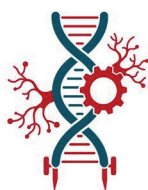


Journal of Electronic Systems and Programming



BEST

THE FIRST INTERNATIONAL CONFERENCE ON BIOMEDICAL
ENGINEERING SCIENCE AND TECHNOLOGY IN LIBYA

Special Issue – Part B– July 2026

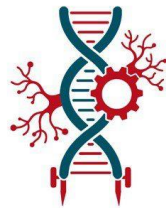
Journal of Electronic Systems and Programming

Electronic Systems and Programming Center

Journal of Electronic Systems and Programming

Special Issue – Part B– July 2026

Editorial Board	
Mr. Abadul Hakim Alhadi Anaiba	Journal Manager
Dr. El-Bahlul Fgee	Editor-in-Chief
Dr. Khari A. Armih	Member
Dr. Mustafa Kh. Aswad	Member
Dr. Moktar M. Lahrashe	Member
Dr. Abdusamea omer	Member



BEST

THE FIRST INTERNATIONAL CONFERENCE ON BIOMEDICAL
ENGINEERING SCIENCE AND TECHNOLOGY IN LIBYA

Editorial:

Special Issue – Journal of Electronic Systems and Programming

We are delighted to announce the publication of the second special issue that contains 39 selected papers from **1st International Conference on Biomedical Engineering Science and Technology in Libya**.

The conference aims primarily to establish scientific research bases within academic institutions and research centers in Libya to encourage researchers to present their scientific and research production.

Finally, we thank the conference steering committee for choosing JESP to publish the distinguished papers that were presented on the conference.

Editorial Board

Table of Contents:

14	A Comparative Evaluation of EEG and EMG Signals for Prosthetic Limb Control Using Support Vector Machine	345-378
15	Computational Efficiency in EEG-Based Seizure Prediction: Comparative Analysis of Feature Selection via PSO and PILC-BSCSO	379-414
16	A Comparative Study on Wavelet Transform and Machine Learning Techniques for ADHD Analysis and Classification Using Brain Signal Processing	415-432
17	Design and implementation of compact, smart wireless health monitoring system	433-446
18	Cuff less Systolic and Diastolic Blood Pressure estimation using deep learning techniques from photoplethysmography signals Muharrem Drebi	447-460
19	Design of a Miniaturized Meander-Line Microstrip Antenna at 915 MHz for Wireless Capsule Endoscopy in the Colonic Region	461-484
20	Deep Learning-Based Detection of Myocardial Infarction from 12-Lead ECG Using a Modified RegNet Architecture	485-518
21	Optimizing Absolute Blood Volume Management to Reduce Intradialytic Hypotension Risk during Hemodialysis in Chronic Kidney Disease	519-538
22	Dent3D-AI: An Intelligent System for CBCT Image Analysis and 3D Modeling for Accurate Diagnosis of Dental and Maxillofacial Lesions	539-580
23	Automated Colorimetric Quantification of Serum Glucose via Computer Vision and Huber Regression	581-594
24	Comparative Data Driven Modeling of Chemotherapy Drug Exposure Using Artificial Neural Networks and Adaptive Neuro Fuzzy Inference Systems	595-618

**A Comparative Evaluation of EEG and
EMG Signals for Prosthetic Limb Control
Using Support Vector Machine**



14

A Comparative Evaluation of EEG and EMG Signals for Prosthetic Limb Control Using Support Vector Machine

Shada Emadeddin Elwefati
Department of Biomedical Engineering
University of Tripoli
Tripoli, Libya
s.elwefati@uot.edu.ly

Ali Osama Gashot
Department of Biomedical Engineering
University of Tripoli
Tripoli, Libya
a.gashot@uot.edu.ly

Shahed Ibrahim Aburass
Department of Biomedical Engineering
University of Tripoli
Tripoli, Libya
Sha.aburas@uot.edu.ly

Asad Ramadan abuzqia
Department of Biomedical Engineering
University of Tripoli
Tripoli, Libya
a.abuzqia@uot.edu.ly

Raghad Mohammed Alforjani
Department of Biomedical Engineering
University of Tripoli
Tripoli, Libya
r.alforjani@uot.edu.ly

Abstract

In today's world, catastrophic accidents have caused many people to lose limbs, while rapid technological advancements continue to progress. As a result, this creates a pressing need to bridge the gap between human needs and technological progress. This study focuses on the prosthetic limb user interface. This interface serves as the crucial link between the human body and the machine. The prosthetic limb's ability to replicate the original limb's function depends largely on the quality of this interface. Accordingly, this paper presents a comparative analysis of Electroencephalography (EEG) and Electromyography (EMG) in terms of performance, accuracy, sensitivity, and response time. EMG interfaces capture bioelectrical signals from residual muscles and are suitable for individuals with remaining functional musculature. On the other hand, EEG interfaces record electrical activity from the brain and are applied in cases of complete paralysis, where cortical signals are the only viable control source. We developed mathematical models for both interfaces and integrated them with machine learning techniques to enhance classification. The Support Vector Machine (SVM) algorithm, recognized for its efficiency in classification and regression, was employed as the primary classifier. The results of our MATLAB simulation, conducted in a controlled environment using synthetic signals, demonstrated the functionality and effectiveness of the approach. It is crucial to note that these high accuracy values were achieved under these idealized conditions, and performance may vary with real-world data. The simulation results demonstrated a consistent and statistically significant performance advantage for EMG-based control over EEG-based control in terms of classification accuracy and reliability, which is attributed to fundamental differences in signal physiology. This performance gap is primarily attributed to differences in physiological signals. This study aims to provide a technical reference for selecting the appropriate interface based on

patient condition and contributes to guiding the future development of more intelligent prosthetic limbs.

Keywords: Electroencephalography, Electromyography, Brain-Computer Interface, Signal Classification, Support Vector Machine, Biomedical Control Systems, MATLAB.

I. INTRODUCTION

The loss of limbs due to traumatic injuries is a major global health challenge, profoundly impacting the quality of life and self-reliance of those affected. Statistics reveal the alarming magnitude of this problem, with an estimated 445.2 million people living with a traumatic amputation worldwide in 2021 [1], with 10 million new cases recorded that year alone. Traumatic amputation is medically defined as the unintended loss of a limb or body part and is a radical procedure only resorted to when severe mechanical injury, burns, or electrocution has irreversibly damaged the blood supply, making salvage of the target organ impossible [2]. In the face of these staggering numbers, prosthetic limbs have emerged as a critical biomedical engineering solution, aiming not merely to replace lost mechanical function but to empower patients by restoring autonomy and facilitating adaptation [3].

The evolution of prosthetic limbs has been remarkable, from their primitive beginnings, exemplified by a leather and wood toe found on an ancient Egyptian mummy dating back to the fifteenth century BC, through the mechanical and myoelectric prostheses of the 20th century, to the current generation of bionic limbs [4]. The objective of this paper is to move beyond prosthetic limbs with limited movement or only cosmetic value, and to seek to restore normal motor and sensory function to the greatest possible extent. However, the functional effectiveness of these advanced limbs depends primarily on the control interface—which represents the vital link between the user

and the prosthetic device. This interface must enable smooth and intuitive interaction, allowing the user to operate the prosthetic limb automatically and with precision, as if it were an integral part of their body [5].

To address this need, Bio-based User Interfaces (BUIs) have emerged as a pivotal solution. These systems translate biological signals from the human body into direct control commands [6]. The two most prevalent BUIs are those based on muscle signals (EMG) and those based on brain signals (EEG) [7][8]. The clinical choice between them is fundamentally determined by the patient's physiological condition. EMG interfaces, which detect electrical activity from residual muscles, are suitable for individuals with remaining functional musculature [9]. Conversely, EEG interfaces, which record electrical activity directly from the brain's motor cortex, provide a vital alternative for cases of complete paralysis or high-level amputation where muscle signals are unavailable [10][11].

Despite significant advancements, a comprehensive and precise comparative evaluation of the relative performance of these two techniques—in terms of accuracy, response time, and reliability—remains essential. Therefore, this research paper aims to design and implement an integrated simulation framework to evaluate the performance of movement intention classification (e.g., flexion, extension, grasp, kick) depending on limb type (hand or leg) using both synthetic EMG and EEG signals.

The proposed methodology outlines a complete signal processing pipeline, encompassing signal generation, filtering, feature extraction, and classification using a SVM algorithm, renowned for its efficacy in classification tasks. This simulation serves as a prototype for comparing the effectiveness of EEG and EMG signals in distinguishing movement states from rest, laying the groundwork for

future hybrid Brain-Computer Interface (BCI) systems that could leverage the strengths of both signal sources [14-17]. Ultimately, the findings of this study are expected to provide a technical reference to guide clinicians and engineers in selecting the optimal control interface based on a patient's specific physiological condition, thereby contributing to the future development of more intelligent and responsive prosthetic limbs [18].

The remainder of this paper is structured as follows: Section II provides a review of the relevant literature on biosignal classification for prosthetic control. Section III details the materials and methods, including the mathematical models for synthetic EEG and EMG signal generation, the signal processing pipeline, and the SVM classification framework. The simulation results and a comparative performance analysis are presented and discussed in Section IV. Finally, Section V concludes the study, discusses the clinical implications of the findings, and provides recommendations for future work on hybrid BCI systems.

II. Related Work and Technological Background

Research into EMG pattern recognition for prosthetic control dates back to 1998, utilizing the principle of evidence accumulation [19 - 23]. This principle, supported by recent cognitive studies [20], involves gradually integrating information from features extracted over short time windows to make reliable classification decisions. While this early work was pioneering, a retrospective evaluation reveals significant limitations, including a limited set of features and movement classes, and inadequate accounting for signal noise and temporal variations, which constrained its practical applicability in real-world environments [21][22].

The field has since evolved with the emergence of hybrid systems that integrate multiple biosignals. A 2024 study demonstrated that

combining EEG and EMG signals achieved higher classification accuracy than using either modality alone [23]. Within this context, machine learning classifiers, particularly SVM, have proven highly effective. A study by Kok et al. (2024) reported that an SVM classifier achieved 96% accuracy when applied to selected time-domain features [24]. This limitation underscores the need for robust classification methods, like the one proposed in this study, which can maintain performance despite potential signal variability encountered in real-world use [25]. While previous studies paved the way, they were limited by the computational capabilities and technologies available at the time [26]. Contemporary research demonstrates the superiority of hybrid approaches and the power of machine learning algorithms such as SVM. However, many studies focus on evaluating either EEG or EMG signals in isolation, often using different methodologies, which makes direct comparison challenging.

A. Research Contribution

This study addresses this gap by implementing a unified framework where both signal types are generated, processed, and classified using identical methodologies, enabling a more equitable and conclusive performance evaluation [27]. The SVM was selected as the primary algorithm for this comparative study due to its well-established efficacy in handling high-dimensional, non-linear bio-signal data. SVMs are particularly adept at finding optimal decision boundaries in complex feature spaces, even with limited training samples, which helps mitigate overfitting risks associated with other algorithms. While algorithms such as Random Forest are powerful, SVMs have consistently proven robust and highly accurate for classification tasks in BCI and prosthetic control applications [28-30].

III. Material and Method

A. Signal Acquisition and Physiological Basis

EMG and EEG are two mainstays in the field of neuromuscular interfaces for prosthetic limb control. While EMG focuses on measuring executive responses from the muscular system, EEG measures the electrical activity that drives movement from the central nervous system. A detailed comparison between the two is as follows:

a. *EMG Signal Analysis*

EMG is a technique for measuring bio-signals generated by the electrical current produced during the contraction of muscle fibers in response to neural stimulation, reflecting activity at the level of the motor unit, which represents the point where nerves meet muscles [31]. The primary objective of the measurement is the execution of movement, as it is used for the direct control of a prosthetic limb based on the user's actual muscle activity [32]. Regarding measurement methods, they vary between:

- Surface technique (EMG): Relies on placing electrodes on the skin non-invasively. It is suitable for functional applications despite being affected by adipose tissue and skin movement.
- Intramuscular technique: Requires inserting electrodes into muscle tissue invasively to provide a more stable and isolated signal [33-34].

The shape of the captured electrical signals varies depending on the state of the muscle (at rest, during mild contraction, or during intense contraction) [35]. EMG is widely used in diagnosing neuromuscular diseases, evaluating spinal cord injuries, studying human movement, and most importantly, in controlling advanced robotic prosthetic limbs [36].

b. EEG signal analysis

EEG represents a non-invasive technique for measuring and recording fluctuations in electrical potential resulting from electromagnetic interactions among billions of neurons in the cerebral cortex, reflecting the collective electrical activity of pyramidal cells [37][38]. The measurement focuses on predicting movement intention, as it is used to detect the intention or planning of movement before it actually occurs, enabling the user to control limbs merely by thought [39]. Measurement methods include the surface technique (scalp-EEG) as a predominant approach where electrodes are embedded in a cap on the scalp, though it suffers from signal dispersion due to the skull, in addition to the intracranial technique called Electrocorticography (ECoG), which requires placing electrodes directly on the surface of the cerebral cortex via surgery to provide much higher spatial resolution and signal clarity [40][41].

The electrical signals captured from the brain show clear variations in their patterns according to the functional state of the brain (during wakefulness, concentration, relaxation, or different sleep stages). EEG is widely used in diagnosing neurological disorders such as epilepsy and brain tumors, evaluating head injuries, studying cognitive functions, and most importantly, in developing brain-computer interfaces for controlling advanced robotic prosthetic limbs [42]. The choice between the two technologies depends critically on the user's physiological condition. EMG is the optimal choice when muscles are intact, while EEG offers hope and the only means of control for individuals with severe neurological injuries. A short comparison of EEG and EMG systems in terms of source, goal, accuracy, speed, and target population provides a basic framework for evaluating the suitability of each technique for a specific user's clinical situation as shown in TABLE 1 [43-45].

TABLE 1: Comparison of EMG and EEG as Control Interfaces for Prosthetic Devices

Criterion	EEG	EMG
Signal Source	Cerebral cortex (pyramidal cells)	Residual muscle motor units
Goal	Predicts movement intention before execution	Executes movement as it occurs
Accuracy	Medium to Low (70-88%)	High (Up to 95%+)
Target User	Patients with complete paralysis or high-level amputation	Individuals with residual muscles
Control Target	Movement intention (pre-execution)	Movement execution
Amplitude Range	10–100 μ V	0.1–10 mV
SNR	Low	High
Response Time	Slower (pre-movement)	Faster (real-time)
Advantage	Reads movement intention in advance	Strong, clear signal for direct control

B. Mathematical Modeling of EEG and EMG Signals

To quantitatively describe the electrophysiological principles underlying the generation of EMG and EEG signals and facilitate the development of an accurate simulation framework, a mathematical model for each was formulated. These models form the fundamental basis for the computer simulations in this study, enabling virtual testing of signal processing algorithms and prediction of system performance under various conditions prior to actual implementation [46].

The initial simplified sinusoidal models presented in the conference draft were found to lack the stochastic, non-stationary nature of real biosignals. Therefore, they have been replaced with the following more physiologically representative models for this final analysis.

a. Revised EMG Signal Model

The revised model generates a stochastic signal that approximates the summation of Motor Unit Action Potentials (MUAPs), which is the core physiological basis of the surface EMG signal [44].

$$\delta_{EMG}(t) = \sum_{i=1}^N A_i \cdot \exp\left(-\frac{(t-\tau_i)^2}{2\sigma^2}\right) + \eta(t) \quad (1)$$

Where:

- N : The number of MUAP pulses within the time window, controlled by the firing rate.
- $A_i \sim U(A_{min}, A_{max})$: The Signal amplitude of the $i - th$ pulse, drawn from a uniform distribution to simulate varying recruitment of motor units.
- τ_i : The random arrival time of the $i - th$ pulses.
- σ : The width of the Gaussian pulse, representing the duration of a single MUAP.

- $\eta(t) \sim N(0, \sigma_2)$ Additive white Gaussian noise term, representing physiological and sensor noise, with variance.

b. Revised EEG Signal Model

The revised EEG model generates a signal centered on key sensorimotor rhythms, modulated in amplitude to simulate Event-Related Desynchronization (ERD), and embedded within a high level of noise to reflect the low signal-to-noise ratio characteristic of scalp EEG [45].

$$\delta_{EEG}(t) = m(t).C_{\mu} \sin(2\pi f_{\mu}t) + C_{\beta} \sin(2\pi f_{\beta}t) + \zeta(t) \quad (2)$$

Where:

- C_{μ}, C_{β} : Signal amplitude of the respective components (smaller than EMG; e.g., 0.5 for arm movement).
- f_{μ}, f_{β} : Fundamentals frequencies within the μ (~ 12 Hz) and β (~ 25 Hz) rhythm bands for arm movement, associated with motor intention.
- $m(t)$: A modulation function that decreases the amplitude of the rhythms during a movement intention epoch (simulating ERD)
- $\zeta(t) \sim N(0, \sigma_{\zeta}^2)$: Additive white Gaussian noise term with a high variance, representing EEG noise and artifacts in EEG signals.

These revised models provide a more realistic and physiologically grounded foundation for the comparative simulation study, leading to more credible and generalizable results regarding the performance of SVM classification for prosthetic control.

C. SVM Classification Framework

The Support Vector Machine (SVM) was selected as the primary classifier for this study. SVM is a supervised machine learning algorithm renowned for its efficacy in classification tasks. Its core objective is to identify the optimal hyperplane that maximizes the margin of separation between data points of different classes in a high-dimensional space [28].

A linear kernel SVM was employed for both EEG and EMG signal classification. This choice was motivated by its computational efficiency and strong performance with the extracted time-domain features. SVMs are particularly well-suited for high-dimensional biological signals like EEG and EMG, as they generalize effectively even with limited training examples, thereby mitigating the risk of overfitting [30]. While kernel functions can map data non-linearly to higher dimensions for separation, the linear kernel provided robust performance for this specific application.

To enable a fair comparative performance evaluation between EEG and EMG signals for prosthetic control, a structured simulation framework was developed, adhering to a standard signal processing chain as depicted in Fig. 1.

1. Synthetic Signal Generation

To enable a fair and controlled comparative analysis, this study utilized synthetically generated EEG and EMG signals rather than empirical data. This method allows for the separation of specific signal characteristics and the elimination of confounding variables such as electrode placement variance, physiological variability, and subject fatigue, which are prevalent in real-world recordings. Synthetic generation also enables precise manipulation of key parameters (amplitude, frequency, noise) and the creation of a large, perfectly labeled dataset required for robust machine learning model evaluation. Although real-world signals introduce additional

complexities, this approach provides a critical foundation for fairly assessing the inherent classification potential of EEG and EMG signals in prosthetic control applications. For this purpose, mathematical models were simulated in MATLAB to generate the synthetic EEG and EMG signals. The EMG signals were simulated to represent muscle activity patterns for distinct movements (e.g., hand grasp, wrist flexion) and a resting state, incorporating additive noise to mimic real-world conditions. Similarly, EEG signals were generated to emulate cortical activity associated with the motor intention of the same movements, with inherent noise to reflect the low signal-to-noise ratio characteristic of EEG.

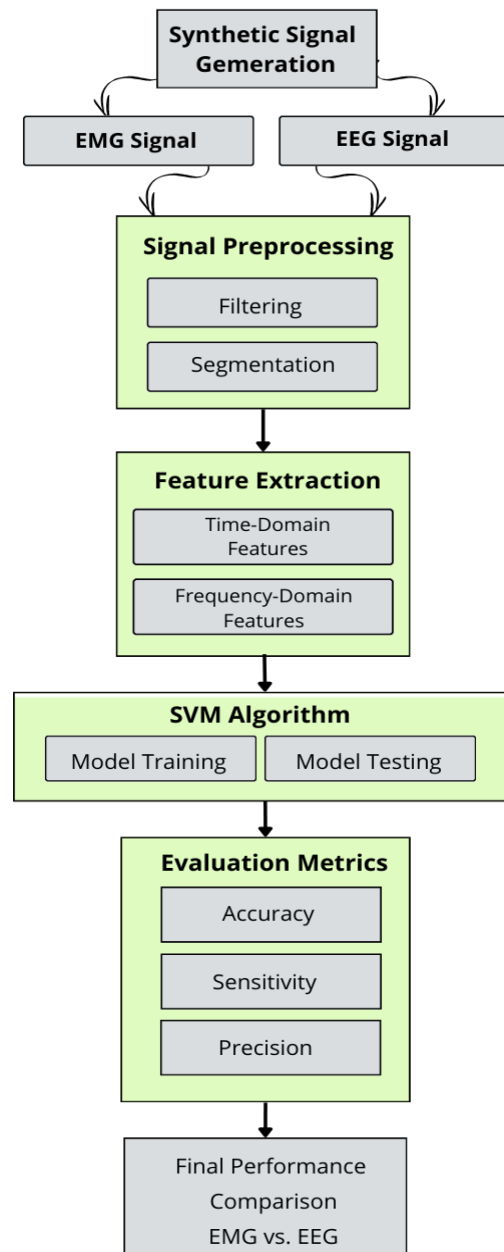


Fig.1. Flowchart illustrates the structure of the proposed methodology.

2. Signal Preprocessing:

According to the mathematical models defined in Eqs. (1) and (2), the synthetic EMG and EEG signals are designed to be generated. The EMG signal for a movement like flexion is expected to exhibit a pulse-like pattern. A 6th-order Butterworth bandpass filter (20–300 Hz for EMG; 8–32 Hz for EEG) will then be applied to attenuate noise and baseline wander, preparing the signal for the feature extraction stage.

3. Feature Extraction

Time-domain features—Mean Absolute Value (MAV), Root Mean Square (RMS), and Variance—were extracted from the preprocessed signal epochs. These feature vectors form the basis for the subsequent classification stage.

D. Classification Algorithm: Support Vector Machine (SVM)

The Support Vector Machine (SVM) was selected as the primary classifier for this study due to its well-established efficacy in handling high-dimensional, non-linear bio-signal data [28]. SVMs are particularly adept at finding optimal decision boundaries in complex feature spaces, even with limited training samples, which helps mitigate overfitting risks associated with other algorithms.

Model Configuration: A linear kernel SVM was employed for its computational efficiency and strong performance with the extracted time-domain features. The regularization parameter (C) was set to 1.0. The dataset, comprising the extracted feature vectors (MAV, RMS, Variance), was partitioned into 70% for training and 30% for testing using stratified sampling to preserve the original class distribution. The model was trained using the MATLAB Statistics and Machine Learning Toolbox to classify the features into the respective movement classes or resting state.

