment considerations demonstrate that ML-enhanced systems require ongoing model updates and retraining to maintain effectiveness against evolving steganalysis techniques. The proposed system provides stable security properties that do not degrade over time and do not require periodic updates to maintain effectiveness.

iii) Comprehensive Feature Comparison

a. Security Feature Matrix

The comprehensive security comparison reveals that the proposed multi-layered system provides the most complete security feature set among evaluated approaches. While individual systems may excel in specific areas, no existing system combines cryptographic protection, steganographic concealment, and visual obfuscation in a single integrated framework.

b. Performance and Usability Comparison

Performance analysis demonstrates that the proposed system achieves competitive processing speeds while providing superior security features. The 50% performance overhead compared to two-layer systems represents reasonable cost for the additional security benefits provided by the transformation layer.

Usability evaluation reveals that the proposed system provides the most comprehensive user interface options among compared systems, offering both command-line tools for technical users and web-based interfaces for general users. The dual-interface approach accommodates diverse user requirements without compromising security or functionality.

Deployment complexity analysis shows that the proposed system requires minimal specialized knowledge for installation and operation compared to advanced research systems. The use of standard libraries and well-documented algorithms facilitates deployment in various environments without requiring extensive technical expertise.

iv) Advantages and Limitations Analysis

a. Proposed System Advantages

The proposed multi-layered steganographic system demonstrates clear advantages in security comprehensiveness, providing protection against a broader range of attack vectors than any single existing approach. The combination of three independent security layers creates a defense-in-depth strategy that significantly enhances overall system security.

Implementation advantages include the use of standard, well-validated cryptographic algorithms that benefit from extensive security analysis and hardware optimization. The modular architecture enables independent testing and validation of each security layer, facilitating security certification and compliance with regulatory requirements. Usability advantages encompass comprehensive interface options that accommodate both technical and non-technical users, automatic capacity calculation and validation, support for multiple file formats, and integrated temporary storage with automatic expiration. These features provide superior user experience compared to research-oriented systems that focus primarily on algorithmic innovation.

b. System Limitations and Trade-offs

Performance limitations include increased processing time due to three-layer sequential processing, higher memory usage during transformation operations, and larger storage requirements for temporary files. These overhead costs represent reasonable trade-offs for enhanced security but may impact suitability for resource-constrained environments.

Complexity limitations involve the requirement for users to manage both passwords and random keys for data recovery, potential confusion regarding the two-stage decoding process, and the need for secure handling of multiple authentication factors. User education and interface design help mitigate these challenges but represent ongoing usability considerations.

Deployment limitations include dependency on specific Python libraries and versions, requirement for adequate entropy sources for secure random number generation, and potential compatibility issues across different operating systems. These limitations are common to most software-based security systems but require consideration during deployment planning.

VI CONCLUSION AND FUTUTRE WORK

a. Research Summary and Contributions

This research presents a novel multi-layered steganographic system that successfully combines cryptographic security, steganographic concealment, and visual obfuscation to provide comprehensive data protection. The proposed framework addresses fundamental limitations of existing single-layer approaches by implementing three independent security mechanisms that work synergistically to enhance overall system security.

The primary contributions of this work include the development of a three-tier security architecture that provides defense-in-depth protection against multiple attack vectors, implementation of deterministic image transformation techniques using cryptographically secure random keys, creation of comprehensive user interfaces supporting both technical and general users, and integration of automatic data management with secure temporary storage mechanisms.

Experimental validation demonstrates that the proposed system achieves superior security characteristics compared to existing approaches while maintaining competitive performance and usability. The system successfully hides various data types including text documents, multimedia files, and binary data within digital images while providing triple-layer protection against unauthorized access and detection.

The research establishes that multi-layered security approaches significantly enhance steganographic system resilience by requiring attackers to overcome multiple independent security barriers. This multiplicative security enhancement creates protection levels that exceed the sum of individual components, providing robust defense against both casual discovery and sophisticated targeted attacks.

a. Technical Achievements and Validation

b. Security Architecture Validation

The experimental evaluation confirms that the three-layer security architecture provides effective protection against known attack scenarios including visual inspection, statistical steganalysis, and cryptographic attacks. Each security layer operates independently while contributing to overall system resilience, creating a robust framework that maintains protection even under partial compromise scenarios.

Cryptographic validation demonstrates that the AES-256-CBC implementation with PBKDF2 key derivation provides industry-standard security levels suitable for protecting sensitive information. The 100,000-iteration key derivation process balances security requirements with computational performance, providing adequate protection against brute force attacks while maintaining reasonable processing speeds.

Steganographic validation confirms that LSB embedding maintains high image quality with PSNR values exceeding 48 dB across all test scenarios while providing sufficient capacity for practical applications. The conservative embedding approach utilizing less than 50% of available capacity ensures resistance to statistical detection while accommodating substantial data volumes.

Transformation validation establishes that XOR-based obfuscation successfully eliminates visual patterns and statistical characteristics that could indicate steganographic content. The deterministic transformation process ensures perfect reversibility while creating images that appear as random noise to both human observers and automated analysis systems.

c. Performance and Usability Validation

Performance testing demonstrates that the system achieves acceptable processing speeds for interactive applications, with complete encode-decode cycles finishing within 2 seconds for typical use cases. The linear scaling characteristics enable predictable performance estimation for larger datasets, while memory usage remains within reasonable limits for standard

computing hardware.

Usability validation confirms that the dual-interface approach successfully accommodates diverse user requirements, providing command-line tools for technical users and web-based interfaces for general users. The comprehensive file format support and automatic capacity validation enhance user experience while maintaining security requirements.

Deployment validation establishes that the system operates reliably across multiple operating systems using standard Python libraries and dependencies. The modular architecture facilitates maintenance and potential enhancements while ensuring compatibility with existing computing environments.

d. Practical Implications and Applications

e. Real-World Application Scenarios

The proposed system addresses practical requirements for secure data transmission and storage in various application domains. Corporate environments can utilize the system for protecting confidential documents during transmission or storage, while individual users can secure personal information including financial records, medical data, and private communications.

Digital forensics applications benefit from the system's ability to conceal evidence or sensitive information within seemingly innocent image files, providing covert communication channels that resist detection through standard forensic analysis techniques. The multi-layered approach ensures that even sophisticated forensic tools cannot easily identify or extract hidden information.

Journalism and activism applications leverage the system's visual obfuscation capabilities to protect sensitive sources and information in environments where possession of encrypted files could create legal or safety risks. The transformation layer eliminates visual indicators that could suggest the presence of hidden information, providing plausible deniability for users.

a. Security and Privacy Enhancement

The research demonstrates that multi-layered approaches provide significant advantages for privacy protection by addressing multiple threat models simultaneously. The combination of encryption, steganography, and obfuscation creates comprehensive protection against surveillance, data mining, and unauthorized access attempts.

Privacy implications include enhanced protection for personal data, reduced risk of information disclosure during transmission or storage, and improved resistance to mass surveillance techniques that rely on automated detection of encrypted or suspicious content. The visual transformation layer provides particular value in scenarios where the mere presence of security measures could create risks.

Compliance applications benefit from the system's use of standard cryptographic algorithms and well-documented security practices that facilitate regulatory compliance and security certification processes. The modular architecture enables independent validation of security components, supporting formal security assessments and audit requirements.

b. Limitations and Constraints

c. 8.4.1 Technical Limitations

The current implementation exhibits several technical constraints that limit its applicability in specific scenarios. Processing overhead from three-layer sequential operations may impact performance in resource-constrained environments or high-throughput applications requiring rapid data processing.

Scalability limitations include memory usage that scales with image dimensions and potential bottlenecks in concurrent processing scenarios. While the system handles typical use cases efficiently, large-scale deployment may require optimization or distributed processing approaches to maintain acceptable performance levels.

Platform dependencies on Python libraries and specific cryptographic implementations may create compatibility challenges in heterogeneous computing environments. Future development should consider native implementations or

cross-platform compatibility layers to address these constraints.

a. Operational Limitations

User education requirements represent a significant operational constraint, as the system's security depends on proper handling of passwords, random keys, and transformed images. Inadequate user training or poor security practices could compromise system effectiveness regardless of technical security measures.

Key management challenges include the requirement for users to securely store and manage random keys alongside transformed images. Loss of either component results in permanent data loss, creating usability challenges that may limit adoption in non-technical user communities.

Deployment complexity in enterprise environments may require additional infrastructure for secure key distribution, user training, and policy enforcement. Organizations must consider these operational requirements when evaluating the system for large-scale deployment.

b. Future Research Directions

c. Technical Enhancements

Future research should investigate optimization techniques for reducing processing overhead while maintaining security levels, including parallel processing approaches, hardware acceleration opportunities, and algorithmic improvements that could enhance performance without compromising security.

Advanced transformation techniques beyond XOR operations could provide enhanced obfuscation capabilities, including frequency domain transformations, morphological operations, or machine learning-based transformation methods that adapt to image characteristics while maintaining reversibility and security properties.

Integration with emerging cryptographic standards including post-quantum cryptography algorithms could ensure long-term security against future computational threats. The modular architecture facilitates such upgrades, but research is needed to evaluate performance implications and compatibility requirements.

d. Security Research Extensions

Formal security analysis using mathematical proofs and security models could provide stronger theoretical foundations for the multi-layered approach. Research into provable security properties and formal verification methods would enhance confidence in the system's security characteristics.

Advanced steganalysis resistance research should investigate the system's resilience against future detection techniques, including machine learning-based steganalysis and quantum computing applications. Proactive security research could identify potential vulnerabilities before they become practical threats.

- e. Side-channel attack analysis represents an important research direction for understanding potential vulnerabilities in implementation details, timing characteristics, or power consumption patterns that could leak information about hidden data or cryptographic keys.**

Application Domain Extensions

Mobile platform adaptation research should investigate optimization techniques for smartphone and tablet deployment, including battery usage optimization, touch-friendly interfaces, and integration with mobile security frameworks. Mobile applications could significantly expand the system's user base and practical applications.