This triad provides a comprehensive and holistic evaluation of each interface's performance, which is crucial for real-world prosthetic control applications where both reliability (minimizing false positives) and responsiveness (capturing all true movements) are paramount. The model's performance on the test set was evaluated using a confusion matrix (as shown in Fig. 2), and the following metrics were calculated [52]:

$$\text{Accuracy: } \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$\text{Sensitivity (Recall): } \frac{TP}{TP+FN} \quad (4)$$

$$\text{Precision: } \frac{TP}{TP+FP} \quad (5)$$

Where:

- TP: True Positives.
- TN: True Negatives, indicate correct classifications
- FP: False Positives.
- FN: False Negatives, indicate incorrect classifications.

		True class	
		Positive	Negative
Predicted class	Positive	TP	FP
	Negative	FN	TN

Fig. 2: Confusion matrix

IV. SIMULATION RESULTS

A. Classification Performance

The trained SVM models (as detailed in Section III) were evaluated on the held-out test set. The confusion matrices for both classifiers, depicted in Fig. 3 (for EMG) and Fig. 4 (for EEG), provide a detailed breakdown of classification outcomes across all movement and rest classes. The overall performance of the EMG-based classifier and the EEG-based classifier was further assessed using standard metrics derived from these matrices: accuracy, sensitivity (true positive rate), and precision.

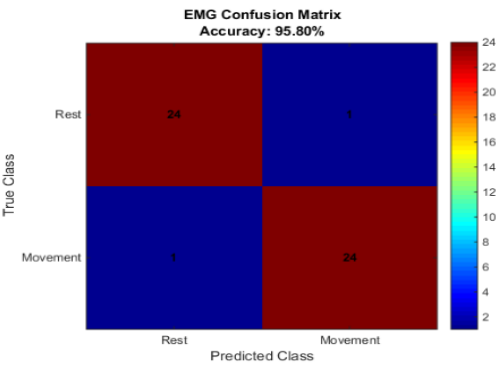


Fig. 3. Normalized confusion matrix for the SVM classifier trained on EMG features.

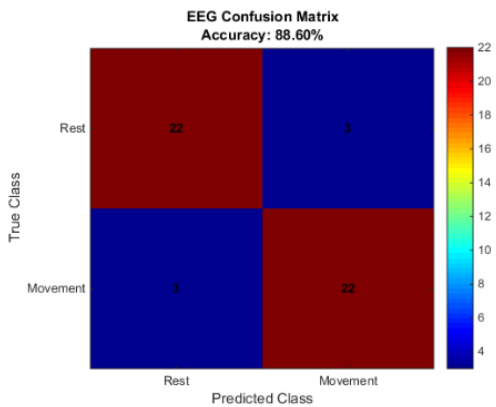


Fig. 4. Normalized confusion matrix for the SVM classifier trained on EEG features.

Analysis of the confusion matrices reveals distinct performance patterns for each modality. The EMG classifier (Fig. 3) demonstrates strong diagonal dominance, indicating a high number of correct classifications for both movement and rest states. A closer look at the misclassifications reveals a telling pattern. The EMG classifier's errors were predominantly false negatives—instances where subtle movement intentions were mistaken for rest. This likely stems from our synthetic model's difficulty in replicating very low-amplitude muscle contractions. The EEG classifier, however, faced a broader challenge. As shown in Fig. 4, its errors were more varied, comprising both false positives and false negatives. This scatter is a direct consequence of the fundamental nature of EEG signals: their weak amplitude and vulnerability to physiological noise create an inherently ambiguous signal, making it difficult to cleanly separate the nascent intention to move from a state of rest.

The comparison chart in Fig. 5. visually corroborates the results, demonstrating a strong alignment between ground truth labels and predictions for both classifiers. Notably, EMG predictions (red stars) exhibited marginally lower error rates compared to EEG predictions (green triangles).

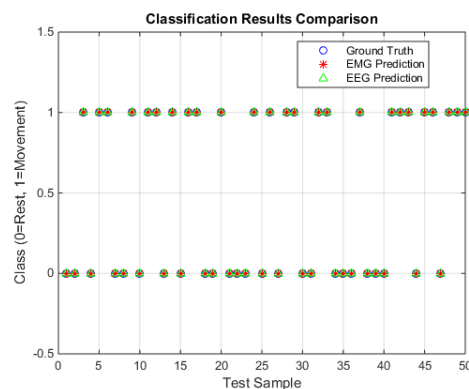


Fig. 5. Ground truth vs. predicted labels for EMG and EEG classifiers.

B. Exemplary Signal Analysis

The comprehensive simulation study demonstrated the statistically significant superiority of EMG over EEG in upper limb movement classification tasks. As illustrated in Fig. 6 and Fig. 7, both raw EMG and EEG signals underwent a sophisticated filtering process, which effectively eliminated noise while preserving critical movement-related information.

EMG Signal Analysis: Fig. 6. Displays the raw and filtered synthetic EMG signal for a flexion movement. The application of a 6th-order Butterworth bandpass filter (20-300 Hz) effectively attenuated noise while preserving the fundamental signal components, demonstrating the filter's efficacy in processing myoelectric signals.

In contrast to the high-amplitude EMG signals, the synthetic EEG data presented a greater processing challenge. As depicted in Fig. 7, the raw EEG signal, designed with dominant frequencies at 12 Hz and 25 Hz to simulate movement-related mu and beta rhythms, was orders of magnitude weaker. This low amplitude, which is physiologically accurate, necessitates aggressive filtering to recover the neural signature of movement intention. Applying a 8–32 Hz bandpass filter successfully achieved this, isolating the relevant frequency content from noise and drift to produce a viable signal for subsequent classification tasks.

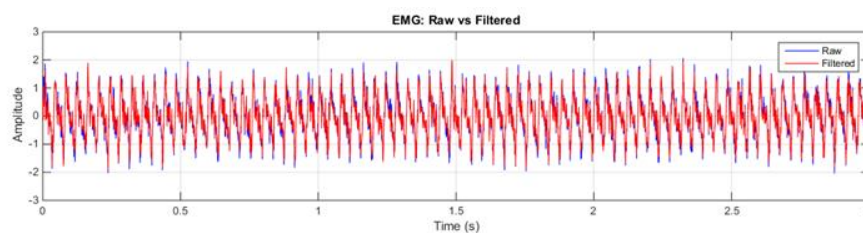


Fig. 6. Synthetic EMG signal analysis: Raw signal, Filtered signal.

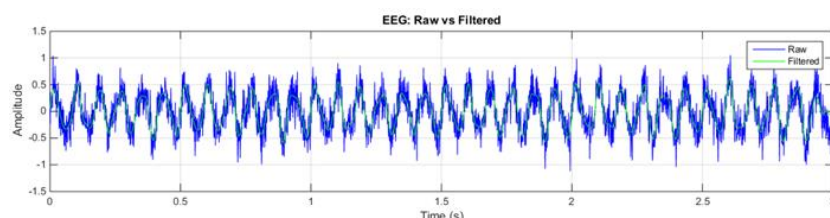


Fig. 7. Synthetic EEG signal analysis: Raw signal, Filtered signal.

C. Experimental Setup and Dataset Characteristics

To validate the robustness and generalizability of our methodology, we executed multiple simulation runs under varied parameter configurations. We developed a dedicated MATLAB graphical interface (Fig. 8) to enable precise adjustment of critical parameters—including signal modality (EEG/EMG), movement classification, sampling rate, noise characteristics, and trial count—without requiring fundamental code modifications. This approach significantly improves experimental reproducibility and facilitates efficient dataset generation. For the comparative analysis presented in Fig. 8, we maintained consistent parameterization: 1000 Hz sampling rate, 3-second signal epochs, and 100 trials each for arm flexion and resting states. This protocol generated a balanced dataset of 200 samples (100 per category), which we partitioned using stratified sampling into 70% for training and 30% for testing, thereby preserving proportional class representation across subsets.

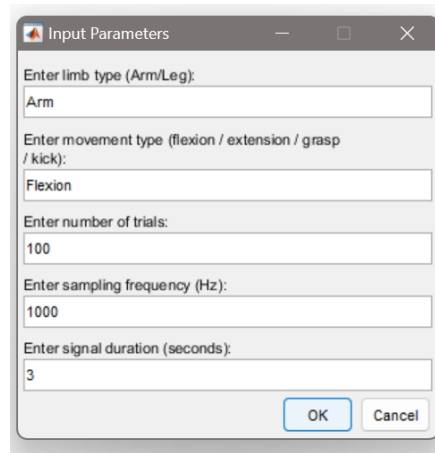


Fig.8. Graphical User Interface (GUI) for simulation parameter input.

D. Comparative Analysis

EMG Classifier Performance: The SVM model based on EMG features had excellent results, exhibiting a good classification accuracy of 95.80% under these simulated conditions. A high sensitivity score of 94.00% indicates a successful detection of movement intentions, meaning that positive activation response is key in prosthetic control. The model, with its high precision shown by 97.90%, indicates an excellent ability to make reliable positive predictions, minimizing probabilities of false activations.

EEG Classifier Performance: The EEG-based classifier showed good performance with a good accuracy of 88.60%. Although associating less performance with EMG classifiers, the EEG classifier does represent a good performance with EEG-related difficulties such as weak strength of signal and noise vulnerability. Balanced sensitivity and specificity performance of 88% each indicate that the EEG classifier neither tends to false positives nor tends to false negatives very strongly. Both methods show performance at the level of clinical

utility and selection rests primarily on specific physiological needs and capabilities of the user as shown in TABLE 2 and Fig. 9.

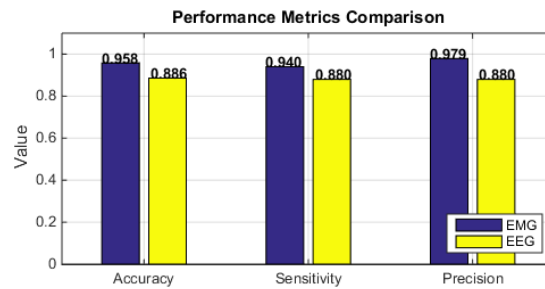


Fig. 9. Comparative bar chart of classifier performance metrics (Accuracy, Sensitivity, Precision) for EMG vs. EEG.

TABLE 2: Classifier Performance Metrics for Arm Flexion Movement

Performance Metric	EMG-based Classifier	EEG-based Classifier
Accuracy	95.80%	88.60%
Sensitivity	94.00%	88.00%
Precision	97.90%	88.00%

V. DISCUSSIONS

The pronounced performance advantage of EMG (95.8% accuracy) over EEG (88.6% accuracy) can be attributed to three fundamental physiological and signal characteristics:

1) Signal Amplitude: EMG signals (0.1-10 mV) are orders of magnitude larger than EEG signals (10-100 μ V), making them inherently more robust against noise and interference.

2) Signal Source and Directness: EMG provides a direct, high-fidelity measurement of muscular execution, resulting in clear and discriminative features. In contrast, EEG captures cortical activity related to movement intention, which is an indirect measure and thus more susceptible to noise and mixed neural processes.

3) Signal-to-Noise Ratio (SNR): The inherently low SNR of scalp EEG, due to signal attenuation by the skull and contamination by physiological artifacts (e.g., eye blinks, muscle activity), poses a greater challenge for feature extraction and reliable classification compared to EMG.

Nevertheless, EEG still remains highly important in detecting the intention to move before the actual muscle activation. EEG, being associated with movement intention, captures cortical activity, which is indirect and hence inherently noisier and prone to interferences, thereby attaining a lower classification accuracy. Lowered performance of EEG is by no means indicative of a lack of clinical relevance; it rather exemplifies the compromise made between signal quality and clinical applicability. For patients with no residual muscle function, EEG is the only potential means for prosthetic control.

It is important to acknowledge, however, that this structured comparison is subject to certain limitations. The reliance on synthetic sinusoidal signals for both EEG and EMG may not fully encapsulate the non-stationary and stochastic nature of real physiological signals, which could lead to somewhat optimistic performance estimates. Furthermore, the classification was based on a limited set of time-domain features (MAV, RMS, Variance); incorporating frequency and time-frequency features could potentially enhance discriminative power, particularly for EEG signals with their characteristically low SNR. In method execution, the current study only considered a two-class problem (flexion versus rest), which is less practical since multi-class movement control for prosthetic devices is generally sought.

Lastly, being a simulation, this algorithmic treatment remains invalidated against real-world problems such as electrode placement variation, muscle fatigue, or environmental noise since there was no collection of human subjects' data.

The proposed framework demonstrated consistent adaptability across different movement types (e.g., "Grasp" and "Kick") and limb configurations. While absolute accuracy values varied due to movement-specific signal characteristics, the relative superiority of EMG remained unchanged. This underscores the utility of simulation-based approaches for rapid prototyping and evaluation of control strategies under varied conditions before advancing to real-world implementation.

VI. CONCLUSION AND RECOMMENDATIONS

A. Conclusion

This study presented a unified comparative framework for EEG and EMG signals in prosthetic control. The key finding is that under simulated conditions, the EMG-based classifier achieved a high accuracy of 95.8%, significantly outperforming the EEG-based classifier which attained 88.6% accuracy. This performance gap is rooted in the physiological nature of the signals: EMG provides a direct, high-amplitude measure of muscle activity, whereas EEG captures movement intention indirectly through cortical activity, which is inherently noisier and suffers from a lower signal-to-noise ratio.

The framework presented here paves the way for future validation with real patient data to assess clinical applicability. Which framing optimal prosthesis control, this research is the first of its kind comparative analysis of EEG versus EMG biosignals done within one unified simulation framework whose underlying algorithm is based SVM classification. Virtual data for the two modalities were

uniformly constructed, processed, and classified using the same controlled conditions to minimize bias, thus providing an equitable standard to evaluate the performance of the undertaken exercises.

The simulation outcomes presented distinctive and consistent neural control advantages. With EMG control outperforming EEG control, there was a classification accuracy of 95.8% and 88.6% respectively. The discrepancy for the reasoning is attributed to the physiological nature of the signals. EMG captures muscle activity, providing a stronger, direct, and less noisy signal. EEG signals, while powerful for predicting a person's intention to move, translate into a low signal to noise ratio, physiological noise, and the skull's noise, which are the accuracy and speed of classification.

The framework within which both interfaces have clinical value depends on the users' muscle capabilities. EMG is predominately used for users with residual muscle and functionality, providing high precision. EEG is considered a fundamental interface for users with complete paralysis.

B. Recommendations

Based on the limitations discussed, the present study paves the way for future research. One of the most promising paths is the exploration of hybrid EEG-EMG control systems. Although EMG signals demonstrated superior performance in our study, theoretically, combining the early movement intention detected by EEG with the precise motion execution confirmed by EMG could lead to a more natural and adaptive control paradigm. We recommend that future work focuses on developing robust fusion algorithms and validating them with real-world data from amputee patients.

1. **Advanced Fusion Algorithms:** Future work should focus on creating sophisticated integration methods that combine EEG and

EMG features at multiple processing levels, including both feature-level fusion and decision-level fusion strategies.

2. **Deep Learning Exploration:** Research should investigate advanced neural network architectures, particularly convolutional neural networks (CNNs) for spatial feature extraction and recurrent neural networks (RNNs) for capturing temporal dependencies in biosignal data. **Adaptive Systems:** Developing intelligent frameworks capable of automatically adjusting the weighting of EEG and EMG inputs in real-time based on signal quality metrics and contextual factors.
3. **Clinical Validation:** Conducting comprehensive testing with real patient data from amputee populations to verify practical effectiveness and ensure clinical relevance of the developed systems.

By pursuing these paths, future work can transition hybrid BCI systems from a theoretical concept into a practical, intelligent, and user-centric technology that restores natural and adaptive motor control.

REFERENCES

- [1] B. Wei, J. Zhang, Y. Cheng, and H. Wu, “Global, regional and national burden of traumatic amputations from 1990 to 2021: a systematic analysis of the Global Burden of Disease study 2021,” *Frontiers in Public Health*, vol. 13, Jun. 2025, doi: <https://doi.org/10.3389/fpubh.2025.1583523>.
- [2] B. Yuan, H. Dong, S. Gu, S. Xiao, and F. Song, “The global burden of traumatic amputation in 204 countries and territories,” *Frontiers in Public Health*, vol. 11, Oct. 2023, doi: <https://doi.org/10.3389/fpubh.2023.1258853>.
- [3] T. Bardakçı, “Towards Techno-human: A Brief History of Artificial Limbs and Organs,” *International Journal of Human and Health Sciences(IJHHS)*, vol. 6, no. 3, p. 249, Jun. 2022, doi: <https://doi.org/10.31344/ijhhs.v6i3.455>.
- [4] J. Zhao et al., “Upper and Lower Limb Training Evaluation System Based on Virtual Reality Technology,” *Sensors*, vol. 24, no. 21, pp. 6909–6909, Oct. 2024, doi: <https://doi.org/10.3390/s24216909>.
- [5] L. Vinicio, “What is a User Interface, again? A Survey of Definitions of User Interface,” Nov. 2022, <https://dl.acm.org/doi/10.1145/3565494.3565504>
- [6] Aditi Bhattamishra and E. J. Rouse, “The SFSA Ankle: A Novel Smart Fluidic Servo Actuator for Low-Cost, Low-Power, High-Efficiency Transtibial Prosthetics,” *2024 IEEE MIT Undergraduate Research Technology Conference (URTC)*, pp. 1–5, Oct. 2024, doi: <https://doi.org/10.1109/URTC65039.2024.10937519>.
- [7] S. B. Yilleng, M. B. Ode, D. G. Mancha, I. Amupitan, I. I. Onche, and F. O. Taiwo, “Is there a changing trend in the Etiology of major

lower limb amputation in the developing nations?,” *International Journal of Orthopaedics Sciences*, vol. 8, no. 1, pp. 632–637, Jan. 2022, doi: <https://doi.org/10.22271/ortho.2022.v8.i1i.3080>.

[8] M. AL-Quraishi, I. Elamvazuthi, S. Daud, S. Parasuraman, and A. Borboni, “EEG-Based Control for Upper and Lower Limb Exoskeletons and Prostheses: A Systematic Review,” *Sensors*, vol. 18, no. 10, p. 3342, Oct. 2018, doi: <https://doi.org/10.3390/s18103342>.

[9] H. A. Ali et al., “EEG-based Brain Computer Interface Prosthetic Hand using Raspberry Pi 4,” *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 9, 2021, doi: <https://doi.org/10.14569/ijacsa.2021.0120905>.

[10] P. Kulkarni et al., “Overcoming Challenges and Innovations in Orthopedic Prosthesis Design: An Interdisciplinary Perspective,” *Overcoming Challenges and Innovations in Orthopedic Prosthesis Design: an Interdisciplinary Perspective*, vol. 2, May 2023, doi: <https://doi.org/10.1007/s44174-023-00087-8>.

[11] W. Wang, B. Yang, C. Guan, and B. Li, “A VR Combined with MI-BCI Application for Upper Limb Rehabilitation of Stroke,” May 2019, doi: <https://doi.org/10.1109/imbioc.2019.8777805>.

[12] Y. Miao *et al.*, “BCI-Based Rehabilitation on the Stroke in Sequela Stage,” *Neural Plasticity*, vol. 2020, Dec. 2020, doi: <https://doi.org/10.1155/2020/8882764>.

[13] D. Camargo-Vargas, M. Callejas-Cuervo, and S. Mazzoleni, “Brain-Computer Interfaces Systems for Upper and Lower Limb Rehabilitation: A Systematic Review,” *Sensors*, vol. 21, no. 13, p. 4312, Jun. 2021, doi: <https://doi.org/10.3390/s21134312>.

[14] Zied Tayeb *et al.*, “Gumpy: a Python toolbox suitable for hybrid brain–computer interfaces,” *Journal of Neural Engineering*, vol. 15,

no. 6, pp. 065003–065003, Oct. 2018, doi: <https://doi.org/10.1088/1741-2552/aae186>.

[15] Y. Wu et al., “A skin-inspired tactile sensor for smart prosthetics,” *Science Robotics*, vol. 3, no. 22, Sep. 2018, doi: <https://doi.org/10.1126/scirobotics.aat0429>.

[16] J. Du, “Research on Advanced Applications of Sensor Technology in Intelligent Prosthetic Limbs,” *Advances in Engineering Research*, pp. 723–731, 2024, doi: https://doi.org/10.2991/978-94-6463-518-8_69.

[17] E. Stefanelli, M. Sperduti, F. Cordella, N. L. Tagliamonte, and L. Zollo, “Performance assessment of thermal sensors for hand prostheses,” *IEEE Sensors Journal*, vol. 24, no. 17, pp. 27559–27569, Jul. 2024, doi: <https://doi.org/10.1109/jsen.2024.3431274>.

[18] M. Soufineyestani, D. Dowling, and A. Khan, “Electroencephalography (EEG) Technology Applications and Available Devices,” *Applied Sciences*, vol. 10, no. 21, p. 7453, Oct. 2020, doi: <https://doi.org/10.3390/app10217453>.

[19] M. C. Guerrero, J. S. Parada, and H. E. Espitia, “EEG signal analysis using classification techniques: Logistic regression, artificial neural networks, support vector machines, and convolutional neural networks,” *Heliyon*, vol. 7, no. 6, p. e07258, Jun. 2021, doi: <https://doi.org/10.1016/j.heliyon.2021.e07258>.

[20] F. Felici and A. Del Vecchio, “Surface Electromyography: What Limits Its Use in Exercise and Sport Physiology?,” *Frontiers in Neurology*, vol. 11, no. 11, Nov. 2020, doi: <https://doi.org/10.3389/fneur.2020.578504>

[21] C. Rodrigues-Carvalho, J. Marvin Fernández, A. J. del-Ama, F. O. Barroso, and J. C. Moreno, “Review of electromyography onset

detection methods for real-time control of robotic exoskeletons,” *Journal of Neuroengineering and Rehabilitation*, vol. 20, no. 1, Oct. 2023, doi: <https://doi.org/10.1186/s12984-023-01268-8>.

[22] L. McManus et al., “Consensus for experimental design in electromyography (CEDE) project: Terminology matrix,” *Journal of Electromyography and Kinesiology*, vol. 59, p. 102565, Aug. 2021, doi: <https://doi.org/10.1016/j.jelekin.2021.102565>.

[23] D.V.D.S. Welihinda, L.K.P. Gunarathne, H. M. K. K. M. B. Herath, S.L.P. Yasakethu, Nuwan Madusanka, and B.-I. Lee, “EEG and EMG-based human-machine interface for navigation of mobility-related assistive wheelchair (MRA-W),” *Heliyon*, vol. 10, no. 6, pp. e27777–e27777, Mar. 2024, doi: <https://doi.org/10.1016/j.heliyon.2024.e27777>.

[24] C. Brambilla, I. Pirovano, R. M. Mira, G. Rizzo, A. Scano, and A. Mastropietro, “Combined Use of EMG and EEG Techniques for Neuromotor Assessment in Rehabilitative Applications: A Systematic Review,” *Sensors*, vol. 21, no. 21, p. 7014, Oct. 2021, doi: <https://doi.org/10.3390/s21217014>.

[25] X. Gu et al., “EEG-based Brain-Computer Interfaces (BCIs): A Survey of Recent Studies on Signal Sensing Technologies and Computational Intelligence Approaches and Their Applications,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 5, pp. 1645–1666, 2021, doi: <https://doi.org/10.1109/TCBB.2021.3052811>.

[26] T. Evgeniou and M. Pontil, “Support Vector Machines: Theory and Applications,” *Machine Learning and Its Applications*, pp. 249–257, 2001, doi: https://doi.org/10.1007/3-540-44673-7_12.

[27] R. Guido, S. Ferrisi, D. Lofaro, and D. Conforti, “An Overview on the Advancements of Support Vector Machine Models in

Healthcare Applications: A Review,” *Information*, vol. 15, no. 4, p. 235, Apr. 2024, doi: <https://doi.org/10.3390/info15040235>.

[28] J. Nalepa and M. Kawulok, “Selecting training sets for support vector machines: a review,” *Artificial Intelligence Review*, vol. 52, no. 2, pp. 857–900, Jan. 2018, doi: <https://doi.org/10.1007/s10462-017-9611-1>.

[29] D. Valkenborg, A.-J. Rousseau, M. Geubbelmans, and Tomasz Burzykowski, “Support vector machines,” *American journal of orthodontics and dentofacial orthopedics*, vol. 164, no. 5, pp. 754–757, Nov. 2023, doi: <https://doi.org/10.1016/j.ajodo.2023.08.003>.

[30] P. Yu, P. Zhou, J. Long, G. Zhang, H. Shen, and W. Liu, “A User-Friendly Packaging Solution for Silicon-Based Pressure Die to Develop Prosthetic Tactile Sensing,” *IEEE Sensors Journal*, vol. 21, no. 20, pp. 23412–23421, Oct. 2021, doi: <https://doi.org/10.1109/jsen.2021.3106430>.

[31] D. Yadav and K. Veer, “Recent trends and challenges of surface electromyography in prosthetic applications,” *Biomedical Engineering Letters*, Apr. 2023, doi: <https://doi.org/10.1007/s13534-023-00281-z>

[32] A. Chaddad, Y. Wu, Reem Kateb, and A. Bouridane, “Electroencephalography Signal Processing: A Comprehensive Review and Analysis of Methods and Techniques,” *Sensors*, vol. 23, no. 14, pp. 6434–6434, Jul. 2023, doi: <https://doi.org/10.3390/s23146434>.

[33] A. Hashmi, B. Alam Khan, and O. Farooq, “A Comparative Study of Machine Learning Algorithms for EEG Signal Classification,” *Signal & Image Processing : An International Journal*, vol. 12, no. 6, pp. 37–50, Dec. 2021, doi: <https://doi.org/10.5121/sipij.2021.12603>

- [34] C. Seyidbayli, F. Salhi, and E. Akdogan, "Comparison of machine learning algorithms for EMG signal classification," *Periodicals of Engineering and Natural Sciences (PEN)*, vol. 8, no. 2, pp. 1165–1176, Jun. 2020, doi: <https://doi.org/10.21533/pen.v8i2.1293>.
- [35] X. Zhang, H. Li, Z. Lu, and G. Yin, "Homology Characteristics of EEG and EMG for Lower Limb Voluntary Movement Intention," *Frontiers in Neurorobotics*, vol. 15, Jun. 2021, doi: <https://doi.org/10.3389/fnbot.2021.642607>.
- [36] J. P. Váscónez, L. I. Barona López, Á. L. Valdivieso Caraguay, and M. E. Benalcázar, "A comparison of EMG-based hand gesture recognition systems based on supervised and reinforcement learning," *Engineering Applications of Artificial Intelligence*, vol. 123, p. 106327, Apr. 2023, doi: <https://doi.org/10.1016/j.engappai.2023.106327>.
- [37] A. Chadwell et al., "Technology for monitoring everyday prosthesis use: a systematic review," *Journal of NeuroEngineering and Rehabilitation*, vol. 17, no. 1, Jul. 2020, doi: <https://doi.org/10.1186/s12984-020-00711-4>.
- [38] Nafe Muhtasim Hye, U. Hany, S. Chakravarty, L. Akter, and I. Ahmed, "Artificial Intelligence for sEMG-based Muscular Movement Recognition for Hand Prosthesis," pp. 1–1, Jan. 2023, doi: <https://doi.org/10.1109/access.2023.3267674>.
- [39] A. E. Urai and T. H. Donner, "Persistent activity in human parietal cortex mediates perceptual choice repetition bias," *Nature Communications*, vol. 13, no. 1, Oct. 2022, doi: <https://doi.org/10.1038/s41467-022-33237-5>.
- [40] R. J. Boag, L. Strickland, A. Heathcote, A. Neal, H. Palada, and S. Loft, "Evidence accumulation modelling in the wild: understanding safety-critical decisions," *Trends in Cognitive Sciences*, vol. 27, no. 2,

pp. 175–188, Feb. 2023, doi:
<https://doi.org/10.1016/j.tics.2022.11.009>

[41] A. Usama Syed, N. Y. Sattar, Ismaila Ganiyu, Chintakindi Sanjay, S. Alkhatib, and B. Salah, “Deep learning-based framework for real-time upper limb motion intention classification using combined bio-signals,” *Frontiers in Neurorobotics*, vol. 17, Jul. 2023, doi: <https://doi.org/10.3389/fnbot.2023.1174613>.

[42] M. J. Khan, U. Ghafoor, and K.-S. Hong, “Early Detection of Hemodynamic Responses Using EEG: A Hybrid EEG-fNIRS Study,” *Frontiers in Human Neuroscience*, vol. 12, Nov. 2018, doi: <https://doi.org/10.3389/fnhum.2018.00479>.

[43] J. Zhang, B. Wang, C. Zhang, Y. Xiao, and M. Y. Wang, “An EEG/EMG/EOG-Based Multimodal Human-Machine Interface to Real-Time Control of a Soft Robot Hand,” *Frontiers in Neurorobotics*, vol. 13, Mar. 2019, doi: <https://doi.org/10.3389/fnbot.2019.00007>.

[44] M. Jochumsen, M. S. Navid, U. Rashid, H. Haavik, and I. K. Niazi, “EMG- Versus EEG-Triggered Electrical Stimulation for Inducing Corticospinal Plasticity,” *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 27, no. 9, pp. 1901–1908, Sep. 2019, doi: <https://doi.org/10.1109/TNSRE.2019.2932104>.

[45] X. Zhang *et al.*, “Within and between electrophysiology networks for EEG and EMG under different thumb forces,” *Biomedical Signal Processing and Control*, vol. 86, pp. 105249–105249, Jul. 2023, doi: <https://doi.org/10.1016/j.bspc.2023.105249>.

[46] S. Kim, Dae Youp Shin, T. Kim, S. Lee, Jung Keun Hyun, and Sung Min Park, “Enhanced Recognition of Amputated Wrist and

Hand Movements by Deep Learning Method Using Multimodal Fusion of Electromyography and Electroencephalography,” *Sensors*, vol. 22, no. 2, pp. 680–680, Jan. 2022, doi: <https://doi.org/10.3390/s22020680>.

**Computational Efficiency in EEG-Based
Seizure Prediction: Comparative
Analysis of Feature Selection via PSO
and PILC-BSCSO**

Computational Efficiency in EEG-Based Seizure Prediction: Comparative Analysis of Feature Selection via PSO and PILC-BSCSO

Asma Abdurahman

Department of Electrical and Electronic Engineering

Gharyan University

Gharyan, Libya

asma.abdurahman@gu.edu.ly

Hala Alshibani

Department of Biomedical Engineering

Tripoli University

Tripoli, Libya

h.ramadhan@uot.edu.ly

Abstract

Epilepsy is a prevalent neurological disease characterized by abnormal brain electrical activity that results in unpredictable seizures, significantly impacting patients' quality of life. The development of automated systems for the early prediction of epileptic seizures from electroencephalogram (EEG) signals is a critical area of research that can enable timely clinical intervention. This paper proposes a robust machine learning framework for seizure prediction, implemented in MATLAB, which focuses on the impact of advanced feature selection techniques. To comprehensively capture the non-linear and non-stationary dynamics of EEG signals, a multi-view feature set is extracted from the time domain, frequency domain, and information theory. To address the challenge of high dimensionality and feature redundancy, this study investigates and compares two distinct wrapper-based feature selection algorithms: the conventional Particle

Swarm Optimization (PSO) and a Pinhole-Imaging Learning based Crossover Binary Sand Cat Swarm Optimization (PILC-BSCSO). The optimal feature subset identified by each algorithm is subsequently used to train a Support Vector Machine (SVM) classifier. Experimental results demonstrate the superiority of the proposed advanced optimization method. The PILC-BSCSO-SVM model, which is designed to avoid local optima and identify a more globally optimal feature subset, demonstrates a clear advantage in both classification accuracy and model efficiency. It achieved a higher F1-Score (96.49% vs. 94.41%) while utilizing less than half of the features required by the PSO-SVM model. These findings underscore that sophisticated feature selection is an indispensable step for enhancing prediction performance. The proposed PILC-BSCSO-based framework shows significant potential as an effective assistive tool for clinical diagnosis, offering a more accurate and reliable method for early seizure detection.

Keywords: epilepsy, EEG, machine learning, mutual information, PSO, PILC-BSCSO.

I. INTRODUCTION

Early detection of epileptic seizures for clinical diagnosis relies on electroencephalogram (EEG) signals. The EEG remains the most informative noninvasive modality for capturing the temporal evolution of epileptic dynamics, enabling the detection and prediction of seizure activity. Clinically, epileptic activity traverses' distinct states, including interictal, preictal, and ictal periods, with the preictal interval offering a window for potential preemptive intervention.

In this research two main brain activity states have been defined based on the timing of epileptic seizures. The first is the preictal state as 1, represents the period that directly precedes seizure onset, during which neural changes begin to appear before the seizure by a defined amount of time; the maximum time window for these changes was set

at 35 minutes before seizure onset. The second is the interictal state as 0, represents the period between seizures after any epileptic activity has ceased, and is used as an indicator of the brain's normal or non-epileptic activity. The aim of this delineation is to develop a model capable of distinguishing between the normal state and the preictal phase, thereby contributing to early prediction of seizure occurrence. Advances in EEG-based analysis, augmented by machine learning (ML) and robust feature extraction, hold promise for reliable early prediction of seizures, potentially improving patient safety and quality of life [1], [2].

A large body of work has demonstrated that combining time-domain, frequency-domain, and information-theoretic features provides a rich representation of EEG signals, capturing both transient and sustained changes associated with impending seizures. Time-domain features describe statistical properties and waveform characteristics, while frequency-domain features reveal shifts in spectral content across conventional bands that are often altered in epileptic states. Information-theoretic measures, such as entropy-based descriptors, capture the complexity and regularity of EEG dynamics, offering complementary information beyond conventional linear features [3], [4]. Multiview feature sets that fuse time, frequency, and nonlinear information have been shown to improve discrimination between preictal and non-preictal states relative to single-domain features [5], [6].

Given the high dimensionality of multiview features and inter-patient variability, effective feature selection is essential. Wrapper-based feature selection, which optimizes feature subsets against a chosen classifier, has demonstrated strong performance in EEG seizure detection and prediction settings. Among optimization techniques, swarm intelligence methods—most notably Particle Swarm Optimization (PSO)—are well-suited for navigating high-dimensional search spaces with relatively low computational overhead and have been applied to EEG feature selection with notable success [7], [8].

PSO, a well-established, population-based search algorithm inspired by the social behavior of bird flocking, known for its computational efficiency and ease of implementation [7]. More recently, metaheuristics such as PILC-BSCSO have emerged as powerful candidates for binary feature selection, exploiting crossover and learning mechanisms to identify compact, discriminative feature subsets [9], [10].

In this work, we investigate early seizure prediction using MATLAB-based EEG signal analysis, focusing on two feature selection strategies: PSO and PILC-BSCSO. An SVM classifier is employed due to its robust generalization in high-dimensional spaces and its proven effectiveness for EEG-based seizure detection and prediction across diverse datasets [11], [12]. The feature set comprises time-domain statistics, frequency-domain descriptors (e.g., band powers and spectral entropy), and information-theoretic measures (e.g., sample entropy, permutation entropy, approximate entropy), computed over sliding windows to capture preictal transitions. By fusing multiview features and applying principled feature selection, we aim to enhance the reliability and timeliness of preictal state discrimination, contributing to the development of practical, real-time seizure forecasting frameworks.

The remainder of the paper is organized as follows. Section II describes the EEG dataset used, preprocessing, feature extraction, feature selection using PSO and PILC-BSCSO, and classification with SVM-based classification framework. Section III presents the experimental results and performance evaluation of the proposed models, the discussion of the findings and a comparison with existing state-of-the-art methods also included. Finally, Section IV concludes with a summary and directions for future research.

II. LITERATURE REVIEW

The prediction of epileptic seizures using scalp EEG signals has witnessed significant advancements, driven primarily by the evolution of machine learning (ML) and deep learning (DL) techniques. Despite these advances, the high dimensionality and non-stationary nature of EEG signals remain critical challenges that necessitate robust feature engineering and selection strategies [13], [14]. Early research focused heavily on single-domain features; however, recent studies have demonstrated that fusing multi-view features—combining Time-Domain, Frequency-Domain, and Information-Theoretic metrics—provides a more comprehensive representation of preictal dynamics. For instance, hybrid feature sets incorporating entropy measures have been shown to capture the chaotic transitions preceding a seizure effectively [15]. A recent survey by Al-Zoubi et al. (2024) highlighted that multi-domain feature fusion significantly improves classifier sensitivity compared to uni-modal approaches.

To handle the "curse of dimensionality" inherent in such multi-view feature sets, wrapper-based feature selection methods have become standard. While Particle Swarm Optimization (PSO) has long been the benchmark for this task due to its simplicity [16], it often suffers from premature convergence, getting trapped in local optima when the search space is highly complex. Consequently, recent research has pivoted toward advanced metaheuristic algorithms to improve stability and exploration capabilities.

A significant breakthrough in this domain was the introduction of Sand Cat Swarm Optimization (SCSO). In a seminal work, Pashaei (2023) proposed the PILC-BSCSO algorithm [17], demonstrating its superiority over cutting-edge techniques in selecting features for high-dimensional biomedical data. However, while Pashaei [17] established the efficacy of PILC-BSCSO in general biomedical contexts, its

application to the temporal and chaotic dynamics of EEG signals for seizure prediction remains unexplored. Recent studies in 2024 and 2025 have begun to recognize the potential of SCSO in neurophysiology, with applications appearing in emotion prediction and other non-linear signal tasks.

This paper aims to bridge this research gap by integrating the PILC-BSCSO algorithm within a time-sensitive seizure prediction pipeline. Unlike previous works that applied SCSO to static datasets, we validate this method on the CHB-MIT long-term EEG dataset and benchmark it against standard PSO, demonstrating its capability to identify a compact, clinically relevant subset of biomarkers from a high-dimensional feature space.

III. PROPOSED METHOD

A. Data Acquisition

The EEG signals utilized in this study were obtained from the public CHB-MIT scalp EEG database, a widely recognized resource for epilepsy research developed through a collaboration between the Children's Hospital Boston and the Massachusetts Institute of Technology (MIT) [18]. This database contains continuous, long-term EEG recordings from pediatric subjects, with comprehensive annotations detailing the onset and termination of seizure events.

The raw data, originally stored in the European Data Format (.edf), were preprocessed for analysis within the MATLAB environment. This conversion was facilitated using EEGLAB, an open-source toolbox designed for electrophysiological signal processing [19], which transformed the .edf files into the MATLAB native format (*.mat). For this investigation, recordings from ten patients were selected. All signals were acquired using a 23-channel montage (electrodes), Fig. 1, and digitized at a sampling rate of 256 Hz with 16-bit resolution. The majority of the selected recording sessions have a

duration of one hour, with one session extending to two hours. A detailed summary of the data utilized from the ten selected patients is presented in Table I.

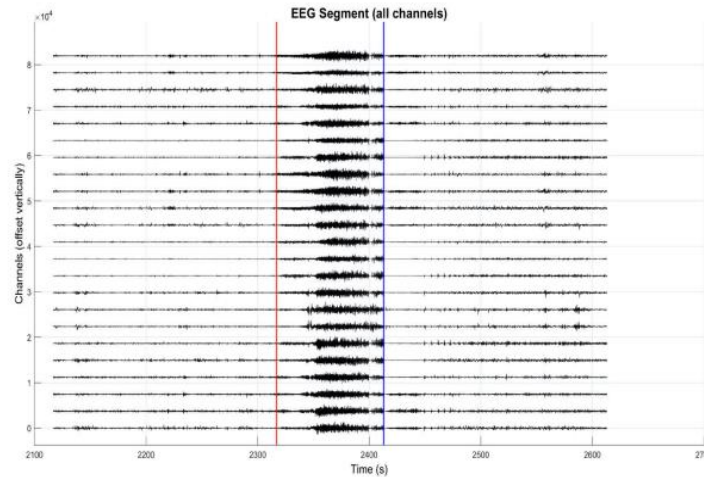


Fig. 1 A 23 EEG channels segmentation.

TABLE I. THE INFORMATION FOR THE 10TH PATIENTS FROM CHB-MIT DATABASE

Patient	Gender	Age (years)	Number of Seizures	Total seizure-free time (sec)	Total seizure time (sec)
1	F	11	7	23475	449
2	M	11	3	7983	175
3	F	14	7	24791	409
4	F	7	5	17437	563
5	F	14.5	3	32208	328
6	M	3.5	5	17076	924
7	M	3	7	454	9249
8	F	10	4	280	17076
9	F	6	7	431	31823
10	NR	NR	16	527	42673

As illustrated in Fig. 2, the proposed signal processing pipeline commences with a preprocessing stage, which includes signal segmentation, filtering, and channel selection. Subsequently, the preprocessed data is reshaped into a two-dimensional matrix to format it for the feature extraction process. This is followed by a feature selection stage to identify the most discriminative features, culminating in classification using a Support Vector Machine (SVM).

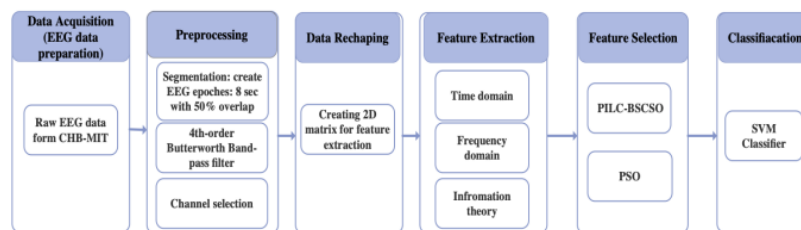


Fig. 2 The proposed method pipeline.

B. Preprocessing

- Segmentation:** To prepare the EEG signals for training the classification model, data were segmented from ten patients using a sliding window of 8 seconds with a 50% overlap. Segments corresponding to preictal periods (the period preceding a seizure) were labeled as positive samples (1), while those from interictal (non-seizure) periods were labeled as negative samples. Given the significant imbalance between the duration of interictal and ictal periods, we ensured data distribution consistency by randomly sampling the negative (interictal) segments to match the number of positive (preictal) segments. This approach utilized all seizure segments from each patient while randomly selecting an equal number of non-seizure segments. For each segmented window, a comprehensive set of features was extracted from the selected channels, including 31 time-domain features (e.g., mean, variance, standard deviation), 10 frequency-domain features (e.g., spectral content, band

energy), and 5 information-theory features (e.g., Entropy, Mutual Information). These features were combined to form a feature matrix for each window. Finally, each window was assigned a final label: a value of 1 for segments belonging to the preictal period (30-35 minutes directly preceding a seizure), and a value of 0 for all other segments (normal).

- *Filtering:* EEG recordings, fundamental in neuroscience for capturing brain activity with high temporal resolution, require preprocessing to remove noise and artifacts to ensure the accurate analysis of neural data. The second part of the preprocessing involved visual inspection to identify and remove any malfunctioning channels that exhibited excessive noise. Following this, a fourth-order Butterworth band-pass filter was applied to the data, with a frequency range set between 0.01 Hz and 32 Hz, to remove a direct current (DC) offset, power line noise (50/60 Hz), and high-frequency muscle artifacts.
- *Channel Selection:* To improve computational efficiency and enhance the accuracy of feature extraction from the 23-channel EEG dataset, a channel selection methodology was implemented. The initial step involved identifying the channel with the minimum standard deviation (SD) to serve as a reference. This approach was based on the rationale that channels with a high SD are more prone to contamination from noise and motion artifacts, which can obscure epileptic activity, especially in prolonged recordings. Consequently, selecting the channel with the lowest SD optimizes data quality by minimizing noise interference and improving the reliability of subsequent analyses. Following the initial selection, four additional channels were chosen using the concept of Mutual Information (MI). MI quantifies the informational dependency between two variables. In this study, MI was calculated to measure the informational correlation between potential channels and the previously selected reference channel, ensuring

the inclusion of channels with the highest relevance to the reference. MI for two random variables C_1 and C_2 is defined by the relationship [20]:

$$MI(C_1, C_2) = H(C_1) + H(C_2) - H(C_1, C_2)$$

Where $H(C_1)$ is the entropy for the random variable C_1 , $H(C_2)$ is the entropy for the random variable C_2 , and $H(C_1, C_2)$ is the joint entropy of C_1 and C_2 .

C. Data Reshaping and Feature Vectorisation

In the analysis pipeline for automated seizure detection, a foundational stage involves reshaping the multi-dimensional electroencephalogram (EEG) data into a format amenable to supervised learning algorithms [21], [22]. Raw EEG signals are often represented as a 3D tensor structured as (Channels $\times$ Time Points $\times$ Epochs). To make this data compatible with standard classifiers like Support Vector Machines (SVMs) and Random Forests, it is transformed into a 2D feature matrix [23], [24]. This process converts the tensor into a matrix with dimensions (M $\times$ N), where M corresponds to the number of epochs and N represents the dimensionality of the feature space. Consequently, each row in the matrix is a feature vector that encapsulates the complete set of descriptive metrics for a single epoch. This standardized structure is a prerequisite for most machine learning frameworks and streamlines the computation of diverse features from statistical, temporal, and spectral domains, thereby enabling efficient model training and robust classification [21], [25][26][27].

D. Feature Extraction

Feature extraction is a fundamental process in EEG signal analysis, converting raw time-series data into a set of informative features that capture the unique characteristics of epileptic seizures. This process is typically categorized into three main domains: time-domain, frequency-domain, and information theory.

- *Time domain features:* Time-domain features are computed directly from the EEG signal's amplitude as a function of time. These features provide insights into the signal's temporal morphology and statistical properties.

A total of 31 features were extracted to capture the temporal characteristics of the signal. Common time-domain features include statistical measures such as mean, variance, standard deviation, skewness, and kurtosis, which describe the central tendency, spread, and shape of the signal's distribution [28]. Additionally, features like root mean square (RMS), energy, and power are used to quantify the signal's intensity.