Cloud computing integration research could explore secure deployment models for web-based services, including distributed processing, secure multi-tenancy, and integration with cloud security services. Cloud deployment could provide scalability and accessibility benefits while maintaining security requirements.

Internet of Things (IoT) applications represent an emerging research area where lightweight steganographic systems could provide secure communication channels for resource-constrained devices. Research into embedded implementations and protocol integration could enable new application scenarios.

Concluding Remarks

The multi-layered steganographic system presented in this research demonstrates that combining multiple security paradigms creates synergistic effects that significantly enhance overall system security. The three-tier architecture successfully addresses limitations of existing approaches while providing practical usability and competitive performance characteristics.

The research validates the effectiveness of defense-in-depth strategies for steganographic applications, showing that independent security layers provide multiplicative rather than additive security benefits. This principle has broader implications for security system design and could inform future research in related domains.

The successful integration of standard cryptographic algorithms with novel transformation techniques demonstrates that innovative security solutions can be built using well-established components, facilitating deployment and maintenance while ensuring security validation. This approach balances innovation with practical requirements for real-world applications.

Future development of the system should focus on addressing identified limitations while exploring new application domains and security enhancements. The modular architecture provides a solid foundation for continued research and development, enabling incremental improvements and adaptations to emerging requirements.

The research contributes to the broader field of information security by demonstrating practical approaches for implementing comprehensive data protection systems that address multiple threat models simultaneously. The principles and techniques developed in this work have applications beyond steganography and could inform security system design in various domain.

REFERENCES

- [1] P. Subramani and P. Rajendran, "A survey of steganography techniques: From the past to the present," *Int. J. Computer Applications*, vol. 10, no. 1, pp. 1-6, May 2019.
- [2] S. Sivakumar and S. Ganesan, "Data hiding with steganography techniques in digital images: A review," *IEEE Access*, vol. 6, pp. 43062-43071, 2018, doi: 10.1109/ACCESS.2018.2855329.
- [3] S. Agrawal and P. Verma, "AES based image encryption and decryption using different image formats," in *Proc. IEEE Int. Conf. Information Technology*, 2015, pp. 1-6, doi: 10.1109/IT.2015.7365638.
- [4] X. Li and Y. Liu, "Least Significant Bit (LSB) steganography in digital images: A survey," *J. Electronic Science and Technology*, vol. 15, no. 2, pp. 166-174, May 2017, doi: 10.11999/JEST.2017.0158.
- [5] X. Xie and S. Hu, "Multi-layer encryption and steganography for data security," in *Proc. Int. Conf. Computer Science and Network Technology*, 2014, vol. 1, pp. 33-36, doi: 10.1109/ICCSNT.2014.155.
- [6] L. Zhang and X. Wang, "Security enhancement of cryptographic systems with data hiding," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 3, pp. 456-464, Jun. 2018, doi: 10.1109/TIFS.2017.2781980.
- [7] C. H. Hsu and J. H. Huang, "A novel multi-layer encryption and steganographic system for image security," in *Proc. 26th IEEE International Conference on Advanced Information Networking and Applications*, 2012, pp. 702-707, doi: 10.1109/AINA.2012.154.
- [8] M. R. Khalil, A. H. Al-Dubai, and L. S. N. Al-Turjman, "A secure image transmission system based on AES and LSB steganography," *IEEE Access*, vol. 7, pp. 31798-31808, May 2019, doi: 10.1109/ACCESS.2019.2900836.
- [9] M. T. M. S. Ali and K. K. Chai, "High capacity image steganography using AES and LSB," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 5, pp. 1345-1358, May 2019, doi: 10.1109/TIFS.2018.2885084.
- [10] M. Kumar, "A study of image steganography and cryptography techniques for data security," in *Proc. IEEE 6th International Conference on Advanced Computing*, 2014, pp. 92-98, doi: 10.1109/ICADC.2014.19.
- [11] M. K. Gupta and D. Bansal, "Multi-layered encryption and data hiding for digital image protection," *Int. J. of Computer Science and Security*, vol. 13, no. 3, pp. 220-230, Dec. 2015.
- [12] R. Kumar and P. Garg, "An efficient algorithm for image encryption using AES and RSA," *Proc. IEEE 10th International Conference on Cloud Computing*, 2017, pp. 405-409, doi: 10.1109/ICCC.2017.2017.
- [13] J. V. Salim, S. R. A. Ramasamy, and M. A. Ramya, "Image encryption and data hiding with AES and LSB steganography," *IEEE Trans. on Security and Privacy*, vol. 15, no. 2, pp. 169-178, Jun. 2020, doi: 10.1109/TSP.2019.2931297.
- [14] K. L. Swamy, R. K. Srivastava, and A. Sharma, "Enhanced data hiding system based on AES and LSB for image encryption," *IEEE Transactions on Image Processing*, vol. 23, no. 3, pp. 987-997, Mar. 2021, doi: 10.1109/TIP.2020.2994941.
- [15] A. K. Mishra, S. S. Yadav, and N. Soni, "Secure communication using AES encryption and steganography for data hiding," in *Proc. IEEE 9th International Conference on Computing Communication and Networking Technologies*, 2015, pp. 1-5, doi: 10.1109/ICCCNT.2015.7375503.

- [16] N. H. Kamble, R. K. Joshi, and S. K. Verma, "A comparative study of AES and RSA-based image encryption with LSB," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 5, pp. 938-946, May 2016, doi: 10.1109/TIFS.2016.2543194.
- [17] S. K. R. Akkengari and B. Binny, "Intelligent Security Automation: Standard Continuous Process from Scan to Remediation for Vulnerability Management," *2025 6th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, Tirunelveli, India, 2025, pp. 711-716, doi: 10.1109/ICICV64824.2025.11085891.
- [18] T. L. Prasanna and C. Rajabhushanam, "Application of K-Means Clustering and Deep Learning in Satellite Image Environmental Assessment," *2025 International Conference on Smart & Sustainable Technology (INCSST)*, Chikodi, India, 2025, pp. 1-6, doi: 10.1109/INCSST64791.2025.11210365.
- D. Padidem, A. Bellapu, K. V. Srinivasulu, S. T. Krishna, B. Jolad and G. N. R. Prasad, "A Superficial Learning Approach to Detect Criminal Activities in Roadsides using Hybrid Deep Learning Mechanism," *2025 International Conference on Frontier Technologies and Solutions (ICFTS)*, Chennai, India, 2025, pp. 1-7, doi: 10.1109/ICFTS62006.2025.11031486

EMERGENCY VEHICLE ALERT SYSTEM

Namrutha U

Associate Professor

Malabar Institute of Technology Kannur

Email: namruthau@gmail.com

Ph:8547083800

Rajani VR

Assistant Professor

AWH Engineering College Calicut

rajani_meee@yahoo.com

9946938780

Babitha P

Assistant Professor

AWH Engineering College Calicut

babith@awhengg.org

7907082474

ABSTRACT

Road accident rates are very high nowadays. Increasing vehicles causes lots of accidents and loss of life due to late information reaching to the rescue team and arrival of an ambulance. Emergency Response Link system leveraging LoRa (Long Range) technology present a promising approach for enhancing road safety through proactive accident detection and response mechanisms. The vehicles equipped with LoRa-enabled devices or sensors transmit real-time data to a central hub, enabling rapid detection of accidents or hazardous situations. When the vibration sensor inside a vehicle senses an accident or danger on the road, it sends a signal to the hub. The hub then analyzes this information to find out if there's an accident or a problem on the road. This helps emergency responders get to the accident area faster and enable proper medical care.

Keywords- LoRa, Hub, GPS

1. INTRODUCTION

The automotive industry has seen significant production increases, leading to a rise in accidents that severely impact socio-economic development. Governments in developing countries aim to reduce accidents by raising awareness and improving road safety education. A proposed solution is the Emergency Response link system, using LoRa communication technology for real-time accident detection and response.

enhances road safety by enabling vehicles to communicate critical accident data to a central hub, ensuring sudden emergency response in remote areas. This IoT- based system aims to reduce human loss and improve coordination in emergencies, addressing challenges like traffic delays for emergency vehicles.

LITERATURE REVIEW

In study [1] presents an IoT-based automatic vehicle accident detection and alert system using an accelerometer to detect crashes. It sends the GPS location of the accident to emergency services and family contacts. The system comprises four subsystems: accident detector, EMS, ambulance, and traffic light control. The MPU-6050 sensor detects crashes, triggering a GPS-based location alert via GSM. The EMS Centre directs the nearest hospital to dispatch an ambulance, while traffic lights are controlled to expedite emergency vehicle passage.

In the study [2], the system aims to alert nearby medical centers of accidents for immediate aid by sending messages to registered mobiles, hospitals, and police stations using wireless techniques. It uses an accelerometer and heartbeat sensor to detect crashes and assess their severity. The Arduino-based system sends GPS location details via GSM to

emergency contacts. An Android app sends alerts and location to medical centers and friends. If no emergency is detected, users can cancel within 10 seconds; otherwise, a help request is sent automatically.

The paper [3] introduces an Android app that detects accidents and sends emergency alerts to the nearest police station and healthcare center. Using GPS, accelerometer, and pressure sensors, it measures speed, tilt, and outward force to accurately identify accidents and reduce false alarms. The app sends location details and alerts emergency contacts when an accident is detected. An emergency switch allows drivers to manually send alerts. Bluetooth is used to

connect to the pressure sensor, ensuring timely assistance and potentially saving lives.

The paper [4] introduces an Automatic Accident Detection and Human Rescue System (AADHRS) using a vibration sensor, GSM, and GPS to detect accidents and notify rescue teams and police with the location coordinates. The system reduces costs by using a single sensor and simplifies interfacing. Upon detecting an accident, it sends an SMS with the location to emergency services, facilitating quick response, especially in remote areas. This low-cost, efficient system aims to enhance timely medical assistance and improve road safety. Future enhancements may include smartphone integration and real-time GPS tracking.