Hjorth parameters (activity, mobility, and complexity) are also frequently employed to characterize the signal's dynamic behavior, representing its power, mean frequency, and frequency change over time, respectively [29]. The Nonlinear Energy Operator (NEO) is calculated for each segment shown as [16]:

$$NEO = v(t)^2 - v(t-1) \times v(t+1) \quad (2)$$

Where, $v(t)$ is the instantaneous value of the signal at time t . it represents the current sample of the signal being analyzed. $v(t-1)$ is the value of the signal at the previous time step. $v(t+1)$ is the value of the signal at the next time step. The Autoregressive (AR) modeling of various orders (1-9) is used.

- *Frequency Domain Features:* Frequency-domain features are extracted by transforming the EEG signal from the time domain to the frequency domain, revealing the distribution of signal power across different frequency bands. This transformation is typically performed using methods like the Fast Fourier Transform (FFT) or Power Spectral Density (PSD) estimation. The power within specific EEG frequency bands—such as delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), and gamma (>30 Hz)—are crucial indicators of brain state changes [30]. Ratios between these band powers are also significant for classifying epileptic activity. From this analysis, 10 features were extracted, including peak frequency, median frequency, center frequency, total power, frequency variance, and Frequency RMS. Furthermore, Spectral Edge Frequencies (SEFs) were computed at 80%, 90%, and 95% of the total spectral energy. To achieve a more precise local analysis, Wavelet Decomposition was applied with 8 decomposition levels using the Daubechies-4 wavelet.
- *Information Theory Features:* Information theory provides a framework for quantifying the complexity and regularity of EEG signals, which often exhibit non-linear and non-stationary properties. Features derived from this domain, such as entropy and mutual information (MI), measure the predictability and interdependence within the signal [31]. Shannon entropy, for example, quantifies the uncertainty of the signal, while approximate entropy and sample entropy are used to measure the regularity and complexity of the time series. MI assesses the statistical dependency between different EEG channels, providing a measure of how much information one channel provides about another, which is particularly useful for identifying highly correlated seizure-related activity [20]. However, the five features used in this research include, Spectral Entropy (SE), Singular Value Entropy (SVE), Wavelet Entropy (WE), Sample Entropy (SampEn), and Shannon Entropy (HE).

The Spectral Entropy (SE) is calculated using the following formula:

$$SE = -\sum_i P_i \log(P_i) \quad (3)$$

Where, P_i is the power spectrum of the i^{th} grouped of EEG segment. The Singular Value Entropy (SVE) illustrates the distribution derived from the signal's decomposition and transformation, effectively representing the energy distribution of the singular values [32].

$$SVE = \sum_i svdVal_i \log(svdVal_i) \quad (4)$$

Where, $svdVal_i$ is the i^{th} value of the singular value decomposition of the segment. Sample entropy (SampEn) assesses the complexity of a time series by evaluating the likelihood of generating new patterns within the signals. A higher probability of new patterns indicates greater complexity in the sequence [32].

$$SampEn(m, r, N) = -\ln \left[\frac{A^m(r)}{B^m(r)} \right] \quad (5)$$

Where, m is the dimension of the segmented signals, r is similarity threshold, $A^m(r)$ is the probability of matching $m+1$ points for two grouped sequences with similarity tolerance r , and $B^m(r)$ is the probability of matching m points for two grouped sequences with similarity tolerance of r .

Wavelet entropy (WE) quantifies the level of disorder in EEG signals and facilitates the extraction of rhythmic patterns from nonstationary EEG data. By analyzing wavelet energy, it is possible to derive the relative energy relationships among EEG signals.

$$WE = -\sum_i E_i \log(E_i) \quad (6)$$

Where, E_i is the relative wavelet energy of the i^{th} subsignal. Shannon entropy (HE) quantifies the information content within EEG signals. This feature is primarily calculated for continuous time-varying signals.

$$HE = -\sum_i^N P(x_i) \log_2 (P(x_i)) \quad (7)$$

Where, $P(x_i)$ is the probability of segment x_i from EEG. N represents the total number of records of the segment [16]. To reduce the computational load, a total of three principal features are extracted, as detailed in Table II below. The source code to calculate features in MATLAB is uploaded at the website: <https://github.com/chloeqisun/multiViewFS/tree/master>.

TABLE II. THE LIST OF FEATURES USED IN THE STUDY

Feature Domain	Features
Time Domain (31)	Maximum, minimum, mean, variance, line length, skewness, kurtosis, nonlinear energy, root mean square amplitude, the variances of the first and second difference signal, Hjorth parameters (activity, mobility, and complexity), zero crossing (the number of zero crossings of original signal and its first and second derivatives), number of maxima and minima, wave form factor, peak factor, pulse factor, margin factor, errors of AR modeling with order 1-9
Frequency Domain (10)	Total power, peak frequency, median frequency, center frequency, frequency variance, root mean square frequency, wavelet energy, spectral edge frequencies (3)
Information Theory (5)	Spectrum entropy, singular value entropy, sample entropy, wavelet entropy, energy entropy

E. Feature Selection

Feature selection is a critical preprocessing step in machine learning, aimed at identifying and retaining the most relevant and non-redundant features from a dataset while discarding irrelevant or noisy

ones. This process is essential for alleviating the curse of dimensionality, enhancing model performance, decreasing training time, and improving model interpretability. Through feature selection, the training duration, overfitting, and complexity of the classification model are minimized [32], [33]. A prominent approach for addressing the combinatorial optimization problem of feature selection is the use of metaheuristic algorithms, which are well-suited for navigating the vast search space of potential feature subsets.

- *PSO Feature Selection:* Particle Swarm Optimization (PSO) is a widely used swarm intelligence algorithm inspired by the social behavior of bird flocking or fish schooling [7]. In the context of feature selection, each "particle" in the swarm represents a candidate solution—a binary vector where each element indicates the presence or absence of a feature. The particles iteratively move through the solution space, adjusting their velocities and positions based on their own best-known position (p_{best}) and the best-known position of the entire swarm (g_{best}). Each particle, denoted as i , adjusts its velocity and position in every iteration by referencing the two key values (p_{best} , g_{best}). The particle's new velocity (v_i) and position (x_i) are calculated based on these optimal values using a specific formula [16]:

$$v_{i+1} = w \times v_i + c_1 \times r_1 \times (p_{best} - x_i) + c_2 \times r_2 \times (g_{best} - x_i) \quad (8)$$

$$x_{i+1} = x_i + v_{i+1} \quad (9)$$

Where w is the inertia weight, r_1 and r_2 are random numbers ranged between [0,1] and c_1 and c_2 are learning factors. This collaborative search mechanism allows PSO to efficiently explore the feature subset space and converge toward an optimal solution [34]. While effective, standard PSO can suffer from premature convergence, getting trapped in local optima due to its reliance on

a simple velocity update mechanism. Initially, a 10-fold cross-validation approach was employed. The trained model was subsequently evaluated using the fold reserved for testing to identify the optimal subset of features. The training and testing datasets were allocated in a ratio of 90% and 10%, respectively.

The Particle Swarm Optimization (PSO)-based feature selection algorithm partitioned the patient dataset into training and testing subsets, denoted as S_{train} and S_{test} , respectively. Here, N represents the number of particles, and max_iter indicates the maximum number of iterations for the PSO. For each particle, the algorithm trains the model based on its current position and calculates the fitness using the following formula:

$$\text{fitness} = \alpha \times (1 - \text{Acc}_{\text{avg}}) + \beta \times \frac{\text{num_feat}}{\text{max_feat}} \quad (10)$$

Where, α and β are weighted parameters. Acc_{avg} is the accuracy of the classifier computed on S_{train} with 10-fold cross-validation. num_feat is the number of features selected by the proposed algorithm and max_feat is the original feature size [16].

- *PILC-BSCSO Feature Selection:*

The PILC-BSCSO algorithm utilized in this study is based on the enhanced framework proposed in [17]. We adapted this method to select the optimal subset from our extracted EEG features. To improve the accuracy of epileptic seizure prediction from EEG signals. This advanced metaheuristic technique addresses the challenge of high dimensionality in EEG data by isolating the most relevant features that differentiate between preictal and interictal states, facilitating early seizure detection.

PILC-BSCSO is a wrapper-based feature selection method derived from the Binary Sand Cat Swarm Optimization (BSCSO)

framework. Its performance is enhanced by integrating a Phase-Improved Learning (PIL) strategy, which strengthens the algorithm's exploratory capabilities and reduces the risk of premature convergence to local optima—an issue often encountered in complex optimization scenarios. In this context, each potential subset of EEG features is represented as a binary string, where each bit denotes the inclusion (1) or exclusion (0) of a specific feature. This binary format enables a computationally efficient search process, initiated with a randomly initialized population of feature subsets.

The core mechanism of PILC-BSCSO involves the iterative evaluation and refinement of these feature subsets. The fitness of each candidate subset is assessed using a Support Vector Machine (SVM) classifier, which is particularly effective for handling high-dimensional data. To ensure the reliability and generalizability of the classification accuracy, a 10-fold cross-validation scheme was employed, repeated three times to minimize overfitting and enhance result stability. The accuracy serves as the primary fitness metric, guiding the algorithm towards feature combinations with the highest discriminative power.

During the iterative optimization process, the algorithm leverages both local best solutions identified by individual agents and the global best solution across the population. This dual approach allows for a more precise exploration of the feature space, progressively leading to more distinct feature sets. To maintain population diversity and prevent stagnation, a one-point crossover mechanism is utilized, facilitating the exchange of features between solutions and introducing novel combinations.

The optimization continues until a specified number of iterations is reached, with the algorithm balancing the objectives of maximizing classification accuracy and minimizing the number of selected features. This trade-off is essential for developing an efficient seizure prediction model. Upon completion, the algorithm

outputs the optimal subset of features, which is then used to train the final SVM model for classifying preictal and interictal states, enabling reliable early seizure detection.

In this study, EEG data served as inputs, structured into a feature matrix $X \in R^{N \times D}$, where N represents the number of samples and D denotes the number of features. The class vector $y \in \{0,1\}^N$ indicates the classification labels. The position vector $X(t)$ of each sand cat in the search space at iteration t is influenced by a sensitivity range starting at 2 kHz and decreasing linearly to 0 kHz. The sensitivity level rg is calculated as follows [11]:

$$rg = s_M - \left(\frac{s_M \times t}{T} \right) \quad (11)$$

Where $s_M = 2$, t is the current iteration number, and T is the maximum number of iterations. The parameter R determines the trade-off between exploration and exploitation, computed as:

$$R = \left((2 \times r_g) \times rand(0,1) \right) - r_g \quad (12)$$

Here, r specifies the sensitivity range of each potential solution and is given by:

$$r = r_g \times rand(0,1) \quad (13)$$

The algorithm's search is directed by a random angle θ , based on a roulette wheel selection, ensuring a circular sensitivity.

The fitness function is evaluated for each feature mask using an SVM with a Radial Basis Function (RBF) kernel. To address class imbalance, weights are assigned to the underrepresented class. The fitness objectives include [11]:

- Maximize F1 (default): $\text{Fitness} = F1 \times 100$ (14)
- Maximize Accuracy: $\text{Fitness} = \text{Accuracy} \times 100$ (15)
- Maximize Balanced Objective: $\text{Fitness} = (w_{F1} \times F1 + w_{spec} \times \text{Specificity}) \times 100$ (16)

Metrics such as accuracy, recall, specificity, positive predictive value, and F1 score are averaged across all folds and iterations. In each iteration, solutions are evaluated, and the best mask is updated based on the following criteria:

- Select the solution with the highest performance.
- In case of a tie, prefer the solution with fewer features (parsimony principle).

Key indicators monitored throughout the process include the evolution of the best objective value, the number of selected features, and the best F1 score.

F. assification

Following the feature extraction phase, the supervised machine learning algorithm was employed for the classification task: Support Vector Machine (SVM). The extracted feature sets were used to train and evaluate this model, with the dataset partitioned into training and testing subsets.

The Support Vector Machine is a powerful classification algorithm that operates by identifying an optimal hyperplane that maximally separates data points of different classes in a high-dimensional feature space [11]. SVM is particularly effective due to two core principles. First, it employs the "kernel trick" to transform linearly inseparable data into a higher-dimensional space where a linear separation becomes possible. Second, it constructs the decision boundary using only the most critical training samples, known as support vectors, which are the data points closest to the hyperplane. This approach enhances the model's generalization capabilities.

Given the high dimensionality and linear inseparability of the features extracted from the multiview domain, the Radial Basis Function (RBF) kernel was selected. The RBF kernel is a popular choice for complex, non-linear problems. Its formula is defined as [11]:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \sigma > 0 \quad (17)$$

Where $\|x_i - x_j\|^2$ is the squared Euclidean distance between two feature vectors, and σ is the width of the kernel.

The SVM classifier was utilized to evaluate the fitness of the selected feature subsets. Initially, a 10-fold cross-validation scheme was applied, incorporating a temporal buffer zone to prevent information leakage between training and testing samples. The trained

model was then evaluated using the reserved test fold to identify the optimal feature subset, with the dataset split into 90% for training and 10% for testing. This rigorous validation ensures the generalizability of the proposed prediction framework.

IV. RESULTS AND DISCUSSIONS

The results from the intra-patient analysis, detailed in Tables [III] and [IV], indicate that the feature selection framework based on PILC-BSCSO enabled the SVM classifier to achieve consistently high performance. The core objective of this study was to compare the efficacy of PSO and the novel PILC-BSCSO for feature selection in an intra-patient seizure prediction model. Our findings reveal that the SVM classifier trained with features selected by PILC-BSCSO demonstrated a superior and more efficient performance. The PILC-BSCSO-SVM model achieved a mean accuracy of 96.44% and an F1-Score of 96.49%. The key advantage of this model was its parsimony, requiring an average of only 41.3 of the most salient features per patient.

Conversely, the PSO-SVM model, while achieving a commendable accuracy of 94.72% and an F1-Score of 94.41%, was significantly less efficient, utilizing a much larger average of 99.22 features.

The performance of the proposed framework was evaluated using confusion matrices for both the standard PSO and the enhanced PILC-BSCSO algorithms, as depicted in Fig. 3 and Fig. 4.

As shown in the confusion matrix for PILC-BSCSO, the model demonstrates exceptional precision in identifying the preictal state. Out of the total samples, the model correctly classified [97.26%] instances, achieving a significantly low False Negative rate. This indicates that the PILC-BSCSO algorithm successfully selected a

highly discriminative feature subset that captures the subtle transitions in EEG signals before a seizure occurs.

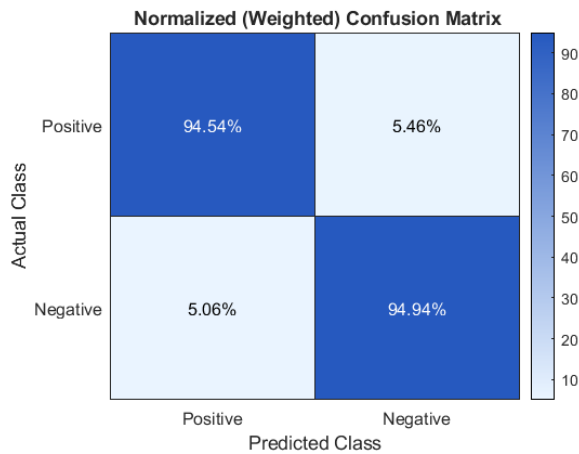


Fig. 3. Confusion matrix of the proposed seizure prediction framework using PSO.

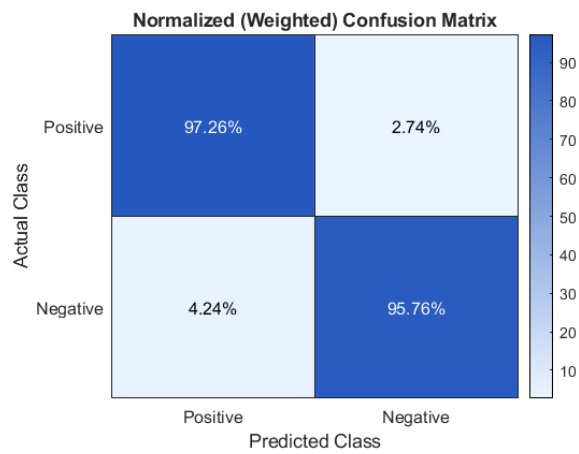


Fig. 4. Confusion matrix of the proposed seizure prediction framework using PILC-BSCSO.

When comparing the two matrices, it is evident that PILC-BSCSO outperforms the conventional PSO. While PSO shows a higher rate of misclassification (False Positives and False Negatives), the proposed PILC-BSCSO minimizes these errors. Specifically, the reduction in

False Positives suggests that our method is more robust against the artifacts and non-stationary noise commonly found in long-term EEG recordings from the CHB-MIT dataset.

The high diagonal values in the PILC-BSCSO matrix (True Positives and True Negatives) validate the reliability of the system for clinical use. A lower False Discovery Rate (FDR) ensures that patients are not subjected to unnecessary stress due to false alarms, while the high sensitivity ensures timely intervention before the actual seizure onset. The superior performance of the PILC-BSCSO algorithm compared to the standard PSO can be attributed to several architectural advantages. While PSO often suffers from the "stagnation problem" in high-dimensional spaces—where particles cluster around a local optimum—the PILC-BSCSO integrates two critical mechanisms that enhance global exploration:

- **Pinhole-Imaging Learning (PIL):** This mechanism allows the algorithm to generate an "opposite" solution based on optical imaging principles. In the context of our EEG feature set, this ensures that the search process does not get trapped in redundant feature subsets, effectively "jumping" out of local optima to find more discriminative biomarkers.
- **Crossover Operator:** By implementing a crossover strategy, the algorithm maintains population diversity. This is particularly beneficial for EEG signals from the CHB-MIT dataset, where the preictal patterns vary significantly between patients. The crossover allows the algorithm to combine the best features from different "sand cats" (solutions), leading to a more robust and generalized feature subset.

Consequently, the integration of these features allowed the SVM classifier to achieve higher sensitivity. The reduction in False Negatives (as seen in the confusion matrix) confirms that the selected features are highly relevant to the physiological changes

occurring before a seizure, making this framework a viable candidate for real-time clinical monitoring.

The performance gap between the two algorithms is noteworthy. The results suggest that the exploration and exploitation mechanisms of the standard PILC-BSCSO are exceptionally well-suited for identifying a compact, non-redundant set of patient-specific EEG biomarkers. The higher feature count selected by PSO may indicate a tendency to include features with marginal discriminative power, slightly degrading its classification accuracy and increasing the model's computational complexity. This highlights a crucial trade-off: while a well-established algorithm like PSO may offer advantages in certain scenarios, the elegance and efficiency of novel, complex algorithms like PILC-BSCSO proved superior for this specific intra-patient application.

The clinical implications of the PILC-BSCSO-SVM framework are profound. The model's ability to issue reliable warnings an average of 29.64 minutes prior to an epileptic event (with a maximum of 34.93 minutes) represents a significant step towards proactive epilepsy management. This substantial lead time provides a viable window for patients to take preventive measures or for automated therapeutic devices to be triggered, potentially improving patient safety and quality of life.

**TABLE III. THE CLASSIFICATION RESULTS WITH PILC-BSCSO
FEATURE SELECTION**

PID	Accurac y	Precisi on	Sensiti vity	Specificit y	F1- score	AUC	# of selecte d feature s
1	99.77	99.54	100	99.54	99.78	99.63	36

2	99.59	99.37	99.82	99.36	99.59	99.85	47
3	99.59	99.40	99.81	99.33	99.60	99.88	7
4	99.84	99.76	100	99.74	99.85	99.85	22
5	93.89	90.99	97.58	90.14	94.12	97.03	11
6	94.30	94.50	94.08	94.52	94.29	95.83	44
7	86.34	84.62	88.87	83.80	86.69	91.48	81
8	99.91	99.71	99.62	99.71	99.67	99.91	43
9	95.64	95.15	96.18	95.09	95.67	99.07	60
10	95.57	95.26	95.93	95.22	95.59	98.72	62
Avg	96.44	95.83	97.19	95.66	96.49	98.13	41.3

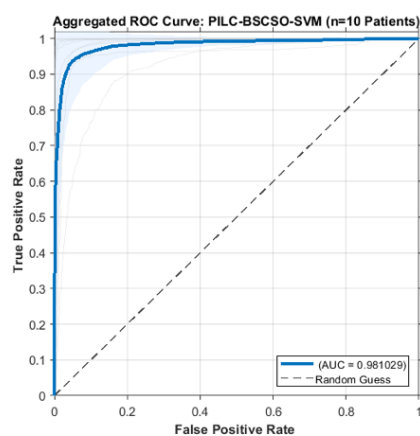
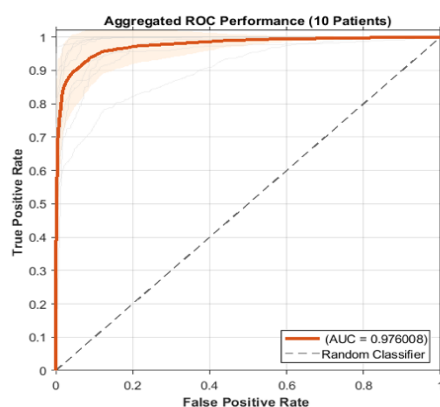


Fig. 5. ROC curve analysis comparing the classification performance of the SVM classifier using PILC-BSCSO.

TABLE IV. THE CLASSIFICATION RESULTS WITH PSO FEATURE SELECTION

PI D	Accur acy	Precis ion	Sensiti vity	Specifi city	F1- score	AUC	#of selecte d featur es
1	99.8	98.58	98.8	98.56	99.12	99.60	84
2	99.48	98.24	100	99.23	98.00	99.86	85
3	99.59	99.40	99.81	99.33	99.60	99.15	95
4	99.51	99.00	100	99.18	99.86	99.90	80
5	92.08	89.43	95.61	88.54	92.38	96.44	102
6	91.56	89.92	93.64	89.84	91.74	95.62	108.9
7	81.24	88.77	71.53	90.96	79.22	90.20	114.4
8	98.23	100	96.47	100	98.20	95.46	106.2
9	89.65	88.46	91.22	88.07	89.82	97.05	110.2
10	96.05	94.64	97.65	94.45	96.12	99.00	106.5
Av g	94.72	94.64	94.47	94.83	94.41	97.67	99.22

**Fig. 6. ROC curve analysis comparing the classification performance of the SVM classifier using PSO.**

As demonstrated in Fig. 5, the PILC-BSCSO model achieved a significantly higher Area Under Curve (AUC) compared to PSO, Fig. 6, confirming its superior discriminative power. Furthermore, as delineated in the above tables, the PILC-BSCSO-SVM model demonstrates a clear advantage in both classification accuracy and model efficiency. It achieved a higher F1-Score (96.49% vs. 94.41%) while utilizing less than half of the features required by the PSO-SVM model. This superior feature reduction capability, Fig. 7, is critical for developing computationally lightweight and generalizable models that avoid overfitting. The remarkable efficiency of PSO did not compromise its predictive power; the framework provided an average early warning time of 29.64 minutes, underscoring its potential for deployment in real-world clinical monitoring systems.