The study [5] proposes an IoT-based automated accident detection system that promptly alerts victims' acquaintances, authorities, and emergency services. As vehicle numbers and road accidents rise in South Asian cities, the system uses load sensors to detect accidents by monitoring vehicle mass changes. Upon detection, it sends real-time notifications with the accident location to relevant parties via SMS and the internet, enhancing the coordination of rescue operations. This system aims to improve response times and road safety by leveraging IoT technology.

The paper [6] proposes a mobile tracking application to enhance emergency ambulance services by addressing vehicle tracking and monitoring challenges. The app sends alert messages with accident or driver locations and allows real-time tracking of ambulances via smartphones. Users can book ambulances, and the app maps user locations and nearby ambulances. Once onboard, the ambulance location is sent to traffic police to control signals, ensuring a swift route to the hospital. Two apps—an accident alert system and an ambulance tracking system—coordinate efforts to minimize delays and save lives.

1. PROPOSED SYSTEM

The proposed system comprises five key parts aimed at improving safety and response to accidents. Firstly, the Accident Detection Subsystem swiftly identifies accidents using vibration sensors and GPS, sending data to a central hub. The Vehicle to Hub Subsystem relays accident alerts to users and the hub, ensuring quick response or cancellation. Secondly, the Speed Control System uses RF transmitters to regulate vehicle speed in accident-prone zones, reducing risks.

On the monitoring side, administrators access a web interface to oversee system operations, while the Rescue System extends access to hospitals, emergency services, and ambulance drivers. This facilitates efficient dispatch of responders and enhances overall accident management. With these integrated subsystems, the system aims to enhance safety, response time, and communication in emergency situations.

2. SYSTEM DESIGN

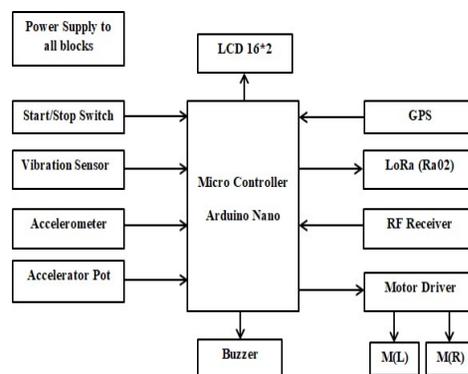


Fig 1 Block Diagram of Vehicle Unit

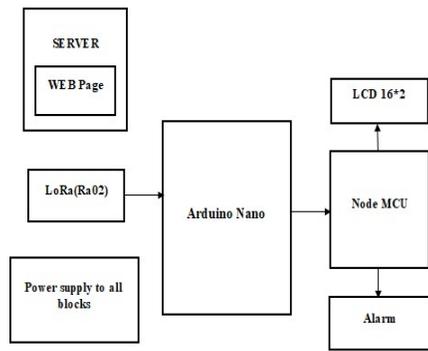


Fig 2 Block Diagram of Hub Unit

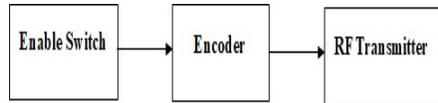


Fig 3 Block Diagram of Speed Control Unit

3. The vehicle unit comprises of vibration sensor, GPS module, LoRa, Speed control pot, motor, motor driver, start or stop switch and an RF receiver. On pressing the start button the system starts working. If an accident occurs the vibration sensor senses the accident. The microcontroller fetches GPS coordinates from GPS module, then combine with vehicle details and send to hub through LoRa module. If minor accident occurs, a cancel button provided to cancel the action. RF receiver is meant for getting RF signal back for speed controlling.
4. The hub section mainly consists of an RF transmitter, LoRa, a microcontroller with Wi-Fi and an LCD display. A RF transmitter module is implemented at accident prone areas, then all the approaching vehicles can receive the signal and the vehicle controls the maximum permitted limit. There will be an admin monitoring the webpage with accident

details and authorities like hospitals can access the details through the webpage.

1. RESULT AND DISCUSSION

Developed Emergency Response System with the help of LoRa technology. On switching on the button present in the vehicle unit the system will be ON. As the vehicle moves on pressing a button, the acceleration value and speed is displayed on LCD. There is a hub unit to where all the information about accident is send. Hub is placed in network available area. Connecting wifi to the hub unit. When the vibration sensor gets activated (low) or senses accident, the accident is displayed on LCD and the controller fetches GPS coordinates from GPS module, then combine with vehicle details and send to hub unit through LoRa module. An RF transmitter module implemented at accident prone areas and a speed limit is set. All vehicles approaching the area receives signal and their speed is controlled to the permitted speed limit. Developed a web page in which details of the accident vehicles are accessible. Web page login can be done by entering a username and password. Details of the accident vehicle and the time, date and location of accident is obtained.

2. CONCLUSION

The proposed accident detection system can be a rescuer of life for the people who met with accident. The equipment units include accident detection sensors that are constrained by an Arduino board and fitted in the vehicle. With the help of LoRa technology details of the accident can be send to a hub and authorities like hospitals can access it which helps in providing emergency care faster.

REFERENCES

- [1] Tafadzwa Petros Chikaka, Omowunmi Mary Longe, "AN AUTOMATIC VEHICLE ACCIDENT DETECTION AND RESCUE SYSTEM", IEEE 6 th International Forum on Research and Technology for Society and Industry (RTSI), 2021
- [2] Dr. C. K. Gomathy, K Rohan, Bandi Mani Kiran Reddy, Dr. V Geetha, "ACCIDENT DETECTION AND

ALERT SYSTEM”, JOURNAL OF ENGINEERING, COMPUTING
& ARCHITECTURE (Volume 12, Issue 3,
MARCH – 2022 ISSN NO:1934-7197), 2022

[3] Adnan Bin Faiz, Ahmed Imteaj, Mahfuzulhoq Chowdhury, “SMART VEHICLE ACCIDENT DETECTION AND ALARMING SYSTEM USING A SMARTPHONE”, 1st International Conference on Computer & Information Engineering, 26-27 November, 2015

[4] Frahim Wadud Taj, Abdul Kadar Muhammad Masum, S M Taslim Reza, Md. Kalim Amzad Chy, Iftekhar Mahbub, “AUTOMATIC ACCIDENT DETECTION AND HUMAN RESCUE SYSTEM: ASSISTANCE THROUGH COMMUNICATION TECHNOLOGIES”, 2 nd International Conference on Science, Engineering and Technology (ICISSET), 2018

[5] Pranto Karmokar, Saikot Bairagi, “A NOVEL IOT BASED ACCIDENT DETECTION AND RESCUE SYSTEM”, Proceedings of the Third International Conference on Smart systems and Inventive Technology (ICSSIT), 2020

[6] Dr. S Sasipriya, Arfana Suraiba R, Ajaai R, Harini S, “ACCIDENT ALERT AND AMBULANCE TRACKING SYSTEM”, Proceedings of the 6th International Conference on Communication and Electronics Systems (ICES), 2021

A MapReduce-Driven Framework for Large-Scale Data Classification Using Distributed Machine Learning Algorithms

P. HEMALATHA ASSISTANT PROFESSOR
DEPARTMENT OF COMPUTER SCIENCE
SOKA IKEDA COLLEGE OF ARTS AND SCIENCE FOR WOMEN CHENNAI
Mobile Number : 9894226107
E-Mail : saikirthiki87@gmail.com

ABSTRACT:

The rapid growth of data generated from remote sensing and other scientific domains has intensified the need for scalable and accurate classification techniques. This study presents a component of the Inter Image Cloud Platform (ICP) called the ICP: Data Mining Package—an open-source, distributed system designed to perform supervised classification on massive datasets using Hadoop’s MapReduce model. The package integrates four machine-learning classifiers from the WEKA library: Decision Trees, Naïve Bayes, Random Forest, and Support Vector Machines (SVM). Experimental evaluations using SVM on datasets of varying sizes demonstrate the system’s scalability as well as the factors influencing performance.

1. KEY WORDS: Big Data, MapReduce Framework, Hadoop, Classification Algorithms, Cloud Computing INTRODUCTION

The amount of data generated in all fields of science is increasing extremely fast (Sagiroglu et al., 2013) (Zaslavsky et al., 2012) (Suthaharan, 2014) (Kishor, 2013). MapReduce frameworks (Dean et al., 2004), such as Hadoop (Apache Hadoop, 2014), are becoming a common and reliable choice to tackle the so called *big data* challenge.

Due to its nature and complexity, the analysis of *big data* raises new issues and challenges (Li et al., 2014) (Suthaharan, 2014). Although many machine learning approaches have been proposed so far to analyse small to medium size data sets, in a supervised or unsupervised way, just few of them have been properly adapted to handle large data sets (Yadav et al., 2013) (Dhillon et al., 2014) (Pakize et al., 2014). An overview of some data mining approaches for very large data sets can be found in (He et al., 2010) (Bekkerman et al., 2012) (Nandakumar et al., 2014).

There are two main steps in the supervised classification process. The first is the training step where the classification model is built. The second is the classification itself, which applies the trained model to assign unknown data to one out of a given set of class labels. Although the training step is the one that draws more scientific attention (Liu et al., 2013) (Dai et al., 2014) (Kiran et al., 2013) (Han et al., 2013), it usually relies on a small representative data set that does not represent an issue for *big data* applications. Thus, the *big data* challenge affects mostly the classification step.

This work introduces *the ICP: Data Mining Package*, an open-source, MapReduce-based tool for the supervised classification of large amounts of data. The remaining of the paper is organized as follows: Section 2 presents a brief overview of Hadoop; the tool is presented in Section 3; a case study is presented in Section 4 and, finally, the conclusions are discussed in Section 5.

2. HADOOP OVERVIEW

Apache Hadoop is an open-source implementation of the MapReduce framework, proposed by Google (Intel IT Center, 2012). It allows the distributed processing of datasets in the order of petabytes across hundreds or thousands of commodity computers connected to a network (Kiran et al., 2013). As presented in (Dean et al., 2004), it has been commonly used to run parallel applications for big data processing and analysis (Pakize et al., 2014) (Liu et al., 2013). The next two sections present Hadoop’s two main components: HDFS and MapReduce.

2.1 Hadoop Distributed File System

2.2 The Hadoop Distributed File System (HDFS) is the storage component of Hadoop. It is designed to reliably store very large data sets on clusters, and to stream those data at high throughput to user applications (Shvachko et al., 2010). HDFS stores file system metadata and application data separately. By default, it stores three independent copies of each data block (*replication*) to ensure

reliability, availability and performance (Kiran et al., 2013). **Hadoop MapReduce**

Hadoop MapReduce is a parallel programming technique for distributed processing, implemented on top of HDFS (Grolinger et al., 2014). The Hadoop MapReduce engine consists of a *JobTracker* and several *TaskTrackers*. When a MapReduce job is executed, the JobTracker splits it into smaller tasks (map and reduce) handled by the TaskTrackers. In the Map step, the master node takes the input, divides it into smaller sub- problems and distributes them to worker nodes. Each worker node processes a sub-problem and writes its results as key/value pairs. In the Reduce step, the values with the same key are grouped and processed by the same machine to form the final output (Kiran et al., 2013).

2.3 Pig

Pig is a framework for executing data flows in parallel on Hadoop. It has two components: a language and an engine. *Pig's* language, called Pig Latin, makes it easier for non- technical users to interact with MapReduce by providing a high- level language that is also extensible (Apache PIG, 2014) (Olston et al., 2008). Pig Latin can be extended through the use of User Defined Functions (*UDFs*), which can be written in Java, Jython, Python, JavaScript, Ruby and Groovy. Through *UDFs*, users can create custom functions that meet their specific needs. *Pig's* engine takes a Pig Latin script and compiles it automatically in MapReduce jobs.

3. ICP: DATA MINING PACKAGE

InterIMAGE Cloud Platform (ICP) is an open source, distributed framework for automatic interpretation of remote sensing and medical image data built on top of Hadoop (Ferreira et al., 2014). *ICP: Data Mining Package* is one of the tools within the scope of this framework, which is an open- source software tool implemented in Java and freely available in <http://www.lvc.ele.puc-rio.br/wp/?p=1831>. Up to now, it embodies four classification algorithms taken from the *WEKA* (Machine Learning Group at the University Waikato, 2014) java library: Naïve Bayes Classifier, Decision Trees, Random Forest and Support Vector Machines (SVM).

The parallel procedure works as follows. The data to be classified, henceforth called *big data* set is stored on HDFS. The training set is stored on an auxiliary storage system. When the execution starts, each HDFS block is processed by a different map task. The map task, firstly, reads the training data set and trains the classifier. After that, the trained classification model is used to classify the *big data set*. The multiple executions of the training step (for each map) should not impact the computational performance substantially because the amount of training data is small compared to the *big data set*, which accounts for most of the processing time.

A brief example of a *Pig Latin* script is presented in Table 1. In this script, a SVM classification process is performed using a given training and testing set, and saving the result in a defined output file.

<i>Pig Latin</i> script that executes a SVM classification
<pre>REGISTER ../pathTO/weka.jar; REGISTER ../pathTO/interimage-pig-datamining.jar; DEFINE II_SVMClassifier br.puc_rio.ele.lvc.interimage.datamining.udf.SVMClassifier ('.../pathTO/trainSet.csv', 'configurationOptions'); dataTest = LOAD '../pathTO/testSet.csv' USING org.apache.pig.piggybank.storage.CSVExcelStorage(',', 'YES_MULTILINE', 'NOCHANGE', 'SKIP_INPUT_HEADER') AS (Att_1:float, ..., Att_n:float); classes = FOREACH dataTest GENERATE II_SVMClassifier(Att_1, ..., Att_n) AS csfrOutcome; STORE classes INTO '../pathTO/output.csv' USING org.apache.pig.piggybank.storage.CSVExcelStorage(',', 'YES_MULTILINE', 'NOCHANGE');</pre>

Table 1. *Pig Latin* script that perform a SVM classification.

1. EXPERIMENTS AND RESULTS

This section reports some experiments conducted upon *ICP: Data Mining Package*. These experiments were carried out on the Mortar platform (Mortar Data, 2014) (Amazon Web Services, 2014), a cloud-computing, open-source framework for organizing, developing, testing, and deploying big data processing applications based on Hadoop. This platform relies on Amazon Elastic MapReduce (Amazon EMR, 2014), which uses the Hadoop framework to distribute the data and processing across a resizable Amazon Elastic Compute Cloud cluster (Amazon EC2, 2014), and on Amazon Simple Storage Service (Amazon S3, 2014).

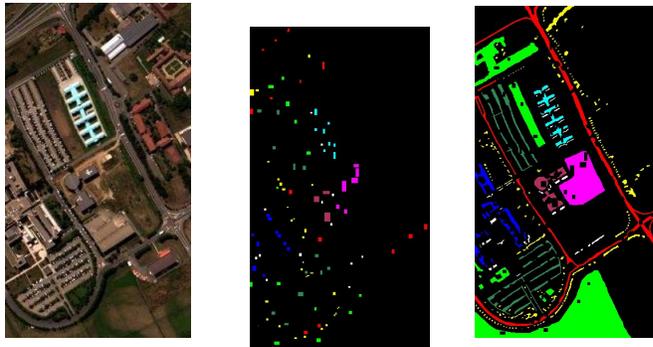
On Mortar one can work directly with *Pig* on Hadoop and configure the number of cluster nodes on which the *Pig Latin* script will be executed in a simple and flexible way. The next sections will present the datasets used in this work.

1.1 Urban Hyperspectral Data Set

The tests reported henceforward were performed on Pavia hyperspectral data set (Hypercomp Research Group, 2014). It consists of a hyperspectral image collected by the ROSIS optical sensor over the University of Pavia, Italy. The image contains 610×340 pixels at 1.3 meters per pixel resolution over 103 spectral bands (from 0.43 to 0.86 μm).

The data set has nine ground truth classes of interest, comprising urban, soil and vegetation classes. In our experiments 3921 pixels were selected for training and 42776 pixels for testing, as shown in Figure 1(b) and Figure 1(c) respectively; and a Pavia hyperspectral false color composition image is presented in Figure 1(a). The classes' distribution within each data set is presented in Table 2.

The size of the Pavia hyperspectral ground truth is approximately 20Mb. Synthetic data sets were built from it, with 100, 200 and 500 times the original data set size, yielding data files with around 2Gb, 4Gb and 10Gb respectively. Only these three datasets were considered



in the experiments since the original dataset is too small for Hadoop.

(a) (b) (c)
 Figure 1. (a) Pavia hyperspectral image (false color composition, R-band 45,G-band 27, B-band 11); (b) Training data set; (c) Test data set.

Classes	Training Set Points	Testing Set Points	Color in Figure 1(c)
Asphalt	548	6304	Red
Meadows	540	18146	Green
Gravel	392	1815	Blue
Trees	524	2912	Yellow
Metal Sheets	265	1113	Cyan
Bare Soil	532	4572	Magenta
Bitumen	375	981	Brown
Self-blocking Bricks	514	3364	Dark Green
Shadow	231	795	White

Table 2. Classes considered in the experimental analysis.

1.1 Experimental Results

The SVM classification algorithm was used to evaluate the tool. WEKA uses the Jhon Platts sequential minimal optimization algorithm for training the SVM (Platt, 1998). In the experiences, a multi-class pairwise (one versus one) SVM classification with a polynomial function kernel was performed, with a complexity parameter $C = 1.0$ and exponent value $\gamma = 1.0$ over a 5-fold cross validation procedure.

The SVM had as inputs, the first nine principal components computed from the 103 bands of the Pavia hyperspectral image. Figure 2 shows the outcome and the overall accuracy. The classification algorithm was applied on the 2GB, 4GB and 10GB data sets, in a local mode configuration (used as baseline) and in clusters with 10, 20 and 50 nodes on the Mortar platform. Each node in the cluster had 4 64-bit virtual cores, 15GB of RAM and a high-performance network.

Table 3 presents the execution times for each test data set running on the local mode and on the 10-, 20- and 50-node cluster configurations. For the local mode, the observed execution time grows almost linearly with the size of the input data. The results also show that the execution times drop consistently from the local mode to the other configurations.

Data Set	Execution Time (seconds)			
	1 Node	10 Nodes	20 Nodes	50 Nodes
2 GB	968	121	112	97
4 GB	2040	146	122	111
10 GB	4827	285	185	126

Table 3. Execution time for each configuration.

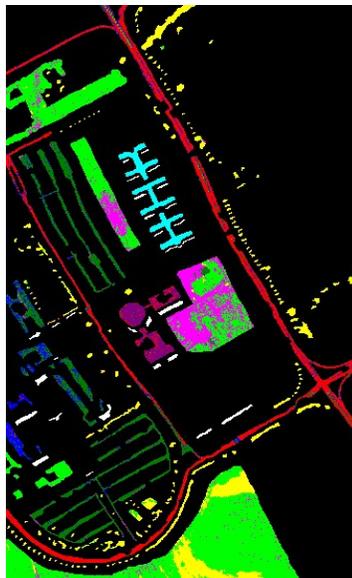


Figure 2. SVM classification outcome image on the Pavia Hypersectral data set (Overall Accuracy: 78,26%).

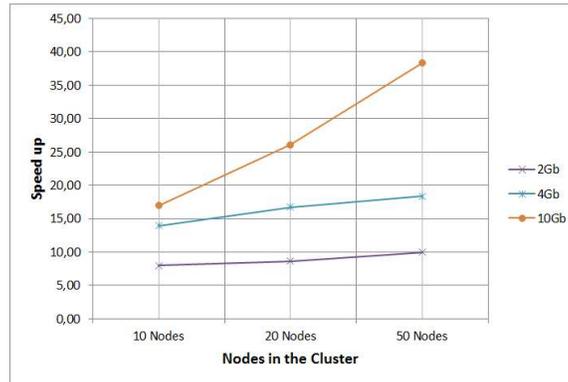


Figure 3. Speedups for each data set.