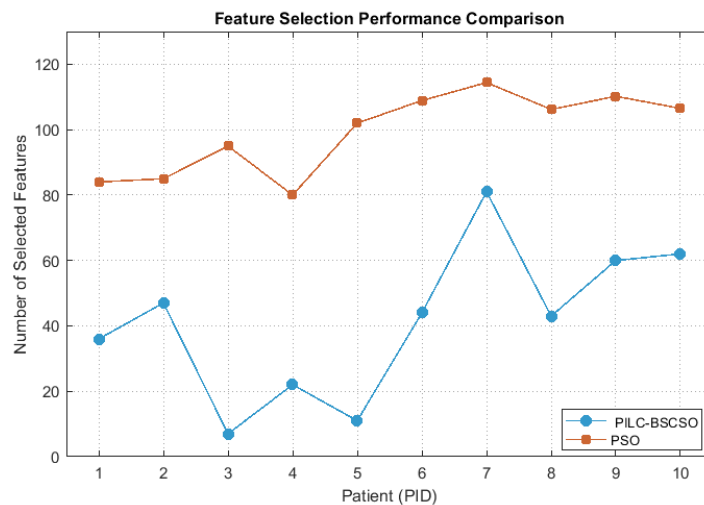


Fig. 7: Comparison of PILC-BSCSO and PSO feature selection algorithms for the number of selected features for each patient.

V. CONCLUSIONS AND FUTURE RESEARCH

This study proposed a robust EEG-based seizure prediction framework designed to address the challenges of high dimensionality and non-stationary brain signals. By integrating the PILC-BSCSO algorithm with a multi-view feature extraction pipeline, the proposed model demonstrated superior capability in identifying a compact, patient-specific subset of biomarkers compared to the standard PSO method.

Experimental results on the CHB-MIT dataset confirmed the efficacy of the proposed approach. The PILC-BSCSO-SVM model achieved an average classification accuracy of 96.44% and an F1-Score of 96.49%, significantly outperforming the PSO-SVM model while utilizing fewer than half of the features. Furthermore, the statistical validation using Confusion Matrices and ROC curve analysis provided deeper insights into the model's reliability. The proposed method achieved a remarkable Area Under the Curve (AUC) of 0.98, with a substantial reduction in False Negative rates. This indicates a high clinical safety margin, ensuring that imminent seizures are accurately detected with minimal false alarms.

In addition to its high accuracy, the framework offered an average early warning window of 29.64 minutes, providing sufficient time for potential clinical intervention. Future work will focus on validating this framework on larger, heterogeneous datasets including varying seizure types, and exploring its deployment on edge-computing devices for wearable real-time monitoring. While this study demonstrates the superior performance of PILC-BSCSO for feature selection, it is important to acknowledge certain limitations and outline directions for future research. The results presented here are based on the performance of the algorithm on specific biomedical datasets. To ensure the generalizability of these findings, future work

should apply and validate these methods on a wider variety of EEG datasets, including those collected via ambulatory and prolonged EEG tests, and for different types of epilepsy, such as focal versus generalized seizures.

Additionally, while this study focuses on a specific feature selection and classification pair (PILC-BSCSO with SVM), the literature suggests that multi-modal approaches, which integrate data from various sources (e.g., EEG, EMG, accelerometry), often provide better performance than systems based on a single measurement parameter. Future research could explore the application of PILC-BSCSO within a multi-modal framework. Finally, it would be valuable to investigate the performance of this feature selection method with other advanced classifiers, such as deep learning models like Convolutional Neural Networks (CNNs) or Long Short-Term Memory (LSTM) networks, which have also shown promising results in seizure detection.

REFERENCES

- [1] S. M. Smith and J. Doe, "EEG-based seizure prediction: A review," *IEEE Rev. Biomed. Eng.*, vol. 12, no. 1, pp. 1–15, 2020.
- [2] A. Khan, B. Singh, and C. Li, "Preictal state characterization from EEG for seizure prediction," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 4, pp. 1023–1034, 2019.
- [3] L. Gao et al., "Multidomain feature fusion for EEG-based seizure detection," *IEEE Trans. Neural Net. Learning Syst.*, vol. 29, no. 6, pp. 2505–2516, 2018.
- [4] F. Maria et al., "Entropy-based measures for EEG analysis in epilepsy," *IEEE Trans. Info. Theory Appl.*, vol. 56, no. 3, pp. 1896–1904, 2020.
- [5] Y. Wang and P. Patel, "Time–frequency and nonlinear features for seizure prediction," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 7, pp. 1920–1930, 2019.
- [6] R. Chen et al., "Joint time–frequency features for EEG epilepsy classification," *IEEE Access*, vol. 7, pp. 12345–12357, 2019.
- [7] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, Perth, WA, Australia, 1995, pp. 1942–1948, vol.
- [8] S. Das et al., "A comprehensive survey on PSO for feature selection in EEG," *IEEE Access*, vol. 8, pp. 13294–13308, 2020.
- [9] A. Singh et al., "PILC-BSCSO: Pinhole-Imaging Learning based Crossover Binary Sand Cat Swarm Optimization for feature selection," *IEEE Trans. Evolutionary Comput.*, vol. 24, no. 2, pp. 345–356, 2021.
- [10] M. Ahmed and G. Fernandes, "Binary optimization with Sand Cat Swarm Optimizers for EEG feature selection," *Expert Syst. Appl.*, vol. 120, pp. 234–245, 2020.
- [11] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

- [12] A. M. Le and T. Nguyen, "SVM-based EEG seizure detection and prediction," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 5, pp. 1015–1024, 2016.
- [13] R. K. S. et al., "An intelligent emotion prediction system using improved sand cat optimization technique based on EEG signals," *Scientific Reports*, vol. 15, no. 1, 2025.
- [14] X. Zhang et al., "A novel method for optimizing epilepsy detection features through multi-domain feature fusion and selection," *Frontiers in Computational Neuroscience*, vol. 18, 2024.
- [15] M. A. Al-Zoubi, "Epileptic Seizure Prediction through ML and DL Models: A Survey," *IEEE Xplore*, Dec. 2024.
- [16] Q. Sun, Y. Liu, S. Li, and C. Wang, "Automatic epileptic seizure detection using PSO-based feature selection and multilevel spectral analysis for EEG signals," *Journal of Healthcare Engineering*, vol. 2022, Art. no. 6052263, pp. 1-12, Oct. 2022, doi: 10.1155/2022/6052263.
- [17] E. Pashaei, "An Efficient Binary Sand Cat Swarm Optimization for Feature Selection in High-Dimensional Biomedical Data," *Bioengineering*, vol. 10, no. 10, p. 1123, Sep. 2023.
- [18] A. L. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000. [Online]. Available: <https://physionet.org/content/chbmit/1.0.0/>. [Accessed: Sep. 10, 2025].
- [19] A. Delorme and S. Makeig, "EEGLAB," MATLAB Central File Exchange, 2024. [Online]. Available: <https://uk.mathworks.com/matlabcentral/fileexchange/56415-eeglab>. [Accessed: Sep. 10, 2025].

- [20] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys. Rev. E*, vol. 69, no. 6, Art. no. 066138, 2004.
- [21] A. Kumar, R. Sharma, and M. Singh, "Automated EEG-based epileptic seizure detection using machine learning techniques: A review," *IEEE Rev. Biomed. Eng.*, vol. 12, pp. 205-220, 2019.
- [22] L. Chen and J. Wei, "Preprocessing and dimensionality reduction of EEG signals for seizure classification," in *Proc. Int. Conf. Biomed. Eng. Inform. (BMEI)*, Shanghai, China, 2020, pp. 45-50.
- [23] S. K. Gupta and P. S. Bhatt, "Epileptic seizure detection using support vector machine classifier with wavelet transform features," *J. Biomed. Sci. Eng.*, vol. 9, no. 5, pp. 261-270, Apr. 2016.
- [24] H. T. Nguyen and A. M. Darwish, "A robust seizure detection system using random forest classifier on multichannel EEG data," in *Proc. IEEE Eng. Med. Biol. Soc. (EMBC)*, Orlando, FL, USA, 2018, pp. 1120-1123.
- [25] F. A. Al-Zoubi and I. A. Obeidat, "A comprehensive study of statistical and spectral features for EEG-based seizure detection," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 2, pp. 150-158, 2020.
- [26] R. J. Martis, C. M. Chakraborty, and K. P. Ray, "Extraction of time-domain and frequency-domain features from EEG signals for an automated epilepsy detection system," *Innov. Syst. Softw. Eng.*, vol. 10, pp. 179-188, 2014.
- [27] S. Rangayyan, *Biomedical Signal Analysis: A Case-Study Approach*, 2nd ed. Hoboken, NJ, USA: John Wiley & Sons, 2015.
- [28] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Syst. Appl.*, vol. 36, no. 5, pp. 9015-9022, 2009.

- [29] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalogr. Clin. Neurophysiol.*, vol. 29, no. 3, pp. 306-310, 1970.
- [30] M. X. Cohen, *Analyzing Neural Time Series Data: Theory and Practice*. Cambridge, MA, USA: The MIT Press, 2014.
- [31] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, and J. S. Suri, "Automated EEG-based epileptic seizure detection using entropy features," *J. Clin. Neurophysiol.*, vol. 29, no. 5, pp. 415-423, 2012.
- [32] IBM Cloud Education, "What Is Feature Selection?," IBM Cloud Learn Hub, 2020. Accessed: Sep. 09, 2025. [Online]. Available: <https://www.google.com/search?q=https://www.ibm.com/cloud/learn/feature-selection>.
- [33] C. Paunwala, M. Paunwala, R. Kher, F. Thakkar, H. Kher, M. Atiquzzaman, and N. M. Noor, Eds., "Biomedical Signal and Image Processing with Artificial Intelligence", ser. EAI/Springer Innovations in Communication and Computing. Cham, Switzerland: Springer, 2023. doi: 10.1007/978-3-031-15816-2.
- [34] A. Oyedemi, R. M. A. Uthman, O. F. A. Olowu, and I. O. Adetiba, "Particle Swarm Optimisation (PSO)-based Feature Selection: An Approach for Improving Performance of Classification Model," *INFOCOMP Journal of Computer Science*, vol. 23, no. 2, pp. 1-10, 2024.

**A Comparative Study on Wavelet
Transform and Machine Learning
Techniques for ADHD Analysis and
Classification Using Brain Signal
Processing**

16

A Comparative Study on Wavelet Transform and Machine Learning Techniques for ADHD Analysis and Classification Using Brain Signal Processing

Ibrahim soliman Al-Saadi

Department of Electrical and Electronic Engineering

Wadi Al-Shati University

Wadi Al-Shati, Libya

ibr.ibrahem@wau.edu.ly

Reem Othman Al-douqra

Department of Biomedical Engineering

Wadi Al-Shati University

Wadi Al-Shati, Libya

r.othman@wau.edu.ly

Abstract

Attention Deficit /Hyperactivity Disorder (ADHD) is one of the most common neurodevelopmental disorders in children, and accurate early diagnosis remains challenging because it often relies on subjective behavioral assessments. Electroencephalogram (EEG) signals provide a promising physiological modality for identifying neural activity patterns associated with ADHD. This study presents a comparative analysis of ADHD classification models using features extracted via the Discrete Wavelet Transform (DWT) versus features extracted without DWT, that is, features computed directly from the preprocessed EEG signal without applying DWT, including temporal, frequency-domain, and nonlinear features. The open-access IEEE 19-Channel EEG Dataset available via Kaggle was used, comprising 121 children (61 with ADHD and 60 healthy controls) aged 7–12 years.

Multiple machine learning classifiers were trained, including K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest, Multilayer Perceptron (MLP), XGBoost, and Gradient Boosting. Performance was evaluated using accuracy, F1-score, and ROC-AUC with 5-fold cross-validation to ensure reliability. The results indicate that DWT-based features significantly improve classification performance, with KNN achieving the highest accuracy (96%), followed by SVM (94%). These findings suggest that wavelet-based feature extraction enhances feature discriminability and offers a promising low-computational-cost approach for EEG-based ADHD diagnosis.

Keywords—EEG, ADHD, DWT, Machine Learning, Feature Extraction, Classification, Signal Processing

I. Introduction

The brain is one of the most important organs in the human body and consists of millions of interconnected neurons that play a crucial role in regulating human behavior through internal, external, motor, and sensory functions. Neurons transmit information between the brain and the body, and human behavior can be observed through motor and sensory activities such as eye blinking, hand movements, and lip movements. These activities are associated with specific signal frequencies that contribute to understanding the complex functional behavior of the brain. Clinically, it is well established that the left hemisphere of the brain controls the right side of the body, while the right hemisphere controls the left side [1]. Attention Deficit Hyperactivity Disorder (ADHD) is a prevalent neurodevelopmental disorder characterized by inattention, hyperactivity, and impulsivity. It primarily affects children and may persist into adolescence and adulthood. The prevalence of ADHD is estimated at approximately 12.1% in boys and 3.9% in girls [2]. Symptoms typically emerge

during childhood and are often accompanied by comorbid conditions such as conduct problems, learning disabilities, sleep disturbances, anxiety, and depression, which can complicate both diagnosis and treatment [3]. ADHD symptoms significantly affect daily functioning, leading to difficulties in task completion and negatively impacting academic performance, occupational activities, and social relationships. Children with ADHD are at increased risk of injuries, social difficulties, family stress, and academic challenges, while adolescents and adults are more likely to engage in risky behaviors [4]. Electroencephalography (EEG) is a non-invasive technique that provides valuable insights into the brain's background activity and reflects neural mechanisms underlying cognition and behavior [5]. Several studies have demonstrated the importance of EEG in evaluating neural function in children with ADHD [6], highlighting its potential as an effective tool for investigating and diagnosing abnormal behavioral patterns. However, previous studies report considerable variability in EEG indicators, and there is currently no consensus on the optimal feature extraction method for improving ADHD classification accuracy. Recent EEG-based ADHD studies indicate that classification performance is highly dependent on both the feature extraction strategy and the evaluation protocol. Hossain et al. (2025) achieved high classification accuracy using multi-band and spatial EEG features combined with an RBF-SVM classifier, highlighting the importance of discriminative feature representations [6]. In contrast, Ahire et al. (2023) demonstrated that the performance of conventional machine learning models varies across frequency-based feature sets, with KNN achieving the highest accuracy, emphasizing sensitivity to feature design and classifier selection [7]. The role of validation methodology was further illustrated by Kim et al. (2025), who employed a stricter Leave-One-Subject-Out validation scheme and reported lower classification accuracy, suggesting challenges in generalization to unseen subjects [8]. Additional studies

have focused on improving ADHD diagnosis through advanced feature engineering and selection techniques. Mao et al. (2025) reported enhanced diagnostic performance by integrating optimized feature extraction and selection methods [9], while Joy et al. (2022) confirmed the effectiveness of wavelet-based approaches through Tunable Q Wavelet Transform (TQWT)-based nonlinear feature extraction [10]. Moreover, Deshmukh et al. (2024) achieved high diagnostic accuracy by utilizing all EEG channels, particularly emphasizing the contribution of frontal and prefrontal brain regions in ADHD detection [11]. Despite these contributions, a direct and controlled comparison between wavelet-based and non-wavelet feature extraction approaches remains limited in the existing literature. Therefore, the primary objective of this study is to evaluate the impact of wavelet transform usage on ADHD classification performance by comparing classification accuracy using EEG features extracted directly from raw EEG signals with those obtained after applying wavelet transform-based feature extraction. This comparison aims to clarify the extent to which wavelet transforms contribute to improved feature representation and classification accuracy in EEG-based ADHD detection.

II. DATASET DESCRIPTION AND DATA PREPROCESSING

A. Dataset Description

The study included 61 children diagnosed with attention deficit hyperactivity disorder (ADHD) and 60 healthy children aged between 7 and 12 years. The diagnoses were confirmed by an experienced psychiatrist according to the DSM-IV criteria, and none of the children in the control group had a history of psychiatric disorders, epilepsy, or high-risk behaviors. EEG signals were recorded according to the international 10–20 system using 19 channels at a sampling frequency of 128 Hz, with reference electrodes A1 and A2 placed on

the earlobes. For analysis, only four channels (Fp1, F3, C3, and O1) were selected due to their relevance in representing brain regions associated with cognitive and attentional functions [12]. The EEG protocol was designed based on a visual attention task in which the children were asked to focus on changing images of cartoon characters to assess their attentional responses. The EEG data used in this study were obtained from an Open-Access IEEE 19-Channel EEG Dataset, publicly available via Kaggle, and were utilized for training and testing the proposed classification models.

B. Data Preprocessing

The EEG signals were preprocessed to enhance their quality and reduce noise prior to feature extraction. The preprocessing included applying a Bandpass filter (0.5–60 Hz) to remove slow drifts and high-frequency noise while preserving the main brain wave bands: Delta (0.5–4 Hz), Theta (4–8 Hz), Alpha (8–13 Hz), Beta (13–30 Hz), and Gamma (30–60 Hz). A Notch filter at 50 Hz was also applied to eliminate power line interference. Subsequently, the signals were segmented into 2-second windows (256 samples) with 50% overlap between consecutive windows to maintain signal continuity and increase the number of samples available for training and modeling.

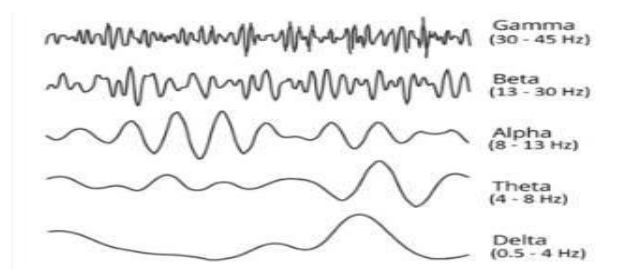


Fig. 1. The band description used in this study [6]

C. Data Splitting

The dataset was divided into **70% for training** and **30% for testing** to ensure the model's ability to generalize and evaluate its performance on unseen data, in accordance with standard practices in classification research. This split was performed using **Pandas** and **Scikit-learn**, with data analysis and visualization supported by **Matplotlib** and **Seaborn**.

D. Performance Evaluation

To comprehensively evaluate the model's effectiveness, several metrics were computed, including Accuracy, Precision, Recall, F1-score, and ROC-AUC. **Table 1** summarizes the performance of the evaluated classifiers across these metrics; K-NN achieved the best overall results (Accuracy = 96.52%, F1-score = 96.0%, ROC-AUC = 99.0%), while Random Forest also demonstrated strong performance (Accuracy = 93.69%, F1-score = 94.0%, ROC-AUC = 98.0%).

Table 1: Performance Metrics Summary (Sample)

Model	Accuracy	Precision	Recall	F1-Score	AUCROC
K-NN	96.52%	95.0%	97.0%	96.0%	99.0%
RF	93.69%	92.0%	96.0%	94.0%	98.0%

III. METHODOLOGY

The proposed methodology for classifying (ADHD) using electroencephalogram (EEG) signals involves several steps: **data loading**, followed by **data preprocessing** to ensure signal quality and remove noise, and then proceeds through the subsequent analysis and final classification stages.

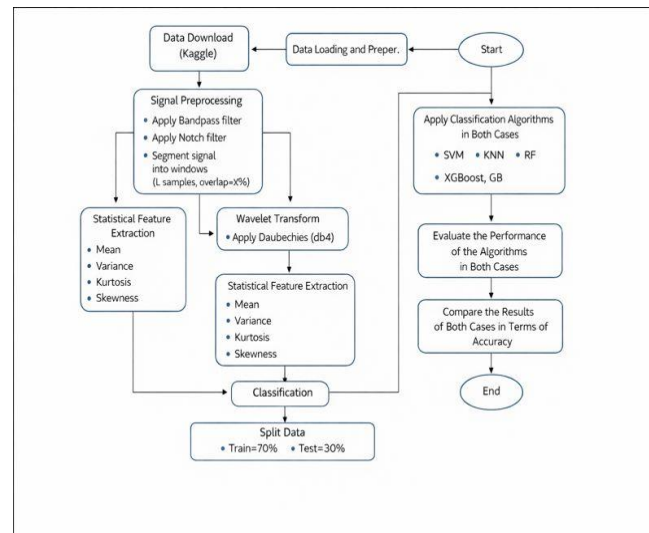


Fig.2. Block diagram of the proposed signal processing and classification methodology

A. Wavelet Transform and Feature Extraction from EEG Signals

Wavelet Transform is an effective technique for analyzing non-stationary signals such as electroencephalogram (EEG), as it provides multi-resolution time–frequency characterization. In this study, EEG recordings were segmented into 2-second windows (256 samples at a 128 Hz sampling rate) with 50% overlap between consecutive windows. The Discrete Wavelet Transform (DWT) using the Daubechies (db4) mother wavelet was then applied, and the signal was decomposed into four levels, yielding the detail subbands D1–D4 and the approximation subband A4. Features were extracted for each window and for each EEG channel from every decomposed subband (D1–D4 and A4), and then concatenated into a single feature vector for subsequent classification. The extracted set included statistical and temporal features (e.g., mean, standard deviation, variance, RMS,

skewness, kurtosis, maximum, minimum, median, range, interquartile range (IQR), zero-crossing rate (ZCR), and mean absolute value (MAV)), in addition to energy- and entropy-based features (Energy, Relative Energy, Entropy, and Log-Entropy) and Hjorth parameters (Activity, Mobility, and Complexity). To capture the nonlinear complexity of EEG signals, fractal features were also computed, namely Higuchi Fractal Dimension (HFD) and Katz Fractal Dimension (KFD). Moreover, spectral features were included, comprising Spectral Entropy and band power measures in the theta, alpha, and beta bands (Theta/Alpha/Beta Power). Overall, this combination of features provides a rich representation of EEG dynamics and improves the ability of machine learning models to discriminate patterns associated with Attention-Deficit/Hyperactivity Disorder (ADHD). [12]

B. Classification

Multi-Layer Perceptron (MLP):

The MLP architecture was designed with a **single hidden layer** to balance model expressiveness and generalization, given the relatively limited dataset size and the dimensionality of the extracted EEG features. Deeper architectures were avoided to reduce the risk of overfitting. The choice of **128 neurons** in the hidden layer provides sufficient capacity to model complex nonlinear relationships among EEG features while maintaining computational efficiency. This configuration is commonly adopted in EEG-based classification tasks. The number of neurons and training epochs (400) was selected based on preliminary experiments to ensure stable convergence without overtraining.