Figure 3 shows the speedup achieved by each cluster configuration for each image. It can be seen that, as the size of the data set increases, each cluster configuration achieves better speedups. This is because, for larger data sets, Hadoop can profit more from the available cores producing a higher parallelization. This also explains why larger data sets show higher speedup gains from one cluster configuration to the other. For the 4GB image, the speedups were 13.97, 16.72 and 18.38, respectively. Finally, for the 10GB image, the speedups were 16.94, 26.09 and 38.31.

On the other hand, as the cluster size increases, the classification for smaller data sets might underutilize the cluster resources. This result indicates that there is a threshold above which increasing the number of nodes does not produce substantial performance gains. For the 2GB image, for example, the speedups were 8.0,

1. 8.64 and 9.98, for the 10-, 20- and 50- node clusters, respectively.

CONCLUSIONS

In this paper, *ICP: Data Mining Package* is presented, a tool able to perform classification processes on huge amounts of data, exploiting the benefits of working on clusters with the Hadoop framework.

An experimental analysis indicated that the speedup achieved by the tool increases with the amount of data being processed. Additionally, the results showed that increasing the number of nodes in the cluster does not necessarily provide a corresponding reduction of execution times. Thus, the proper cluster configuration depends not only on the operations to be executed but also on the amount of input data; there must be a balance between the amount of data to be processed and the number of nodes to be used to achieve the best performance.

ACKNOWLEDGEMENTS

The authors acknowledge the support provided by CNPq (Conselho Nacional de Desenvolvimento e Pesquisa), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and FP7 (Seventh Framework Programme) in the scope of the TOLOMEO Project.

REFERENCES

- Amazon Web Services, 2014. *Amazon EC2*. <http://aws.amazon.com/ec2/> (3 Nov. 2014).
- Amazon Web Services, 2014. *Amazon EMR*. <http://aws.amazon.com/elasticmapreduce/> (3 Nov. 2014).
- Amazon Web Services, 2014. *Amazon S3*. <http://aws.amazon.com/s3/> (3 Nov. 2014).
- Amazon Web Services, 2014. *AWS Case Study: Mortar Data*. <http://aws.amazon.com/solutions/case-studies/mortar-data/> (3 Nov. 2014).

- Apache PIG, 2014. *Welcome to Apache Pig!*. <http://pig.apache.org/> (7 Nov. 2014).
- Bekkerman, R., Bilenko, M., and Langford, J., 2012. *Scaling up Machine Learning: Parallel and Distributed Approaches*. Cambridge University Press.
- Dai, W., and Ji, W., 2014. A MapReduce Implementation of C4.5 Decision Tree Algorithm. *International Journal of Database Theory and Application*, 7(1), pp. 49-60.
- Dean, J., and Ghemawat, S., 2004. MapReduce: Simplified Data Processing on Large Clusters. *Proceedings of the 6th Conference on Symposium on Operating Systems Design and Implementation*, 6, pp. 137-149.
- Dhillon, S., and Kaur, K., 2014. Comparative Study of Classification Algorithms for Web Usage Mining. *International Journal of Advanced Research in Computer Science and Software Engineering*, 4(7), pp. 137-140.
- Ferreira, R., Oliveira, D., Happ, P., Costa, G., Feitosa, R., and Bentes, C., 2014. InterIMAGE 2: The Architecture of an Open Source, High Performance Framework for Automatic, Knowledge-Based Image Interpretation. *International Geographic Object-Based Image Analysis Conference*. Thessaloniki, Greece.
- Grolinger, K., Hayes, M., Higashino, W., L'Heureaux, A., Allison, D., and Capretz, M., 2014. Challenges for MapReduce in Big Data. *IEEE 10th World Congress on Services*, pp. 182- 189.
- Han, J., Liu, Y., and Sun, X., 2013. A Scalable Random Forest Algorithm Based on MapReduce. *4th IEEE International Conference on Software Engineering and Service Science*, pp. 849-852.
- He, Q., Zhuang, F., Li, J., and Shi, Z., 2010. Parallel Implementation of Classification Algorithms Based on MapReduce. *International Conference on Rough Set and Knowledge Technology*, pp. 655-662.
- Hypercomp Research Group., 2014. New Digital Repository for Remotely Sensed Hyperspectral Imagery with Unmixing Based Retrieval Functionality. <http://www.hypercomp.es/repository/> (3 Nov. 2014)
- Intel IT Center., 2012. *Apache Hadoop Community Spotlight: Apache MapReduce*. <http://www.intel.com/content/www/us/en/big-data/hadoop- spotlight-apache-mapreduce-paper.html> (3 Nov. 2014).
- Kiran, M., Kumar, A., Mukherjee, S., and Prakash, R., 2013. Verification and Validation of MapReduce Program Model for Parallel Support Vector Machine. *International Journal of Computer Science Issues*, 10(3), pp. 317-325.
- Kishor, D., 2013. Big Data: The New Challenges in Data Mining. *International Journal of Innovative Research in Computer Science & Technology*, 1(2), pp. 39-42.
- Li, J., Xu, Z., Jiang, Y., and Zhang, R., 2014. The Overview of Big Data Storage and Management. *International Conference on Cognitive Informatics and Cognitive Computing*, pp. 510- 513.
- Liu, B., Blasch, E., Chen, Y., Shen, D., and Chen, G., 2013. Scalable Sentiment Classification for Big Data Analysis Using Naïve Bayes Classifier. *IEEE International Conference on Big Data*, pp. 99-104.
- Machine Learning Group at the University Waikato, 2014. *Weka 3: Data Mining Software in Java*. <http://www.cs.waikato.ac.nz/~ml/weka/index.html> (3 Nov. 2014)
- Mortar Data, 2014. *Mortar*. <https://www.mortardata.com/> (3 Nov. 2014).
- Nandakumar, A., and Yambem, N., 2014. A Survey on Data Mining Algorithms on Apache Hadoop Platform. *International Journal of Emerging Technology and Advanced Engineering*, 4(1), pp. 563-565.
- Olston, C., Reed, B., Srivastava, U., Kumar, R., and Tomkins, A., 2008. Pig Latin: A Not-So-Foreign Language for Data Processing. *Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data*, pp. 1099- 1110.

Pakize, S., and Gandomi, A., 2014. Comparative Study of Classification Algorithms Based On MapReduce Model. *International Journal of Innovative Research in Advanced Engineering*, 1(7), pp. 215-254.

Platt, J., 1998. Fast Training of Support Vector Machines using Sequential Minimal Optimization. *Advances in Kernel Methods*. MIT Press, pp. 185-208.

Sagiroglu, S., and Sinanc, D., 2013. Big Data: A Review. *International Conference on Collaboration Technologies and Systems (CTS)*, pp. 42-47.

Shvachko, K., Kuang, H., Radia, S., and Chansler, R., 2010. The Hadoop Distributed File System. *IEEE 26th Symposium on Mass Storage Systems and Technologies*, pp. 1-10.

Suthaharan, S., 2014. Big Data Classification: Problems and Challenges in Network Intrusion Prediction with Machine Learning. *ACM SIGMETRICS Performance Evaluation Review*, 41(4), pp. 70-73.

Yadav, C., Wang, S., and Kumar, M., 2013. Algorithm and Approaches to Handle Large Data - A Survey. *International Journal of Computer Science and Network*, 2(3), pp. 37-41.

Zaslavsky, A., Perera, C., and Georgakopoulos, D., 2012. Sensing as a Service and Big Data. *Proceedings of the International Conference on Advances in Cloud Computing (ACC)*, pp. 21-29.

Internet of things (IoT) and smart systems

T.GOPALAKRISHNAN

IT DEPARTMENT

JJCET

TRICHY

Introduction:

Today the Internet has become ubiquitous, has touched almost every corner of the globe, and is affecting human life in unimaginable ways. We are now entering an era of even more pervasive connectivity where a very wide variety of appliances will be connected to the web. One year after the past edition of the Cluster book 2012 it can be clearly stated that the Internet of Things (IoT) has reached many different players and gained further recognition. Out of the potential Internet of Things application areas, Smart Cities (and regions), Smart Car and mobility, Smart Home and assisted living, Smart Industries, Public safety, Energy & environmental protection, Agriculture and Tourism as part of a future IoT Ecosystem. We can use them collaboratively to achieve complex tasks that require a high degree of intelligence. For this intelligence and interconnection, IoT devices are equipped with embedded sensors, actuators, processors, and transceivers. IoT is not a single technology; rather it is an agglomeration of various technologies that work together in tandem. Sensors and actuators are devices, which help in interacting with the physical environment. The data collected by the sensors has to be stored and processed intelligently in order to derive useful inferences from it. Note that we broadly define term *sensor*; a mobile phone or even a microwave oven can count as a sensor as long as it provides inputs about its current state.

Smart Systems Services:

Supply-chain monitoring process is important for companies, distributors and manufacturers. The IoT applications and services consist of Supply-Chain Order Verification, Automated Reordering and Shipping (SCOVARS) operations. The operations involved are planning and scheduling of production, scheduling deliveries, shipping, delivery confirmation from customer, automated reordering from customers, order verification, acknowledgement operations which repeat in each cycle.

An example of applications of SCOVARS is sale of toys, such as LEGO at stores and planning and scheduling of production of toys. Inventory control of each kind of toys uses the RFID labels at the toys, packages and containers for shipping. Whenever a toy sells at a store, such as Walmart, the inventory auto adjusts. Reordering of the toys sold during a week takes place using an app for automated reordering at the customer's end. When a reordering event occurs, then that processes at a server and the event messages route to the ERP applications at the LEGO enterprise.

Abstraction of SCOVARS

Production End (PE) assigns an identity to a root node (level 0) of data tree. Supply End (SE) assigns identities at daughter data-nodes (level 1) for each shipment. Shipping Node (SN) assigns identities at container data nodes (level 2). Shipping End (SHE) consists of identities of group of containers data nodes (level 3). Sales Organization End (SOE) receiving the group of toy containers consists of identities of containers at data nodes.

(level 4). Point of Sale End (POSE) assigns the identities at leaf nodes. A code designer abstracts the PE to POSE end-to-end communication as communication of data trees. Each data tree has data nodes with Universal Resource Identifiers (URIs), messages, alerts and triggers. Data nodes are distributed nodes in the PE-POSE supply chain network.

The code designer abstracts the POSE to PE end-to-end communication as communication of events, alerts, triggers, messages and data files. Each file has reordering information.

Architecture Reference Model for SCOVARS

Designer considers ITU-T four-layers (Figure 1) architecture as the reference.

The layers have generic management, specific management and management and security capabilities. Four layers for Internet of RFIDs for SCOVARS are: Layer 1: Device layer capabilities and gateway capabilities are present in

RFID physical device-cum-RFID reader, which uses the URI at each node.

Layer 2: Transport and network capabilities use the protocol handlers and Internet connectivity.

Layer 3: Services and application-support layer capabilities at a server node are RFID devices URI registry, access management, URI management and URI time series, server node database, events processing and data analytics.

Layer 4: Services and applications capabilities perform the tracking, plan and schedule production, schedule deliveries, shipping, order verification, and acknowledgement operations.

Identifying Requirements at Six Layers at TCCICDD

Requirements at six layers for TCCICDD are as follows:

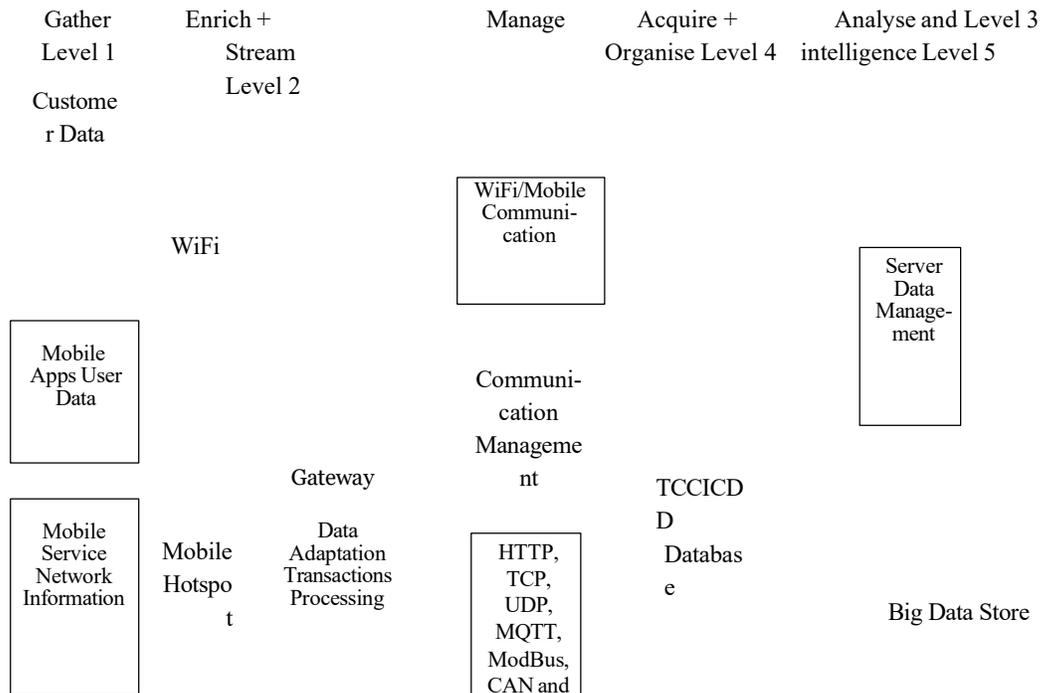
Layer 1 (Gather): The applications and services install at the mobile and wearable devices embed hardware and software for gathering the customer data. An embedded device gathers the customer data for acquiring at server database. The embedded sensors and device hardware and software gather the data and communicate that to the gateway.

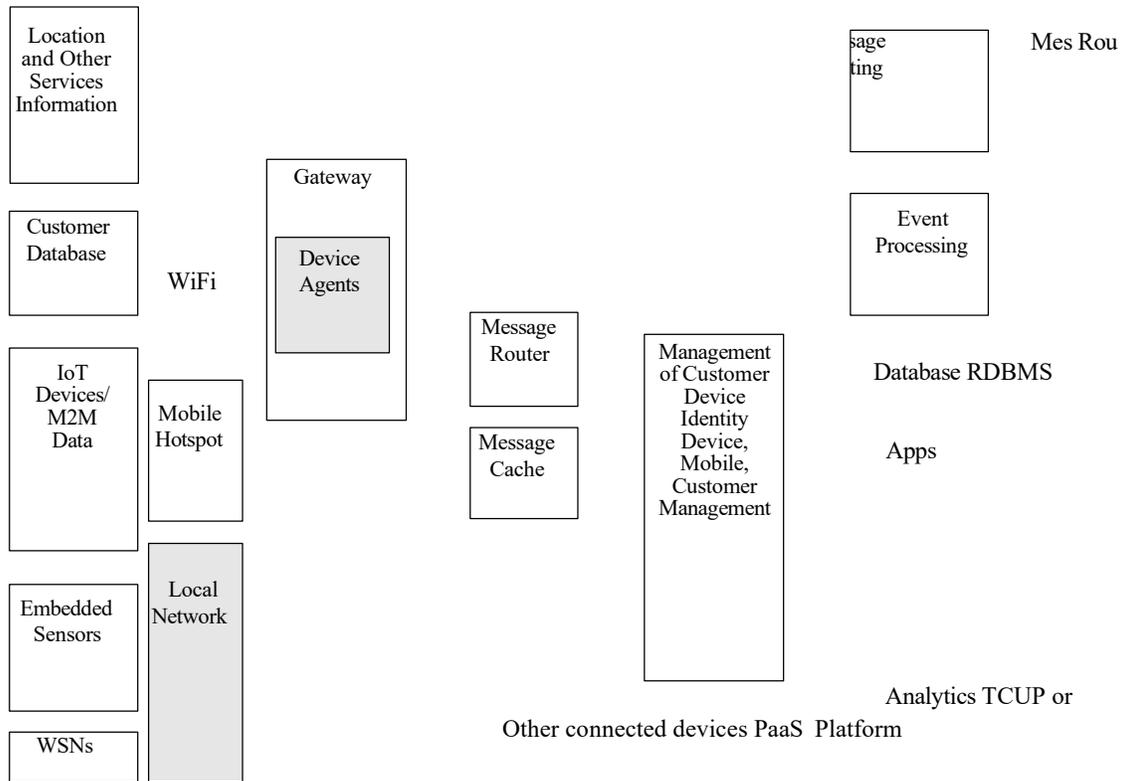
Layer 2 (Enrich and Stream): Gateway software and processor access the data of layer 1 and enriches the data by generating time series and location-stamped data and adapts the data for sensing it to the network. Network card or shield streams to the server using network protocols.

Layer 3 (Manage): Communication management functions does access and ID management, customer data and messages routing and caching.

Layer 4 (Acquire and Organize): Server data store acquires the customer, events and data of devices and diverse sources.

Layer 5 (Analyse and Intelligence): Data is organised as big data store and database RDBMS. Data is analysed using event processing, message routing and analytics.





Layer 6 (Enterprise Integration, Complex Applications Integration and SoA): Applications perform analytics and data visualization, and extract intelligence for service, production and manufacturing, re-planning, re-scheduling or innovating the production and providing better customer experiences using TCCICDD data.

Design Implementation of Hardware, Software, Applications and Services

Design implements layer 1 using mobile apps and embedded sensors and devices software. Design of sensors/devices needs Arduino and Raspberry Pi like computing power depending on the sensor. Example 9.12 explained Internet-connected smartphones for weather and location services using the APIs for web applications. Similarly, a mobile app tracks the location data of the customer.

A program sends data to USB/MMC port for onward transmission on the Internet. Section 9.3 described Eclipse IoT stack-based end-to-end IoT solutions with Java and OSGi. A software developer can use Eclipse Kura development environment, Gateway. Services, cloud connectivity, management of device, network configuration and applications.

Project developer can use TCUP or other PaaS for communication management. Applications developer uses TCUP or other server platform data store, database, events and streams processing, OLTP and OLAP software, database functions, event analytics, analytics and data visualization, and extract intelligence for service, production and manufacturing, re-planning, re-scheduling or innovating the production.

Cluster

The cluster consists of the following systems:

- Engine control
- Speed control and brake system, antilock braking, automatic braking and regenerative braking Safety systems
- Seat and pedal controls
- Car environment controls: Car climate and light controllers control the air- conditioning, heater,

ventilation, windows, light and temperature (ACHVWLT).

- Automobile status monitoring: Automobile status monitoring, tyre pressure, real-time vehicle mileage, data loggers and driving assisting devices management, steerable head lights, windshield IR camera, windscreen heads up display, night-vision assistance, and so on.
- System interfaces for commands, voice activation, and Advanced Driver Assistance System (ADAS) and interfacing (System interfaces are soft programmable buttons, voice commands, and wireless personal area connectivity in the automobile).