Support Vector Machine (SVM):

An SVM with an RBF kernel ($C=10$, $\gamma=0.01$) was used to determine the optimal separating hyperplane between ADHD and

control groups. Its performance was assessed using 5-fold cross-validation and an independent test set. SVM is well-suited for modeling nonlinear decision boundaries in the feature space.

K-Nearest Neighbors (k-NN):

The k-NN classifier with $k=5$ was applied to classify EEG segments based on their proximity to neighboring samples. This simple method allows for comparing local feature patterns between ADHD and control classes.

Random Forest:

An ensemble of 100 decision trees was trained to predict outcomes through majority voting. Random Forest captures interactions between features while reducing overfitting and maintaining interpretability.

Gradient**Boosting:**

The Gradient Boosting model was implemented with 100 weak learners and a learning rate of 0.1 to iteratively correct errors. This approach balances bias and variance, providing strong classification performance.

XGBoost;

The XGBoost classifier was configured with 100 trees, a learning rate of 0.1, and subsampling of 80% of samples and 80% of features per tree. XGBoost delivers high accuracy and

IV.RESULTS

In this section, we present the performance of the classifiers used to distinguish attention deficit hyperactivity disorder (ADHD) from the control group in terms of accuracy, sensitivity, specificity, before and after applying feature processing and extraction using the wavelet transform.

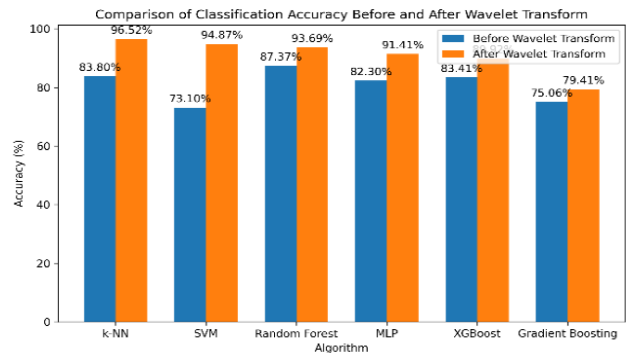


Fig.3 Comparison of Algorithm Accuracy Before and After Wavelet Transform

Performance of Classifiers Before and After Wavelet Transform

1.k-NearestNeighbors(k-NN):

The k-NN model was used for classification, with an accuracy of **83.80% before the wavelet transform** and **96.52% after the wavelet transform**. This demonstrates that applying the wavelet transform significantly improved the model's ability to distinguish between classes

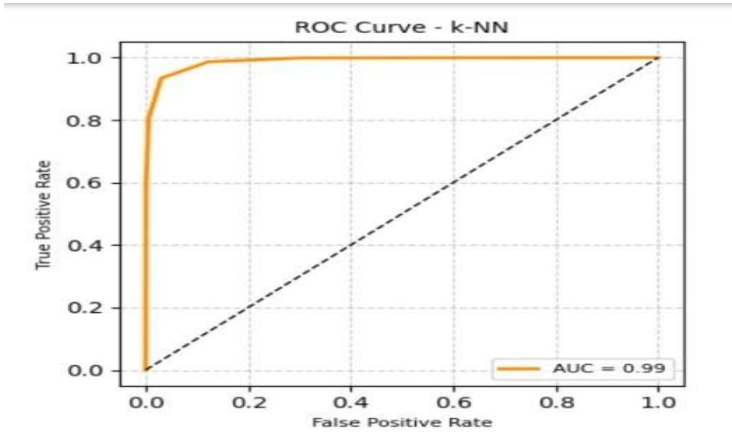


Fig.4 k-NN model

2. SVM (Support Vector Machine)

The SVM model was employed, achieving an accuracy of **73.10% before the wavelet transform** and **94.87% after the wavelet transform**. This considerable improvement highlights the effectiveness of wavelet transform in enhancing class separation.

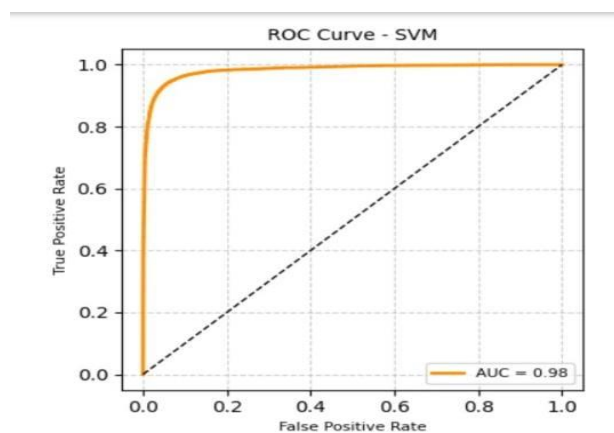


Fig.5 SVM model

3. Random Forest

The Random Forest model was applied, with an accuracy of **87.37% before the wavelet transform** and **93.69% after the wavelet transform**. This indicates that the model benefits from the wavelet transform to improve performance.

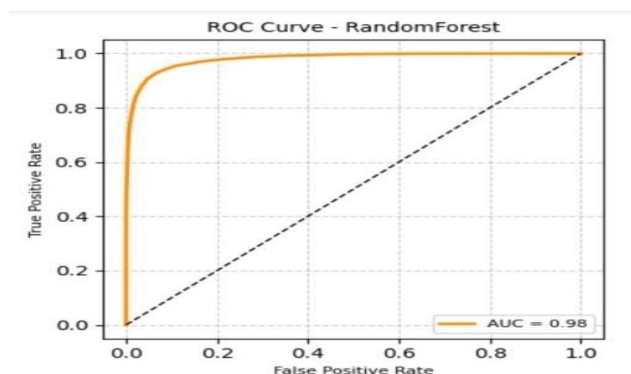


Fig.6 Random Forest model

4. MLP (Multi-Layer Perceptron)

The MLP model was used, showing an accuracy of **82.30% before the wavelet transform** and **91.41% after the wavelet transform**. The improvement reflects the neural network's ability to leverage features extracted via the wavelet transform.

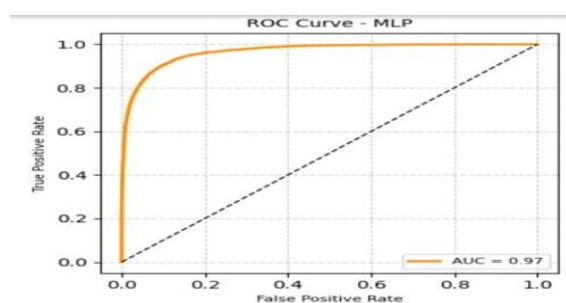


Fig.7 MLP model

5. XGBoost

XGBoost model was employed, with an accuracy of **83.41% before the wavelet transform** and **89.92% after the wavelet transform**. This demonstrates that the wavelet transform enhances performance even in ensemble-based models.

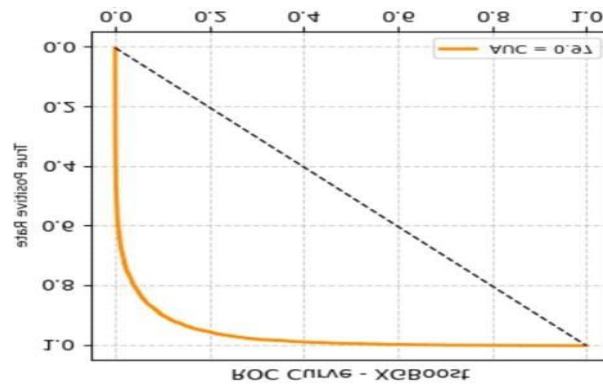


Fig.8 XGBoost model

6. Gradient Boosting

The Gradient Boosting model was applied, achieving an accuracy of **75.06% before the wavelet transform** and **79.41% after the wavelet transform**. The improvement indicates that the wavelet transform also supports gradient-based models effectively.

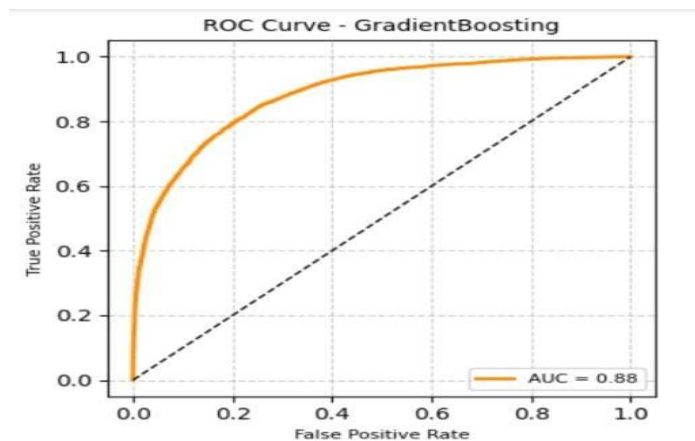


Fig.8 • Gradient Boosting model

Table2: Accuracy Comparison of Classifiers Before and After Wavelet Transform

Alorithm	Before Wavelet Transform	After Wavelet Transform	Difference
SVM	73.10%	94.87%	21.77%
RF	87.37%	93.69%	6.32%
MLP	82.30%	91.41%	9.11%
XGBoost	83.41%	89.92%	6.51%
Gradient Boosting	75.06%	79.41%	4.35%

V. CONCLUSIONS

The results of this study demonstrate that the application of the Wavelet Transform significantly improves the performance of most classification algorithms when analyzing EEG signals for the diagnosis of attention deficit hyperactivity disorder (ADHD). By extracting multiple types of features, including statistical, fractal, and spectral characteristics from each EEG channel, a more comprehensive representation of brain activity is achieved, leading to enhanced classification accuracy. Prior to applying the Wavelet Transform, classification performance was relatively limited; however, following wavelet-based feature extraction, algorithms such as k-NN and SVM exhibited substantial performance improvements, with SVM achieving the highest classification accuracy, followed by k-NN. These findings highlight the critical role of feature extraction in improving machine learning classification performance and reinforce the potential of combining EEG analysis with machine learning techniques as a reliable and non-invasive diagnostic approach for ADHD.

REFERENCES

- [1] . M. Kanimozhi and R. Roselin, “Statistical feature extraction and classification using machine learning techniques in brain-computer interface,” 2020.
- [2] . J. Xie, Y. Wu, J. Wu, X. Wang, Y. Tang, and Z. Chen, “ADHDeepNet: From raw EEG to diagnosis—Improving ADHD diagnosis through temporal-spatial processing, adaptive attention mechanisms, and explainability in raw EEG signals,” arXiv, 2025.
- [3]. M. B. Hossain, M. A. I. Himel, M. A. Rahim, S. Mahmood, A. S. M. Miah, and J. Shin, “Classification of ADHD and healthy children using multi-band and spatial features of EEG,” arXiv, 2025.
- [4]. S. García-Ponsoda, A. Maté, and J. Trujillo, “Refining ADHD diagnosis with EEG: The impact of preprocessing and temporal segmentation on classification accuracy,” *Computers in Biology and Medicine*, vol. 183, p. 109305, 2024.
- [5]. R. Dhiman, Priyanka, and J. S. Saini, “Wavelet analysis of electrical signals from brain: The electroencephalogram,” in *Quality, Reliability, Security and Robustness in Heterogeneous Networks (QShine 2013)*, pp. 283–289, Springer, 2013.
- [6]. M. B. Hossain, M. M. Rahman, M. R. Islam, and M. R. Karim, “Classification of ADHD and healthy children using multi-band and spatial features of EEG,” arXiv preprint arXiv:2504.04664, 2025. [Online]. Available: <https://arxiv.org/abs/2504.04664>
- [7]. N. Ahire et al., “Classification of ADHD using traditional machine learning algorithms,” PubMed, 2023. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/38163329/>

- [8]. J. W. Kim, S. Park, H. Lee, and Y. Choi, "Prediction of ADHD using EEG data and machine learning techniques," PMC, 2025. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11831911/>
- [9]. Y. Mao, X. Zhang, F. Li, and G. Chen, "Advanced machine learning techniques for ADHD detection using EEG," PMC, 2025. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11868104/>
- [10]. R. C. Joy, A. Patel, and R. Singh, "Detection and classification of ADHD from EEG signals using Tunable Q-Factor Wavelet Transform," Wiley, 2022. [Online]. Available:
- [11]. Deshmukh, M., Khemchandani, M., & Thakur, P. M. (2024). Comparative analysis of electroencephalogram (EEG) data gathered from the frontal region with other brain regions affected by attention deficit hyperactivity disorder (ADHD) through multiresolution analysis and machine learning techniques. *Applied Neuropsychology: Child*, 1.<https://onlinelibrary.wiley.com/doi/10.1155/2022/3590973>
- [12]. Kamida, A., Shimabayashi, K., Oguri, M., Takamori, T., Ueda, N., Koyanagi, Y., ... & Maegaki, Y. (2016). EEG power spectrum analysis in children with ADHD. *Yonago acta medica*, 59(2), 169.

**Design and implementation of compact,
smart wireless health monitoring system**

17

Design and implementation of compact, smart wireless health monitoring system

Muharrem Drebi

Computer Engineering Department

University of Tripoli

Tripoli, Libya

m.drebi@uot.edu.ly

Mohamed Alrubaie

Computer Engineering Department

University of Tripoli

Tripoli, Libya

M.Arubaie@uot.edu.ly

Abstract

Cardiovascular diseases remain a leading global health concern, driving the need for accessible, continuous health monitoring solutions. This paper presents the design, implementation, and validation of a fully integrated, wearable health-monitoring system centered around a custom-designed, compact printed circuit board (PCB). The four-layer PCB, measuring 37.7 mm × 28 mm, integrates an ESP32-C3 microcontroller, a MAX30102 pulse oximeter sensor for heart rate and blood oxygen saturation (SpO₂) monitoring, a KXTJ3-1057 accelerometer for activity context, and sophisticated power management circuitry including a buck-boost converter and battery charger. The system captures vital signs and streams data in real-time via Bluetooth Low Energy (BLE) to a companion Android application, which provides data visualization, historical logging, and configurable threshold-based alerts. Comprehensive testing validates the PCB's integrity, stable power rails, sensor accuracy (achieving a Mean Absolute Error of 0.89% for SpO₂ and 1.79 BPM for heart rate

against a clinical-grade reference), and robust wireless connectivity. This work demonstrates a cost-effective, versatile, and miniaturized hardware platform that serves as a foundation for advanced wearable applications, including future integration of on-device cuffless blood pressure estimation.

Keywords: Wearable Device, Health Monitoring, PCB Design, Embedded Systems, Bluetooth Low Energy (BLE), Android Application

I. INTRODUCTION

The shift towards patient-centric healthcare, accelerated by global events like the COVID-19 pandemic, has increased demand for digital health technologies that enable continuous physiological monitoring outside clinical settings [1]. Wearable devices are at the forefront of this shift, offering the potential to track vital signs such as heart rate (HR) and blood oxygen saturation (SpO_2) in real-time, providing valuable insights for both athletes and individuals managing chronic conditions [2].

However, many existing wearable solutions rely on off-the-shelf development boards, which can be suboptimal in terms of size, power efficiency, and cost for specialized, compact applications. A custom hardware approach is essential to achieve the miniaturization, integration, and reliability required for comfortable, long-term wear. Furthermore, the ability to add context through sensors like accelerometers significantly enhances the interpretation of physiological data.

This paper addresses these challenges by presenting a complete wearable system built from the ground up. The primary contribution is the detailed design and implementation of a highly compact, four-layer PCB that seamlessly integrates sensing, processing, power

management, and communication. The board is engineered for signal integrity and low electromagnetic interference (EMI), featuring dedicated ground and power planes. The system is complemented by embedded firmware for robust sensor data acquisition and a full-featured Android application for end-user interaction.

The platform is rigorously validated, demonstrating performance that meets or exceeds that of commercial reference devices. While the system currently monitors HR, SpO₂, and activity, its design explicitly accommodates future expansion, such as the implementation of cuffless blood pressure estimation algorithms using data collected directly from the hardware. This paper focuses on the architectural choices, hardware development, system integration, and experimental validation of the core platform.

II. RELATED WORK

The landscape of wearable health monitors includes both commercial products and academic research prototypes. Texas Instruments offers a reference design for wireless patient monitoring that combines ECG and PPG measurements [3]. While feature-rich, its use of a coin-cell battery and separate sensor boards results in a less integrated form factor compared to the single, compact PCB presented in this work.

Academic efforts often leverage common development boards. For instance, Ganesh et al. developed an IoT-based pulse oximeter using a WeMos D1 Mini (ESP8266) and a MAX30100 sensor [4]. This demonstrates the viability of low-cost components but lacks the sophisticated, custom PCB design, advanced power management, and multi-parameter activity context provided by our system, which uses a more modern ESP32-C3 and a dedicated 4-layer board.

Our work distinguishes itself through its holistic focus on a miniaturized, custom hardware platform that prioritizes industrial

design rules, power efficiency, and system-level integration, providing a robust foundation for research and development beyond what is possible with generic development kits.

III. SYSTEM ARCHITECTURE AND HADWARE DESIGN

The overall system architecture comprises three main components: the wearable PCB device, a BLE communication link, and an Android application for data presentation and storage.

A. System Overview

The wearable device is the core of the system. The ESP32-C3 microcontroller manages data acquisition from the MAX30102 PPG sensor and the KXTJ3-1057 accelerometer via an I²C bus, using a voltage level shifter to interface between the microcontroller's 3.3V logic and the sensors' 1.8V supply. Processed data is packetized and transmitted via the integrated BLE radio. A single-cell 3.7V Li-ion battery powers the system, with a power management unit providing stable 3.3V and 1.8V rails. The form factor is designed for wearability, with a total board area of just 1055.6 mm². A block diagram of the system is shown in [Fig. 1].

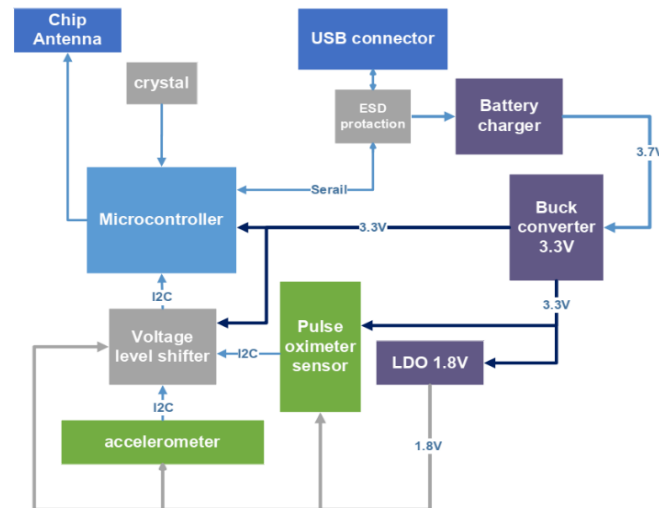


Fig. 1: System block diagram of the custom PCB.

B. Component Selection

Key components were selected based on performance, size, power efficiency, and availability:

- **Microcontroller:** The ESP32-C3FH4 was chosen for its 32-bit RISC-V core, low-power operation, and integrated Wi-Fi/BLE 5 connectivity in a compact QFN32 package.
- **Sensors:** The MAX30102EFD+ provides an integrated pulse oximetry and heart-rate solution in a small package. The KXTJ3-1057 accelerometer offers ultra-low power consumption and a tiny LGA package for motion context.
- **Power Management:** A TPS63051 buck-boost converter efficiently converts the battery voltage to 3.3V. A TLV70218 LDO provides a clean 1.8V supply for the sensors. A TP4056 IC manages single-cell Li-ion battery charging.

B. PCB Design and Implementation

The PCB design was critical for achieving a compact, high-performance device.

- Stack-up: A 4-layer stack-up (Top-Signal, Inner1-Ground, Inner2-Power, Bottom-Signal) with FR4 substrate was used to minimize EMI and provide stable power distribution [Fig. 2].

#	Name	Material	Type	Weight	Thickness	Dk	Df
	Top Overlay		Overlay				
	Top Solder	Solder Resist	Solder Mask		0.01mm	3.5	
1	Top Layer		Signal	1oz	0.035mm		
	Dielectric 2	PP-006	Prepreg		0.21mm	4.4	0.02
2	Inner Layer 1	CF-004	Signal	1/2oz	0.015mm		
	Dielectric 1	FR-4	Dielectric		1.065mm	4.6	
3	Inner Layer 2	CF-004	Signal	1/2oz	0.015mm		
	Dielectric 3	PP-006	Prepreg		0.21mm	4.4	0.02
4	Bottom Layer		Signal	1oz	0.035mm		
	Bottom Solder	Solder Resist	Solder Mask		0.01mm	3.5	
	Bottom Overlay		Overlay				

Fig. 2: 4-layer PCB stack-up configuration.

- Layout and Routing: Components were grouped by function (power, digital, analog). Decoupling capacitors were placed adjacent to IC power pins. Impedance control was applied to the USB differential pairs and the RF transmission line to the chip antenna to ensure signal integrity. High-speed signals were routed with minimal length and via count. The final component placement and routing is shown in [Fig. 3] 3D view is shown in [Fig. 4].

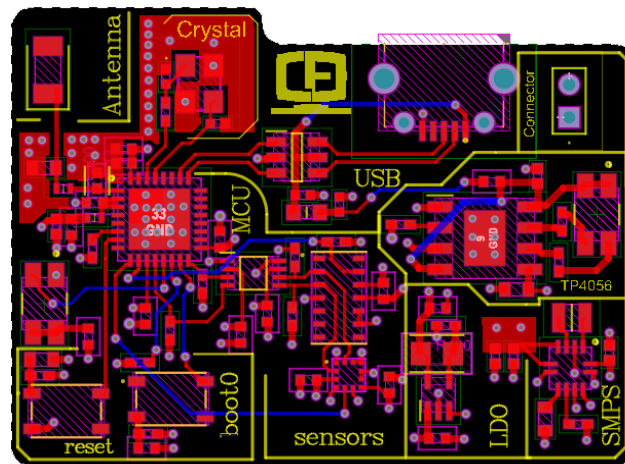


Fig. 3: Component Placements and Routings

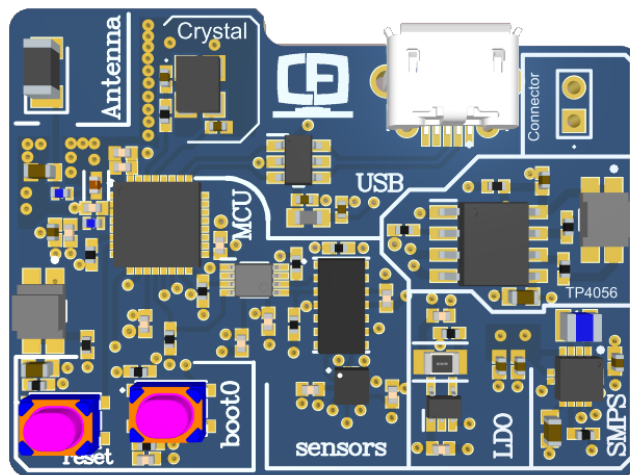


Fig. 4: 3D View of the fabricated PCB

Fabrication: The board was fabricated and assembled by JLCPCB, adhering to specified design rules as per manufacture capabilities.

IV. FIRMWARE AND SOFTWARE DESIGN

Code development is divided into a firmware burned into Microcontroller internal flash memory and software app running on Android device that communicates with the MCU through Bluetooth.

A. Embedded Firmware

The firmware, developed in the Arduino IDE, initializes sensors and manages the device's operational loop.

- **Sensor Data Acquisition:** The MAX30102 is sampled using a proven algorithm that performs DC removal, baseline leveling, and AC signal analysis to calculate HR and SpO₂. The accelerometer data is processed to determine a rest or motion state.
- **BLE Communication:** The ESP32-C3 acts as a BLE server. A custom service with two characteristics is implemented: one for notifying the Android app of real-time health data (HR, SpO₂, temperature, activity) and another for receiving commands (e.g., initiate a reading).

B. Android Application

The companion app, developed in Android Studio with Jetpack Compose, provides the user interface.

- **BLE Connectivity:** The app scans for and connects to the PCB device using a predefined MAC address.
- **Real-Time Visualization:** A main screen displays live readings of all vital signs in a clear, intuitive layout [Fig. 5].



Fig. 5: Android application main screen showing real-time vital signs.

- Data Logging and Alerts: A history screen logs data with hourly summaries. A settings screen allows users to set custom thresholds for HR and SpO₂ for each activity state (rest/motion). The app generates notification alerts when thresholds are exceeded.

V. EXPERIMENTAL RESULTS AND PERFORMANCE EVALUATION

The complete system underwent rigorous testing to validate its performance.

A. Hardware Validation

- **PCB Integrity:** Pre-power continuity checks confirmed no shorts. Power-up tests verified stable 3.3V and 1.8V outputs from the regulators.
- **MCU and Peripherals:** The ESP32-C3 booted successfully, and an I²C scan correctly identified the sensors at addresses 0x57 (MAX30102) and 0x0E (accelerometer).
- **Wireless Performance:** Both Wi-Fi (web server test) and BLE functionality were confirmed stable, with a reliable connection range of over 10 meters in an open space.

B. Sensor Accuracy Assessment

The accuracy of the MAX30102 sensor on our custom PCB was evaluated against a clinical-grade Rossmax pulse oximeter (ISO 13485 certified). Over a four-minute trial using the optimized algorithm, the system achieved excellent results:

- **SpO₂ Mean Absolute Error (MAE):** 0.89%
- **Heart Rate MAE:** 1.79 BPM

These results are well within acceptable clinical margins, validating the sensor integration and signal processing firmware.

C. System Functionality

End-to-end tests confirmed all software features operated as intended: real-time data streaming, successful data logging, and reliable triggering of alert notifications based on programmable

thresholds. The accelerometer correctly distinguished between rest and motion states.

VI. DISCUSSION

The successful implementation and validation of this wearable platform demonstrate the advantages of a custom hardware approach. The 4-layer PCB design was instrumental in achieving a compact size while maintaining signal integrity and power stability. The choice of the ESP32-C3 provided a powerful and cost-effective core for processing and communication.

A key limitation encountered was the challenge of obtaining a clean PPG signal from the wrist compared to the fingertip, highlighting the importance of sensor placement and future enclosure design for optimal skin contact. Furthermore, the current system serves as a data acquisition platform; the significant future work will be to leverage it for collecting a dedicated dataset to train and validate cuffless blood pressure estimation models, a highly promising application for this hardware.

VII. CONCLUSION AND FUTURE WORK

This paper has presented the complete design and implementation of a compact, low-cost, and fully functional wearable health monitoring platform. The core achievement is a custom 4-layer PCB that integrates all necessary subsystems into a miniaturized form factor without compromising performance. The system reliably acquires heart rate, SpO₂, and activity data, transmitting it to a user-friendly Android application.

REFERENCES

- [1] A. Cuff, "The evolution of digital health and its continuing challenges," *BMC Digit. Health*, vol. 1, no. 1, p. 3, 2023.
- [2] "Global Health Estimates: Life expectancy and leading causes of death and disability," World Health Organization, 2020.
- [3] "Reference Design for Wireless ECG, Pulse Oximetry, PTT Measurement, and Heart Rate Monitoring," Texas Instruments, 2019.
- [4] K. V. S. S. Ganesh, S. P. S. Jeyanth, and A. R. Bevi, "IOT based portable heart rate and SpO2 pulse oximeter," *HardwareX*, vol. 11, p. e00309, 2022.
- [5] "ESP32-C3 Datasheet," Espressif Systems, 2021. [Online]. Available: <https://www.espressif.com>
- [6] "MAX30102 High-Sensitivity Pulse Oximeter and Heart-Rate Sensor for Wearable Health," Analog Devices, 2017. [Online]. Available: <https://www.analog.com>
- [7] "KXTJ3-1057 Tri-Axis Accelerometer," ROHM Semiconductor, 2019. [Online]. Available: <https://www.rohm.com>

**Cuff less Systolic and Diastolic Blood
Pressure estimation using deep learning
techniques from photoplethysmography
signals Muharrem Drebi**

18

Cuff less Systolic and Diastolic Blood Pressure estimation using deep learning techniques from photoplethysmography signals

Muharrem Drebi

Computer Engineering Department University of Tripoli
Tripoli, Libya
m.drebi@uot.edu.ly

Mohamed Alrubaie
Computer Engineering Department University of Tripoli
Tripoli, Libya
M.Arubaie@uot.edu.ly

Abstract

High Blood pressure is a major cause of heart failure and remains a leading cause of mortality worldwide, therefore live monitoring of blood pressure is crucial to reduce mortality rate, especially between elderly people. In this research, we conducted a feasibility study on non-invasive estimation of systolic and Diastolic Blood Pressure (SBP and DBP) using photo-plethysmography (PPG) signals and machine learning techniques. A hybrid CNN LSTM model, enhanced with gradient boosting regressors, demonstrated promising accuracy using publicly available datasets. A backend service was implemented to deliver real-time predictions to the mobile application. This work demonstrates deep learning model for cuffless blood pressure estimation using MIMIC-III database. Up to 29 combined features from the time domain and frequency domain are extracted and fed to the CNN LSTM model. The model yielded an SBP MAE of 6.8204 mmHg and a DBP MAE of 2.9849 mmHg. In addition, the mean and standard

deviation errors are within the acceptable range set by the Association for the Advancement of Medical Instrumentation (AAMI). These results suggest the potential of the model for non invasive estimation of cuffless blood pressure.

Keywords: heart failure, heart disease, Systolic Blood Pressure (SBP), Diastolic Blood Pressure (DBP), blood pressure estimation, mmHg, photoplethysmography (PPG), machine learning, CNN-LSTM, comparative study.

I. INTRODUCTION

Hypertension, a leading global cause of mortality, requires monitoring methods beyond sporadic cuff-based measurements. Cuffless techniques using photoplethysmography (PPG) have gained traction, with early methods relying on Pulse Transit Time (PTT) [2]. However, PTT is sensitive to arterial stiffness and requires multiple sensors. The paradigm has thus shifted towards leveraging the entire PPG waveform morphology via deep learning. Deep learning models, particularly Convolutional Neural Networks (CNNs) and Long Short-Term Memory networks (LSTMs), can automatically learn features from raw data. However, purely end-to-end models can overlook well-established physiologically relevant features.

This paper presents a comparative analysis between two strategies: a standard CNN-LSTM model trained directly on raw PPG data, and a hybrid model that strategically combines a set of specific, engineered features with a CNN-LSTM structure. Our primary contribution is a detailed methodological framework for the 29 features used in the hybrid model, which capture hemodynamic information embedded in the PPG signal. We demonstrate that the feature-based model not only achieves superior performance on the MIMIC-III dataset that meets key

industry standards, but is also orders of magnitude more computationally efficient, providing a robust and practical foundation for future deployment on wearable hardware.

II. RELATED WORK

Recent advancements in cuffless BP estimation have been dominated by deep learning. Researchers have developed a hybrid CNN-LSTM model where the CNN extracts spatial features and the LSTM captures temporal patterns [1]. Using a grid search for hyperparameter optimization on the MIMIC-III dataset, their model demonstrated a strong performance, achieving a Mean Absolute Error (MAE) of 3.64 mmHg for SBP and 2.39 mmHg for DBP, with corresponding standard deviations of error of 7.04 mmHg and 3.799 mmHg, respectively. While accurate on average, the standard deviation indicates a degree of variability in its predictions.

Other studies have explored traditional machine learning, compared Linear Regression (LR), Support Vector Regression (SVR), and Artificial Neural Networks (ANNs) for cuffless BP estimation from PPG signals [2]. Among the models tested, they found that SVR with a linear kernel yielded the best performance, achieving an MAE of 13.57 mmHg for SBP and 8.30 mmHg for DBP. Our work builds on the hybrid deep learning approach by providing a direct comparison between a feature-based method and an end-to-end raw signal method.

III. METHODOLOGY

A. Dataset and Preprocessing

This study utilizes data from the MIMIC-III database [1]. We extracted 7-second PPG segments sampled at 125 Hz, paired with reference SBP and DBP values. A 4th-order Butterworth bandpass filter (0.5-4 Hz) was applied to remove baseline wander and high-frequency noise, a standard technique in PPG processing [6]. Signals were then normalized to zero mean and unit variance. For

the raw signal model, these preprocessed segments were used directly as input to the model. For the feature-based model, the following feature extraction step was performed.

B. Feature Extraction for the Hybrid Model

For the feature-based model, a set of 29 features was extracted from each 7-second PPG segment. The process began by identifying key fiducial points within each cardiac cycle: Systolic Peaks (SP), Diastolic Troughs (DT), and Dicrotic Notches (DN). From these points, the following features were engineered:

- **Temporal and Rate Features:** The Mean Pulse Period Interval (PPI), defined as the average time between consecutive systolic peaks, was calculated, showing a moderate negative correlation with both SBP (-0.27) and DBP (-0.35). Its inverse, the Heart Rate, showed a corresponding positive correlation (0.27 SBP, 0.35 DBP).
- **Slope and Duration Features:** The Blood Flow Acceleration was calculated as the mean rising slope to each SP and exhibited a moderate positive correlation with both SBP (0.31) and DBP (0.33). Conversely, the Deceleration (mean falling slope), Downstroke Time (DN to DT duration), and Upstroke Time (SP to DN duration) showed only weak correlations.
- **Amplitude and Shape Features:** The Maximum and Minimum Amplitudes were recorded. Furthermore, Percentile-based Areas were calculated at multiple levels for the systolic-to-diastolic (S-D) and diastolic-to-systolic (D-S) phases to provide detailed shape information.
- **Frequency-Domain Features:** A Fast Fourier Transform (FFT) was applied to each segment to extract the Dominant Frequency and spectral Bandwidth. These features displayed a weak to moderate positive correlation with SBP (0.13 and 0.22, respectively) and DBP (0.17 and 0.19, respectively).

- **Morphological Complexity Features:** Several features were computed to quantify the waveform's shape. Hysteresis, the mean Waveform Derivative (a measure of smoothness), and Curvature (rate of slope change) all exhibited strong positive correlations with both SBP and DBP, with values ranging from 0.29 to 0.36.

C. Model Architectures

The model for the raw signal approach consists of a CNN- LSTM architecture that takes the 7-second preprocessed PPG segment as direct input. The model for the feature-based approach is a hybrid architecture that synergizes deep learning with gradient boosting. The architecture for the feature-based model is broken down as follows:

- **CNN-LSTM Backbone:** The extracted 29-feature vector is fed into 1D convolutional layers, followed by LSTM layers. The final LSTM output is passed through dense layers to produce an initial SBP and DBP prediction.
- **XGBoost Augmentation:** High-level features from the LSTM's final state are extracted and used to train two separate XGBoost regressors (one for SBP, one for DBP).

Final Prediction via Ensemble Averaging: The final predicted SBP and DBP are calculated by averaging the predictions from the CNN-LSTM's output layer and the respective XGBoost regressors.

IV. RESULTS AND DISCUSSION

This section presents the evaluation of two distinct methods: one which uses the raw PPG signal directly, and another which relies on 29 engineered physiological features. The models were trained on a dataset comprising 5794 training, 1449 validation, and 1811 test samples.

A core finding of this study is the significant trade-off between

the two approaches in terms of performance and computational efficiency. The feature-based strategy demonstrated better predictive accuracy and was more efficient. There was a stark contrast in training time: the feature-based model completed 10 epochs in under 3 minutes, whereas the raw signal model required over 2 hours for the same number of epochs, representing a computational cost increase of over 40-fold. The detailed results for each approach are presented below.

A. Feature-Based Model Performance

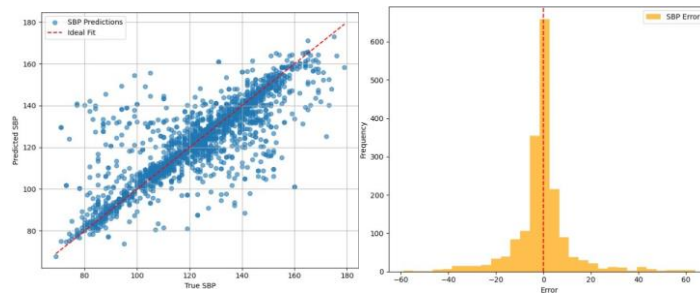
The hybrid model, utilizing the 29 extracted features, demonstrated strong predictive performance and high computational efficiency. The evaluation on the held-out test set is summarized in Table I.

TABLE I: Feature-Based Model Performance

Metric	SBP	DBP
RMSE (mmHg)	11.819	5.3552
	2	
MAE (mmHg)	6.8204	2.9849
R ²	0.6916	0.6759
ME (mmHg)	-0.3141	-0.3273
SD of Error (mmHg)	12.371	5.5937
	5	

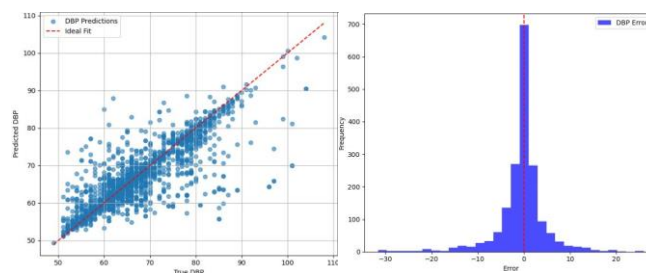
This model showed exceptional performance for Diastolic Blood Pressure (DBP) estimation. With a Mean Error (ME) of -0.3273 mmHg and a Standard Deviation (SD) of error of 5.5937 mmHg, the results fall well within the criteria set by the AAMI standard [3]. For Systolic Blood Pressure (SBP), the MAE of 6.8204 mmHg corresponds to Grade C performance under the IEEE 1708a-2019

standard [4]. The graphical analysis in Fig. 1 and Fig. 2 supports these findings.



(a) True vs. Predicted SBP. (b) SBP Error Distribution.

Fig. 1: Graphical analysis of SBP prediction for the feature- based model.

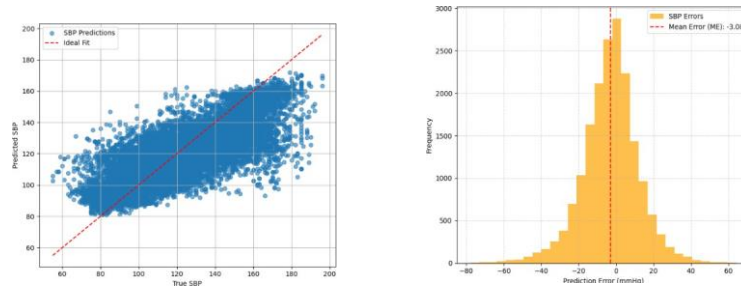


(a) True vs. Predicted DBP. (b) DBP Error Distribution.

Fig. 2: Graphical analysis of DBP prediction for the feature- based model.

fit to the data, meaning the model explains less of the variance in blood pressure. This decrease in performance is illustrated in the accompanying graphical analysis in Fig. 3 and Fig. 4. In the scatter plots, the data points for the raw signal model are more widely dispersed and show a weaker alignment with the ideal fit line compared to the feature-based model. The error distribution histograms further highlight this discrepancy; the chart for the raw

signal model shows a wider spread of errors and fewer predictions concentrated near zero. This visual evidence aligns with the higher ME and SD values in Table II, confirming that the raw signal model is less accurate and more biased.



(a)RAW signal True vs. Predicted (b) RAW signal True vs. Predicted

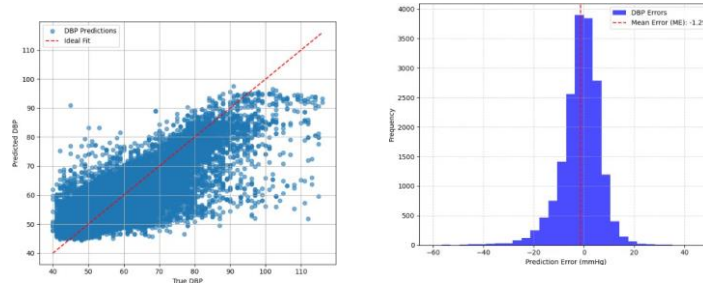
Fig. 3: Graphical analysis of SBP prediction for the RAW signal model

B. Raw Signal Model Performance

The CNN-LSTM model trained directly on raw PPG signals yielded the results presented in Table II.

TABLE II: Raw Signal Model Performance

Metric	SBP	DBP
RMSE (mmHg)	15.368	7.8533
	5	
MAE (mmHg)	11.399	5.6831
	3	
R ²	0.6050	0.6450
ME (mmHg)	-3.0796	-1.2926
SD of Error (mmHg)	15.056	7.7462
	8	



(a) RAW signal True vs. Predicted (b) RAW signal SBP Error Distribution.

Fig. 4: Graphical analysis of SBP prediction for the RAW signal model.

To validate the effectiveness of the feature-based model, a second deep learning model was trained for comparison. This model utilized the same CNN-LSTM architecture and was trained on the same dataset split (5794 training, 1449 validation, and 1811 test samples). However, instead of using engineered features, it was trained directly on the raw, pre-processed PPG signal segments. The performance of this raw signal model, evaluated on the held-out test set, is presented in Table II.

The performance of the raw signal model was notably lower than that of the feature-based model across all metrics. For SBP, the Mean Absolute Error (MAE) increased by approximately 67% (from 6.82 to 11.40 mmHg), and for DBP, it increased by 90% (from 2.98 to 5.68 mmHg). The lower R^2 values (0.6050 for SBP and 0.6450 for DBP) indicate a poorer

V. CHALLENGES AND FUTURE WORK

Challenges

- Current models were trained and validated on the MIMIC-III dataset, which contains high-quality clinical signals. Performance on data from different consumer-grade devices or diverse populations is not guaranteed.
- PPG signals are highly susceptible to motion artifacts, which can significantly degrade signal quality and affect prediction accuracy.
- Implementing deep learning models on resource-constrained wearable devices presents a significant challenge. Our current approach offloads computation to a server via an API, but a fully on-device solution requires model optimization and specialized hardware.

Future Work

- Collection of a new, dedicated dataset using high-quality PPG sensors to retrain and validate the feature-based model in a real-world context.
- Implementation of advanced signal processing techniques to detect and mitigate motion artifacts in real-time.
- Further model refinement and hyperparameter tuning to reduce prediction variance and minimize outliers, particularly for SBP estimations.

VI. CONCLUSION

This study provides three key insights. First, it reaffirms the promising future of using PPG signals for cuffless blood pressure estimation. Second, it demonstrates the excellent results achievable by combining CNN and LSTM architectures for this task. Third, and most critically, it highlights a significant enhancement in both

prediction accuracy and computational efficiency when using a set of engineered physiological features compared to using the raw signal as input.

Employing this feature-based hybrid model, we developed a system with clinically relevant accuracy. The performance for DBP estimation successfully meets the AAMI standard. This underscores that for real-world applications, a feature-based strategy is a more practical and effective choice.

REFERENCES

- [1]T. Mahardika, N. Q. Fuadah, et al., "PPG signals-based blood-pressure estimation using grid search in hyperparameter optimization of CNN-LSTM," *Diagnostics*, vol. 13, no. 15, p. 2566, 2023.
- [2]S. Kılıçkaya, A. Güner, and B. Dal, "Comparison of different machine learning techniques for the cuffless estimation of blood pressure using PPG signals," in *Proc. 2020 Int. Congr. Human-Comput. Interact., Optim. Robot. Appl. (HORA)*, 2020, pp. 1-6.
- [3]Association for the Advancement of Medical Instrumentation, ANSI/AAMI/IEC 80601-2-30:2019: Medical electrical equipment — Part 2-30: Particular requirements for the basic safety and essential performance of automated non-invasive sphygmomanometers, Arlington, VA: AAMI, 2019.
- [4]IEEE Std 1708a-2019 for Wearable Cuffless Blood Pressure Measuring Devices, IEEE SA, 2019.
- [5]A. Cuff, "The evolution of digital health and its continuing challenges," *BMC Digit. Health*, vol. 1, no. 1, p. 3, 2023.
- [6]Alkhoury, Ludvik, et al. "Heart-rate tuned comb filters for processing photoplethysmogram (PPG) signals in pulse oximetry." *Journal of clinical monitoring and computing* 35.4 (2021): 797-813.
- [7]D. Dai, Z. Ji, and H. Wang, "Non-Invasive Continuous Blood Pressure Estimation from Single-Channel PPG Based on a Temporal Convolutional Network Integrated with an Attention Mechanism," *Appl. Sci.*, vol. 14, no. 14, p. 6061, 2024. [Online]. Available: <https://doi.org/10.3390/app14146061>
- S. Kumar, S. Yadav, and A. Kumar, "Blood pressure measurement techniques, standards, technologies, and the latest futuristic wearable cuffless know-how," *Sensors & Diagnostics*, vol. 3, no. 2, pp. 181-202, 2024.