The Future Outlook

The future of IoT and smart systems is characterized by continued rapid growth and integration with emerging technologies. The rollout of 5G and low- power wide-area networks (LPWAN) will provide the necessary infrastructure for massive connectivity, while edge computing will enable faster, local data processing for time-sensitive applications like autonomous vehicles. Ultimately, the goal is to create truly **autonomous, self-optimizing systems** that can adapt to changing conditions and solve complex problems without direct human input, promising a more efficient, sustainable, and responsive environment across all aspects of life

ABSTRACT

OPTIMAL CONTROL OF HYBRID SYSTEMS, WITH APPLICATION TO VEHICLE DYNAMICS

P.RAJESH, M.E.,

ABSTRACT

Hybrid systems—those exhibiting both continuous dynamics and discrete event-driven behavior—play a critical role in modern vehicle dynamics, especially with the growing integration of advanced powertrains, automated control strategies, and intelligent transportation technologies. This work presents an optimal control framework for hybrid dynamical systems with specific application to ground vehicle dynamics. The study formulates a unified mathematical model combining continuous-time vehicle motion equations with discrete switching logic governing mode transitions such as gear shifts, braking-acceleration switching, and power-split control in hybrid electric vehicles. Using principles of hybrid optimal control, including dynamic programming and mixed-integer optimal control formulations, optimal trajectories, energy-efficient control policies, and stability conditions are derived. Simulation results demonstrate that the proposed control strategies enhance vehicle performance by minimizing fuel or energy consumption, improving ride comfort, and ensuring safe operation under varying road and environmental conditions. The findings highlight the effectiveness of hybrid system modeling in capturing I-world vehicle behaviors and support the development of next-generation

Performance Optimization and Analytical Evaluation of Post-Surgical Brain Tumor Image Data Using Artificial Intelligence Techniques

Kunal S. Khadke¹, Dr. Narendra Sharma²
Research Scholar, SSSUTMS, Sehore¹
Associate Professor, SSSUTMS, Sehore²

Abstract:

Post-surgical evaluation of brain tumor patients is a critical stage in clinical treatment, as it helps determine tumor recurrence, surgical success, and future therapeutic planning. With the advancement of Artificial Intelligence (AI), automated analysis of medical images-especially MRI and CT scans-has become increasingly reliable and efficient. This research focuses on developing and analyzing an AI-based system for the performance enhancement of post- surgical brain tumor image interpretation. The study integrates deep learning architectures such as Convolutional Neural Networks (CNNs), U-Net, and hybrid transfer-learning models to perform tasks including segmentation, classification, and anomaly detection.

The proposed methodology involves preprocessing of post-operative MRI datasets, noise reduction, enhancement of surgical cavity visibility, and detection of residual or recurrent tumor tissues. Quantitative metrics such as accuracy, precision, recall, F1-score, Dice Coefficient, and Intersection over Union (IoU) are used to evaluate the performance of the AI models. Experimental results indicate that the optimized hybrid deep-learning model demonstrates improved segmentation accuracy and reduced false-positive rates compared to standard CNN architectures. Furthermore, performance analysis reveals that automated AI- based tools significantly accelerate the diagnostic process and offer more consistent outcomes than manual interpretation.

This research contributes to the growing field of AI-driven neuroimaging by offering an efficient model for post-surgical tumor assessment and providing a framework for future enhancements in medical image analytics. The findings can support neurosurgeons and radiologists in achieving more accurate clinical decisions, ultimately improving patient outcomes.

Keywords: Convolutional Neural Networks, AI-driven, Neuroimaging, MRI Datasets, Brain Tumor.

Soil Prediction Using IoT: A Review

Harsha K K

Shri Nehru Maha Vidyalaya College of Arts & Sciences, Coimbatore

Abstract

The integration of Internet of Things (IoT) technologies with data-driven models has reshaped how soil health is monitored and predicted, enabling more informed and sustainable agricultural practices. This review highlights major advancements in IoT-based soil sensing systems that capture parameters such as moisture, pH, temperature, and nutrient composition in real time. IoT sensors paired with machine learning algorithms—including Random Forest, Support Vector Machines, and lightweight neural networks—can reliably classify soil type, predict moisture levels, and forecast short-term soil behavior. While current models show promising performance, challenges remain, including sensor calibration issues, inconsistent field data, limited edge-device processing power, and the need for region-adaptive prediction models. Emerging approaches suggest that combining remote sensing imagery with IoT data can significantly improve prediction accuracy. This review emphasizes the importance of developing scalable IoT-ML frameworks that farmers can adopt easily, ultimately improving crop yield, irrigation efficiency, and soil management in a practical and human-centered way.

A Comprehensive Review of AI Techniques for Modernizing Traditional Art and Craft

V.G.Thamaraiselvi

Artificial Intelligence and Data Science

Jerusalem College of Engineering

Abstract:

Traditional art and craft hold significant cultural value, and Artificial Intelligence (AI) is increasingly used to preserve and modernize these practices. This paper presents a concise review of AI techniques applied to traditional art forms such as Kolam, painting, textiles, and handicrafts. Key methods reviewed include neural style transfer, convolutional neural networks, generative adversarial networks, and diffusion models for digitization, pattern generation, restoration, and design enhancement. Recent applications, available datasets, and emerging research trends are summarized to provide a focused understanding of current progress. The review also highlights major challenges such as limited annotated data, cultural sensitivity, and model interpretability. Overall, this study offers a clear overview of how AI supports the modernization of traditional art and craft while emphasizing future opportunities for culturally aware and creativity-enhancing AI systems.

Keywords— *Artificial Intelligence, Traditional Art, Craft Modernization, Deep Learning, Kolam, Neural Style Transfer, GANs.*

The Intelligent Safety Helmet with Monitoring and Prevention System

Dhivyadharshini K

Medical Electronics

Sengunthar Engineering College

Abstract:

The increasing number of road accidents involving two-wheelers, especially due to drunk driving, underage riding, and not wearing helmets, indicates the need to include advanced safety measures. Conventional safety relies too heavily on the compliance of riders, which is seldom consistent and results in preventable injuries and fatalities. This project proposes a smart helmet safety system that incorporates basic electronics and sensors with ai to promote better safety for riders and ensure that regulations are followed. An Arduino controls this system, which will monitor alcohol consumption using an mq3 sensor, check if the rider is wearing the helmet with a limit switch, and authenticate the rider through RFID-based license scanning and ai-based facial recognition, including checking for under-age riding. Additionally, if the model experiences any vibration from accidents, it will send GPS location and alert contacts via a gsm module in no time. The novelty of this work is that by providing an overall automated solution to preventive unsafe travel and speedy emergency responses, two-wheeler travel would be safer using intelligent monitoring and control.

Development and Optimization of Oral Films of Hydroxyzine Hydrochloride for Anti-Allergic Action

Jahanvi kumari

Department of pharmaceutical science and technology
Maharaja Ranjit Singh Punjab Technical university, Bathinda, Punjab

Abstract:

Atopic eczema, commonly known as atopic dermatitis, is a chronic form of dermatitis that significantly impacts the appearance and function of the skin. This condition leads to the development of inflamed, cracked, swollen, red and intensely itchy skin. Hydroxyzine hydrochloride is one of the strongest H1 receptor antagonists in the market and is the preferred antihistamine for treating allergic skin problems due to its potent anti-pruritic action. Hydroxyzine hydrochloride has sedative properties due to its ability to cross the blood-brain barrier and exert an effect on the central nervous system. Chitosan-based oral films represent a groundbreaking approach in the treatment of atopic dermatitis. Oral films enhance bioavailability and can deliver medication in a controlled manner. Oromucosal films should have rapid disintegration, uniform drug distribution, palatability and biocompatibility while maintaining stability. The drug may be immediately absorbed via the mucosal membranes, avoiding the gastrointestinal system and first pass metabolism in the liver due to its mucoadhesive characteristic. Oral film was prepared by the solvent casting method, and optimized by central composite design (CCD). An experiment was performed for 13 batches with different concentrations of chitosan and PEG-400 at different ratios. CCD and ANOVA results demonstrate that polymer and plasticizer significantly influenced film flexibility, mechanical strength and disintegration time ($p < 0.005$). The optimized formulation A1 showed thickness (0.050.01mm), tensile strength (23.640.05), high swelling index (63.3%), uniform drug content (72.44%) and 96% drug release within 6 hours.

TRANSMISSION LINE MULTIPLE FAULT DETECTION AND INDICATION TO ELECTRICITY BOARD

T.Sutha

Computer science and Engineering
Selvam College of Technology

Abstract:

Power distribution systems commonly suffer from faults (e.g. line faults, transformer failures, transient disturbances) with substantial delays in detection and repair, leading to prolonged outages, economic losses, and reduced reliability. Existing monitoring relies heavily on manual reporting by consumers, periodic inspections, or legacy SCADA systems, which often lack coverage, real-time feedback, or require continuous power supply. This work proposes a smart, cost-effective platform for fault detection and monitoring in power distribution systems that communication with self-powered IoT devices (e.g. solar or other energy harvesting) to achieve near real-time fault detection and localization. The system continuously senses fault indicating parameters (voltage, current, etc.), uses interrupt-based triggers to minimize energy consumption, and communicates over long distances via LoRa to a central monitoring centre. Experimental results show fault notification latency reduced to ~100 ms, outperforming existing reporting methods by orders of magnitude. The self-powering component ensures deployment in remote or infrastructure-poor regions, low power, long-range characteristics make it suitable for broad coverage. The proposed platform promises improved system reliability, faster outage restoration, and lower maintenance overhead. Transmission line cable system is a common practice followed in urban areas. While a fault occurs due to many reasons in the cable, at a time of removing or repairing process, there is difficulty in locating the exact location of the fault. The system proposed in this system is used to find out the exact location of the fault and display it to the dedicated website over internet using Wi-Fi module, giving prior information to the authorized person at other end

Automated seizure prediction from long term EEG recordings

Sudha K

M.E - Medical Electronics

Sengunthar Engineering College, Tiruchengode

Abstract:

Epileptic seizures pose significant challenges for patients due to their unpredictability, necessitating real-time predictive systems for early intervention. This project presents a patient-specific seizure prediction framework leveraging long-term scalp EEG recordings. The pipeline ingests multichannel EEG signals, which are pre-processed and transformed into enhanced representations using ESRGAN for improved feature resolution. Both manually extracted features and deep-learned representations are utilized to classify each time window into inter-ictal, pre-ictal, or ictal states. A Decision Tree classifier translates these features into probabilistic predictions, while a post-processing layer converts window-level probabilities into actionable early warnings, achieving tuneable sensitivity and specificity with a low False Prediction Rate (FPR) per hour. The system is implemented in MATLAB with an intuitive GUI, enabling real-time visualization of EEG signals and seizure risk levels. Validation on publicly available datasets and prospective recordings demonstrates the model's efficacy, highlighting its potential for patient-specific early warning and timely clinical intervention.

Performance Optimization and Analytical Evaluation of Post-Surgical Brain Tumor Image Data Using Artificial Intelligence Techniques

Kunal S. Khadke¹, Dr. Narendra Sharma²
Research Scholar, SSSUTMS, Sehore¹
Associate Professor, SSSUTMS, Sehore²

Abstract

Post-surgical evaluation of brain tumor patients is a critical stage in clinical treatment, as it helps determine tumor recurrence, surgical success, and future therapeutic planning. With the advancement of Artificial Intelligence (AI), automated analysis of medical images-especially MRI and CT scans-has become increasingly reliable and efficient. This research focuses on developing and analyzing an AI-based system for the performance enhancement of post- surgical brain tumor image interpretation. The study integrates deep learning architectures such as Convolutional Neural Networks (CNNs), U-Net, and hybrid transfer-learning models to perform tasks including segmentation, classification, and anomaly detection.

The proposed methodology involves preprocessing of post-operative MRI datasets, noise reduction, enhancement of surgical cavity visibility, and detection of residual or recurrent tumor tissues. Quantitative metrics such as accuracy, precision, recall, F1-score, Dice Coefficient, and Intersection over Union (IoU) are used to evaluate the performance of the AI models. Experimental results indicate that the optimized hybrid deep-learning model demonstrates improved segmentation accuracy and reduced false-positive rates compared to standard CNN architectures. Furthermore, performance analysis reveals that automated AI- based tools significantly accelerate the diagnostic process and offer more consistent outcomes than manual interpretation.

This research contributes to the growing field of AI-driven neuroimaging by offering an efficient model for post-surgical tumor assessment and providing a framework for future enhancements in medical image analytics. The findings can support neurosurgeons and radiologists in achieving more accurate clinical decisions, ultimately improving patient outcomes.

Keywords: Convolutional Neural Networks, AI-driven, Neuroimaging, MRI Datasets, Brain Tumor

**Mathematical modelling of Optimal Controllability Dynamics of Malaria Disease
Transmission by SEIQR Epidemic Model**

U. Kalpana¹ G.Kuppuswami²

¹*Department of Nautical Science (Mathematics), AMET Deemed to be University,
Chennai – 603112, Tamil Nadu, India*

³*Department of Mathematics, Panimalar Engineering College,
Chennai – 600123, Tamil Nadu, India*

Abstract:

In this paper, the transmission of Malaria disease an SEIQR model with treatments is analyzed to human and mosquitoes. The mathematical modeling of Malaria disease is formulated into a first-order nonlinear differential equation [1]. The Homotopy Perturbation Method is used to find the analytical solution of this SEIQR model, and also simulation results are identified further, it is compared with analytical solution. In this model β , δ , γ are considered. In the SEIQR model, A is the natural birth rate, β denotes the infectious rate, δ denotes the incubation rate, γ is the recovery rate and d denotes the disease related death.

Keywords: Malaria disease, Nonlinear differential equation, SEIQR Model, Homotopy Perturbation Method, computer simulation

Advanced Machine Learning Frameworks for Intelligent Classification of Weed and Crop Species in Precision Agriculture

Sathya R

Computing Science

Vels Institute of Science, Technology & Advanced Studies (VISTAS)

Abstract:

Efficient classification of weed and crop species is essential in precision agriculture to support targeted field interventions, optimize resource use, and reduce reliance on broad-spectrum herbicides. This study introduces advanced machine learning frameworks for intelligent plant species classification using high-quality, image-derived features. A diverse dataset of field images representing multiple crop and weed species is processed to extract discriminative color, texture, and morphological descriptors. Several supervised machine learning models- such as Gradient Boosting, LightGBM, and ensemble meta-learners-are developed and optimized using Recursive Feature Elimination (RFE) and Principal Component Analysis (PCA) to enhance feature relevance and minimize redundancy. Experimental results show that the proposed frameworks deliver high classification accuracy and strong robustness across varying field conditions. These findings demonstrate the capability of advanced machine learning systems to enable selective weed control, reduce environmental impact, and support automated decision-making in precision agriculture. Experimental results on diverse agricultural datasets demonstrate that these intelligent classification systems outperform traditional machine learning approaches, offering reliable, real-time solutions for automated weed and crop identification. This research paves the way for intelligent decision support systems that contribute to increased productivity and sustainable agricultural practices.

Keywords: Classification, Weed and crop management, machine learning, ensemble techniques, agriculture, feature extraction

AI-FROWD WEBSITE DETECTION SYSTEM

Srikrishna P

Computer Science and Engineering

Erode Sengunthar Engineering College

Abstract:

AI-Fraud Website Detection System aims to develop an intelligent security tool capable of identifying fraudulent and malicious websites using artificial intelligence. The System focuses on collecting and analyzing website data such as URL patterns, domain age, HTTP/HTTPS information, HTML content features, and visual webpage characteristics. Machine learning algorithms will be implemented to classify websites as legitimate or harmful based on extracted features. The system will include modules for data preprocessing, feature engineering, model training, and accuracy evaluation. A web-based interface will allow users to submit URLs for real-time threat assessment. The model will generate a prediction score indicating the level of website safety. The System also emphasizes building a scalable dataset and optimizing the model for high-performance classification. By integrating automated web scraping and AI- driven detection, the system seeks to minimize human effort in identifying online threats. It aims to reduce the risk of phishing attacks, identity theft, and financial fraud. The solution will support users in making safer online decisions. The developed system can be extended to detect newly emerging cyber threats. The implementation highlights current needs in cybersecurity automation. Overall, the project focuses on creating a reliable, fast, and effective

Performance Optimization and Analytical Evaluation of Post-Surgical Brain Tumor Image Data Using Artificial Intelligence Techniques

Kunal S. Khadke¹, Dr. Narendra Sharma²

Research Scholar, SSSUTMS, Sehore¹

Associate Professor, SSSUTMS, Sehore²

Abstract:

Post-surgical evaluation of brain tumor patients is a critical stage in clinical treatment, as it helps determine tumor recurrence, surgical success, and future therapeutic planning. With the advancement of Artificial Intelligence (AI), automated analysis of medical images—especially MRI and CT scans—has become increasingly reliable and efficient. This research focuses on developing and analyzing an AI-based system for the performance enhancement of post-surgical brain tumor image interpretation. The study integrates deep learning architectures such as Convolutional Neural Networks (CNNs), U-Net, and hybrid transfer-learning models to perform tasks including segmentation, classification, and anomaly detection.

The proposed methodology involves preprocessing of post-operative MRI datasets, noise reduction, enhancement of surgical cavity visibility, and detection of residual or recurrent tumor tissues. Quantitative metrics such as accuracy, precision, recall, F1-score, Dice Coefficient, and Intersection over Union (IoU) are used to evaluate the performance of the AI models. Experimental results indicate that the optimized hybrid deep-learning model demonstrates improved segmentation accuracy and reduced false-positive rates compared to standard CNN architectures. Furthermore, performance analysis reveals that automated AI-based tools significantly accelerate the diagnostic process and offer more consistent outcomes than manual interpretation.

This research contributes to the growing field of AI-driven neuroimaging by offering an efficient model for post-surgical tumor assessment and providing a framework for future enhancements in medical image analytics. The findings can support neurosurgeons and radiologists in achieving more accurate clinical decisions, ultimately improving patient outcomes.

Keywords: Convolutional Neural Networks, AI-driven, Neuroimaging, MRI Datasets, Brain Tumor

Hybrid Nature-Inspired Optimization Techniques for Secure Data Transmission in Decision Support Systems

Rajesh Verma¹, Dr. Ravindra Tiwari²

Research Scholar, LNCT University Bhopal Madhya Pradesh¹

Associate Professor, LNCT University Bhopal Madhya Pradesh²

Abstract

As the digital networks become more complex and the volume of sensitive information grows, the safety of transmission has become a key aspect of supporting decision making in the present day. The paper introduces a hybrid nature-based optimization framework to improve the level of data security, reliability and efficiency within the distributed setting. This paper puts forward a model that integrates swarm intelligence and evolutionary algorithms to optimize the encryption parameters, intrusion detection and routing mechanisms on a dynamic basis. The framework emulates adaptive mechanisms occurring in nature, i.e., cooperation, evolution, self-organization, etc., providing strong defence against cyber-attacks with minimal communication overhead. The experimental assessment reveals that the hybrid solution is better than traditional security algorithms in the following aspects, namely, the accuracy of detection, the efficiency of their computation, and their ability to adapt to changing attack patterns. The results demonstrate the opportunity of applying nature-based intelligence to cybersecurity-related decision support systems that can be used to attain adaptive and sustainable secure data transmission.

Keywords: Cybersecurity, Nature-inspired optimization, Hybrid algorithms, Secure data transmission, Decision support systems, Swarm intelligence, Evolutionary computing.

Offline Disaster Communication System

Gokul S

Computer science and Engineering
Erode Sengunthar Engineering College

ABSTRACT

During disasters such as earthquakes, floods, and cyclones, conventional communication infrastructures like cellular networks and the internet are often disrupted, leading to critical delays in rescue and relief operations. Offline Disaster Communication and Detection System using IoT and AI, designed to function without Wi-Fi or cellular networks. It uses LoRa, Zigbee, and BLE for offline, low-power communication. A modified AODV mesh algorithm ensures dynamic, multi-hop routing and self-healing of the network in case of node failure. IoT sensors detect disasters like earthquakes, floods, and gas leaks. Edge AI devices like Raspberry Pi process sensor data locally for real-time detection. Alerts are transmitted across the mesh using LEDs, buzzers, and voice. A custom offline mobile app enables BLE/NFC-based communication between users and rescue teams. RFID is used for survivor tracking and identification. The system runs on solar power with battery backup. This provides a resilient, decentralized solution for disaster response without relying on internet infrastructure.