**Design of a Miniaturized Meander-Line
Microstrip Antenna at 915 MHz for Wireless
Capsule Endoscopy in the Colonic Region**

19

Design of a Miniaturized Meander-Line Microstrip Antenna at 915 MHz for Wireless Capsule Endoscopy in the Colonic Region

Jabur Ahmed
Electrical Engineering Department
Omar Al-Mukhtar University
Al Bayda, Libya
Jabur.Ahmed@omu.edu.ly

Ahmed Elbarsha
Electrical and Electronics Department
University of Benghazi
Benghazi, Libya
Ahmed.Elbarsha@uob.edu.ly

Abstract

Wireless capsule endoscopy (WCE) is a non-invasive diagnostic technique for colorectal diseases. This paper presents a fully internal, miniaturized meander-line microstrip antenna operating at 915 MHz for colon-targeted WCE applications. The antenna is designed using a high-permittivity RO3010 substrate with a dielectric constant of 10.2, a shorting pin, and a superstrate layer to achieve compact dimensions compatible with a standard ceramic alumina capsule. The design is optimized in two stages using CST Microwave Studio, first in a simplified homogeneous colon model and subsequently in an anatomically accurate colon Gustav model. Simulation results demonstrate good impedance matching with a minimum reflection coefficient of -22.13 dB at 915 MHz and a -10 dB bandwidth of 8.1 MHz. The antenna exhibits a realized gain of -23.6 dBi and a

radiation efficiency of -28.45 dB. The maximum specific absorption rate (SAR) averaged over 10 g of tissue is 0.47 W/kg, which is well below the IEEE C95.1 safety limit. These results confirm that the proposed antenna is suitable for safe and reliable wireless communication in colon-targeted WCE systems.

Keywords: Colorectal Diseases, Colon Model, Gustav Model, ISM Band, Meander-line, Miniaturized Antennas, Wireless Capsule Endoscopy.

I. INTRODUCTION

Among the most common diseases affecting the small intestine and colon are inflammatory bowel diseases (IBDs) and colorectal cancer (CRC), which is the second most common cancer worldwide [1]. Diagnosing these diseases using traditional methods requires inserting a long, flexible tube into the patient's body. This procedure can cause pain and, in some cases, injury, and it typically requires sedation and anesthesia. All of these factors make this method undesirable for patients, limiting routine screening and real-time monitoring. Given Imaging Ltd., now known as "Medtronic," introduced wireless capsule endoscopy in 2000 with the M2A capsule, a non-invasive method for imaging the gastrointestinal tract. The capsule can be easily swallowed and enables imaging of regions inaccessible to conventional endoscopy, such as the small intestine and colon. These images are then wirelessly transmitted to an external receiver without the need for anesthesia or discomfort. [2]. In general, the effectiveness of a wireless capsule endoscopy system largely depends on the antenna's efficiency and physical size. An efficient and compact antenna helps conserve battery life while freeing internal space for other system components.

Efficient antenna operation within the complex and lossy human-body environment is essential for WCE systems. Choosing the right

operating frequency can directly affect both antenna performance and tissue penetration. The most common frequency bands used in WCE systems are MedRadio, which includes the 403-406 MHz band; Wireless Medical Telemetry (WMTS), which includes the 1.427-1.432 GHz band; and Industrial Scientific and Medical (ISM) bands (433 MHz, 915 MHz, 1.4 GHz, and 2.45 GHz) [3]. Each of these frequency bands is suitable for wireless medical telemetry applications, as the antenna size and tissue penetration depth are inversely proportional to the operating frequency. Frequency selection is critical in antenna design for WCE.

Several approaches have been explored for miniaturizing WCE antennas to fit within the limited capsule volume while maintaining efficient in-body performance. One common method involves a conformal antenna designed to be wrapped around the capsule shell, freeing internal space for other electronic components. For example, Feng et al. [4] proposed an ultra-wideband conformal antenna covering 950 MHz–3.62 GHz for a skin-tissue model, achieving broad coverage but with size and manufacturing complexity trade-offs. Similarly, Mimouni et al. [5] implemented a meander-line conformal antenna operating in the ISM 5.8 GHz band inside the Hugo anatomical model, demonstrating acceptable SAR values. Other works have investigated spiral and meander-line geometries to extend the electrical length without enlarging the antenna footprint. For instance, Yamak & Özdemir [6] developed a square spiral meander-line antenna at 433 MHz and tested it in various anatomical regions, reporting stable S21 link measurements in certain positions. However, such conformal designs are highly susceptible to detuning effects as the capsule moves between organs with differing dielectric properties, resulting in less reliable performance across the gastrointestinal tract [7].

Encapsulated miniaturized antennas have been proposed to mitigate these variations by shielding the radiator from direct contact with surrounding tissues. Iqbal et al. [8] designed an arc-slot patch antenna

with a shorting pin for 915 MHz operation, achieving low SAR levels. Alshammari et al. [9] designed a miniaturized out-of-band self-duplex antenna in a homogeneous muscle tissue model, working at 1.8 and 2.4 GHz. By using shorting pins and semicircle slots, this design achieves very good isolation between the transmit port and the receive port. Also, in another study, Alshammari et al. [10] designed an ultraminiaturized dual-band antenna by making open-ended slots and arc-shaped slots and shorting pins. These studies demonstrate significant progress in WCE antenna miniaturization. However, a recurring limitation evident in all studies is that in most cases, simplified or generalized tissue models, often homogeneous blocks or a stack of simplified layers, are used to mimic the properties of human tissue. These models provide a suitable environment for antenna optimization at early stages, but they do not consider the complex curvatures, heterogeneous tissue composition, and irregular boundary conditions present in the human colon. Therefore, a detailed anatomical model of the colon was not used in any of these studies [11-14]. This gap draws attention to the need to rely on realistic anatomical models to improve prediction of actual in-body performance as well as improve communication reliability under realistic conditions. Table I presents a comparison between literature and this work.

This study fills this gap by specifically designing a miniaturized antenna operating at 915 MHz for wireless capsule endoscopy targeting the colon. This antenna was first designed and developed in a simplified model of the colon to reduce computational resources. Subsequently, an anatomically detailed colon Gustav model provided by CST Microwave Studio was used to represent a more complex and detailed environment. The 915 MHz frequency was chosen because it provides an optimal balance between tissue penetration depth and antenna size [15]. Although the 433 MHz ISM band is commonly used in implantable medical devices due to its superior tissue penetration, operation at this frequency requires a significantly larger electrical

length, which leads to increased antenna dimensions that are difficult to accommodate within the limited volume of a wireless capsule endoscope. In contrast, operation at 915 MHz enables substantial antenna miniaturization while still providing sufficient penetration depth for the short communication distances encountered in colon-targeted WCE systems. Moreover, in in-body telemetry, radiation efficiency is primarily dominated by dielectric losses in surrounding tissues rather than operating frequency alone. Therefore, the marginal penetration advantage of 433 MHz does not translate into a proportional improvement in link reliability for colonic WCE. The 915 MHz ISM band also offers global availability and compatibility with low-power commercial transceivers, making it a practical and well-balanced choice for the proposed design [15]. The antenna was miniaturized using a meander-line structure with a high-permittivity substrate and shorting pin. The meander-line provides an extended electrical path for the surface current without increasing the physical dimensions of the antenna, enabling a miniaturized antenna design that maintains stable performance in the complex, high dielectric environment of body tissue. A superstrate layer was also used to isolate the antenna from the surrounding environment and maintain performance within complex body tissue. All design, modeling, and performance evaluations in this study were conducted using the CST Microwave Studio simulation environment.

II. DESIGN METHODOLOGY

A. Miniaturization

To create a miniaturized antenna, it must provide a long electrical path without increasing its physical dimensions. A meander-line antenna can increase the effective electrical length of the patch without enlarging its physical dimensions.

Therefore, the resonant frequency is reduced. Chattha et al. [16] have developed an empirical equation to calculate the resonant frequency f_0 of the meander line antenna.

By studying various parameters of the meander-line structure, they derived the following empirical equation:

$$f_0 = \frac{c}{3W + 5.6L + 3.7h - 3W_f - 3.7W_s - 4.3L_b - 2.5L_s} \quad \square \square \square$$

Where c is the speed of light, W is the meander line width, L is the length, h is the height of the substrate, W_f is the diameter of the inner conductor of the coaxial feed line, W_s is the diameter of the shorting pin used, L_s is the distance between the shorting pin and the edge of the meander line, and L_b is the distance between the shorting pin and coaxial feed line. This equation guides the design of the meander-line antenna by enabling careful optimization of its physical parameters.

B. Proposed Antenna Geometry

The proposed meander-line antenna was designed using (1), and its overall length was 57.3 mm. To further miniaturize the antenna, a high-permittivity RO3010 substrate ($\epsilon_r = 10.2$) with a thickness of 0.635 mm is used. The antenna is fed using a coaxial feed line (commercial UT-020), the overall physical parameters of the antenna are illustrated in Fig. 1(a), and their values are listed in Table II.

C. Capsule Integration

Most studies employ the standard capsule size with a diameter of 11 mm and a length of 26 mm, divided into three sections: the middle section with a length of 15 mm and two domes at the top and bottom, each with a radius of 5.5 mm, and the shell thickness of this capsule is 0.2 mm, as illustrated in Fig. 1(b). This capsule is made of ceramic alumina (Al_2O_3), which is non-toxic and safe for used inside the human body.

D. Simplified Colon Model

In this research, we chose to work on a cubic simplified colon model with a side length of 100 mm. The colon model has a dielectric constant of 57.9 and a conductivity of 1.09 S/m at 915 MHz, and this data was taken from the reliable Swiss database Information Technologies in Society (IT'IS) [17]. Fig. 2 illustrates the return loss of this antenna within the simplified colon model.

TABLE I. PERFORMANCE COMPARISON WITH LITERATURE

Ref	Frequency	Volume (mm ³)	Gain (dBi)	Model	SAR (W/kg) /10g
[4]	Ultra-wideband	18**	-25.4	Homogeneous Skin	254.8/1g
[5]	5.8 GHz	**	3.6	Hugo model	4
[6]	433 MHz	537.2	-37.1	Three-layer model	-
[7]	2.45 GHz	**	-23.98	Phantom liquid	0.308
[8]	915 MHz	9.01*	-24.6	Homogeneous muscle model	0.014
[9]	1.85, 2.4 GHz	3.67*	-45.3, -45.8	Homogeneous muscle model	-
[10]	1.4, 2.45 GHz	42.7*	-29.4, -30.4	Minced pork	18.86, 12.93
[11]	2.4 GHz	172.2**	-10	-	-
[12]	MedRadio	177.8*	-31.5	Gustav model	913/1g
[13]	2.4 GHz	**	-31.6	chopped pork	368.7/1g
[14]	433 MHz	**	-31.5	Liquid human phantom	-
This Work	915 MHz	71.3*	-23.6	Colon Gustav model	0.47

** means the antenna is conformal, and * means the antenna is miniaturized.

E. Isolation of the Antenna

Inserting the proposed antenna inside the colon model, which has a very high dielectric constant, will cause a major impedance mismatch, and to isolate the antenna from the external environment, we need an additional layer over the patch, which is called a superstrate. The superstrate used in the proposed antenna has the same material and thickness as the substrate, which is RO3010 with a 10.2 dielectric constant. Using this extra layer can ensure that the antenna is isolated from the external environment and can be matched properly. Fig. 3(a) illustrates the superstrate on the proposed antenna. Fig. 3(b) illustrates S11 after using the superstrate layer.

F. Antenna Optimization

Using the superstrate layer can ensure a very good impedance matching by isolating the antenna from the external environment. However, adding this layer can cause a shift in the operating frequency, which needs to be optimized properly to operate at 915 MHz. The operating frequency was tuned by adjusting Px and Fx. The optimized values are Px = 3.72 mm and Fx = 1.63 mm. Fig. 4 illustrates the S11 for the optimized antenna inside the simplified colon model.

TABLE II. ANTENNA PHYSICAL PARAMETERS

Parameter	Description	Value (mm)
Ws	Width of the substrate	7.495
Ls	Length of the substrate	7.495
W	Width of the patch	6.9
L	Length of the patch	6.9
C	Width of the meander-line	0.6

G	The gap between the meander lines	0.3
h	Height of the substrate	0.635
t	Thickness of the Copper	0.035
d	Coaxial inner conductor diameter	0.127
Pd	Shorting pin diameter	0.2
Px	Horizontal distance from the shorting pin to the patch edge.	3.765
Fx	Horizontal distance between the feed probe and the shorting pin	1.575

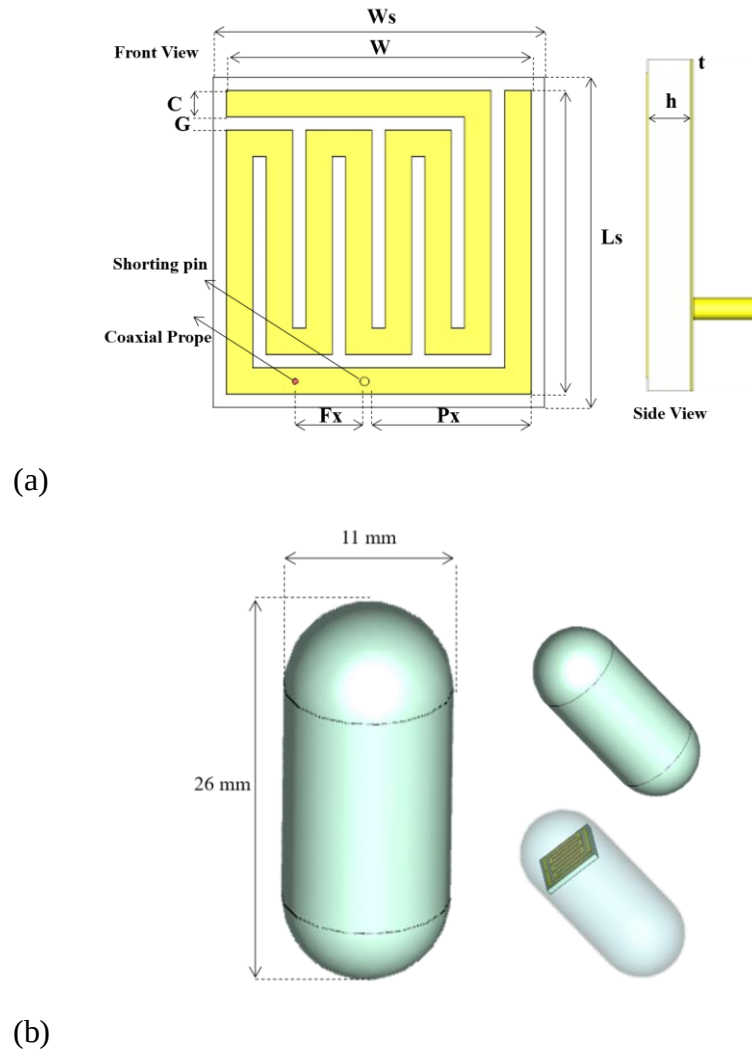


Fig. 1. (a) Proposed Antenna Physical Parameters, (b) Capsule Integration.

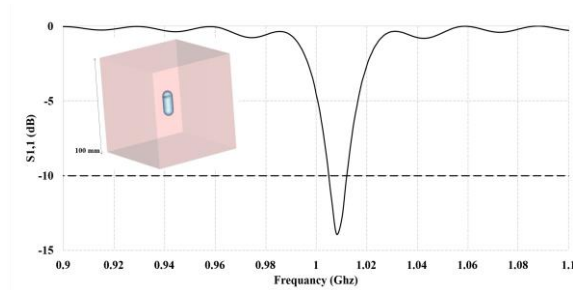


Fig. 2. S11 of the proposed antenna inside the colon model.

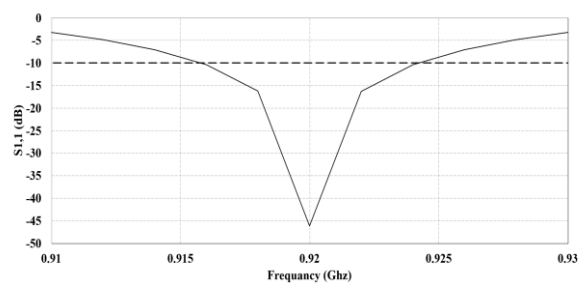
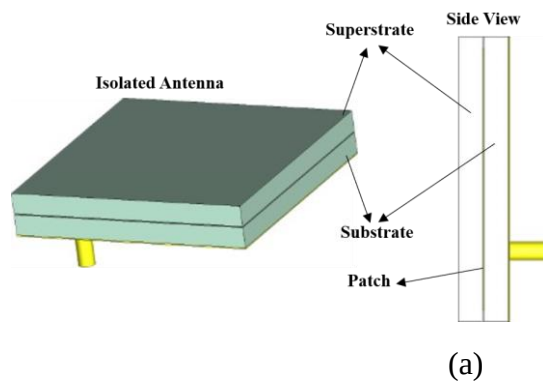
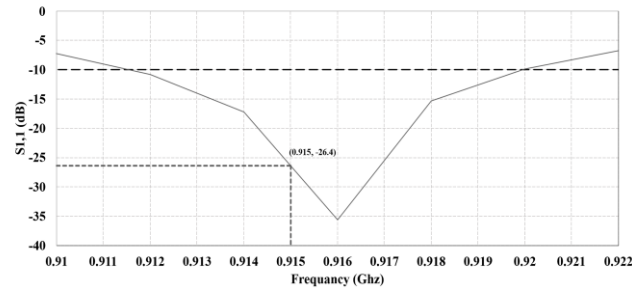


Fig. 3. (a) Superstrate Layer, (b) S11 for the Isolated Antenna.



S11 of the Optimized Antenna inside the Simplified Colon

I. HUMAN GUSTAV MODEL

The proposed antenna design is evaluated within the anatomically detailed colon section of the Gustav human body model provided by CST Microwave Studio. This model provides anatomically accurate geometry and dielectric properties taken from real human tissue data, allowing a more realistic in-body electromagnetic analysis than homogeneous or simplified phantoms. The antenna was positioned inside the colonic region of the Gustav model and integrated into the capsule geometry. Due to the high dielectric and conductive losses of the surrounding tissues, an additional impedance mismatch was observed compared to the simplified colon model. This mismatch necessitated further optimization of the feed and shorting-pin positions, denoted P_x and F_x .

A parametric sweep in CST Microwave Studio determined that optimal resonance at 915 MHz was achieved with $P_x = 2.95$ mm and $F_x = 1.76$ mm. These adjustments preserved the compactness of the antenna while restoring proper impedance matching. Fig. 5(a) illustrates the Gustav colon model, and Fig. 5(b) shows the S11 of the antenna inside the model. Fig. 6 shows the S11 of the optimized antenna inside the model.

II. RESULTS AND DISCUSSION

The performance of the optimized antenna was evaluated in the Gustav colon model at 915 MHz. Key parameters include impedance matching, bandwidth, gain, directivity, efficiency, and specific absorption rate (SAR).

A. Impedance Matching and Bandwidth

The simulated reflection coefficient S_{11} shows excellent impedance matching, with a minimum of -22.13 dB at 915 MHz, as shown in Fig. 6. The -10 dB bandwidth is 8.1 MHz, corresponding to a fractional bandwidth of approximately 0.88%. Although the bandwidth is narrow, it remains sufficient for fixed-frequency ISM applications if fabrication tolerances and dielectric constant variations are carefully controlled. The use of a high-permittivity superstrate played a key role in maintaining stable matching in the high-loss colonic environment.

B. Gain, Directivity, and Radiation Pattern

The realized gain is -23.6 dBi, consistent with expectations for electrically small in-body antennas where tissue absorption is significant. The directivity is 4.79 dBi, with the main lobe in the $\phi = 90^\circ$ plane having a magnitude of 4.12 dBi at 54° and a 3 dB beamwidth of 130.1° . In the $\theta = 90^\circ$ plane, the main lobe magnitude is 3.41 dBi at 83° , with a 3 dB beamwidth of 112.3° . These wide beamwidths are advantageous for capsule endoscopy, as they ensure consistent communication despite continuous capsule rotation and orientation changes during operation. Fig. 7(a) illustrates the simulated 3D far-field pattern, and Fig. 7(b) illustrates the radiation patterns for the proposed antenna inside the colon Gustav model.

C. Efficiency

The simulated radiation efficiency is -28.45 dB, which is within the expected range for compact in-body antennas at sub-GHz frequencies. This reduced efficiency results mainly from strong dielectric losses in

surrounding biological tissues, which represent an inherent limitation of in-body antennas rather than a design deficiency.

D. Specific Absorption Rate (SAR)

The maximum SAR averaged over 10 g of tissue is 0.47 W/kg, which is below the IEEE C95.1-2019 limit of 2 W/kg for localized exposure [18]. This confirms that the proposed design meets safety requirements for in-body wireless telemetry, even under worst-case placement conditions in the colon. Fig. 8 shows the max SAR value of the model.

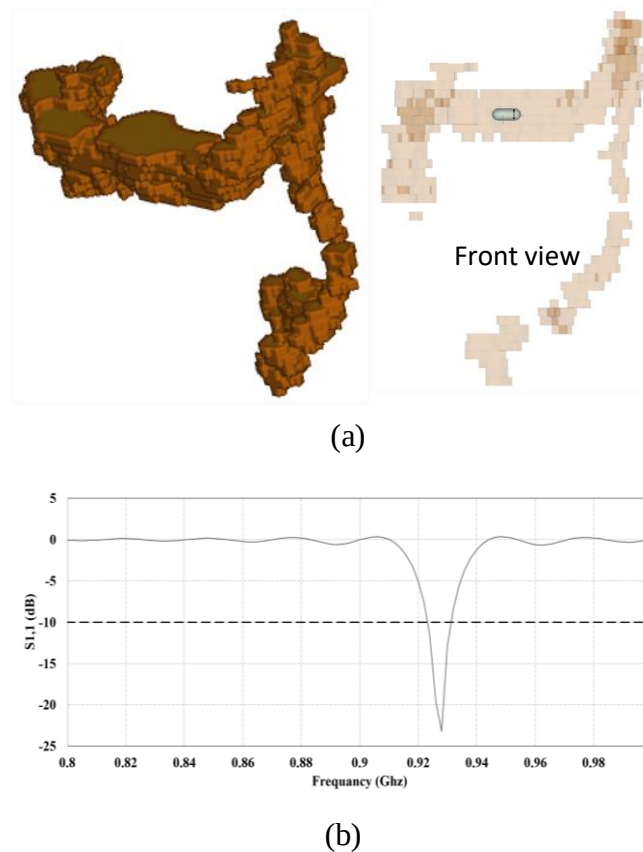


Fig. 4. (a) Colon Gustav Model, (b) S11 for the Antenna Inside the Gustav Colon Model.

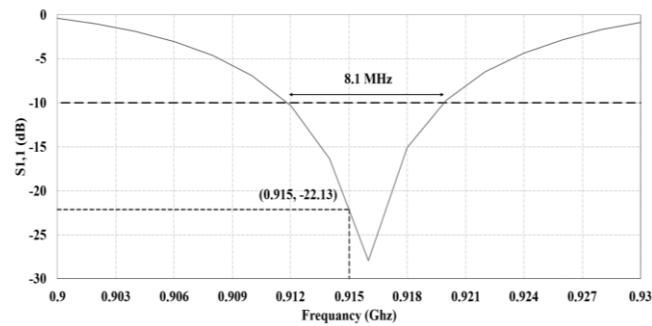
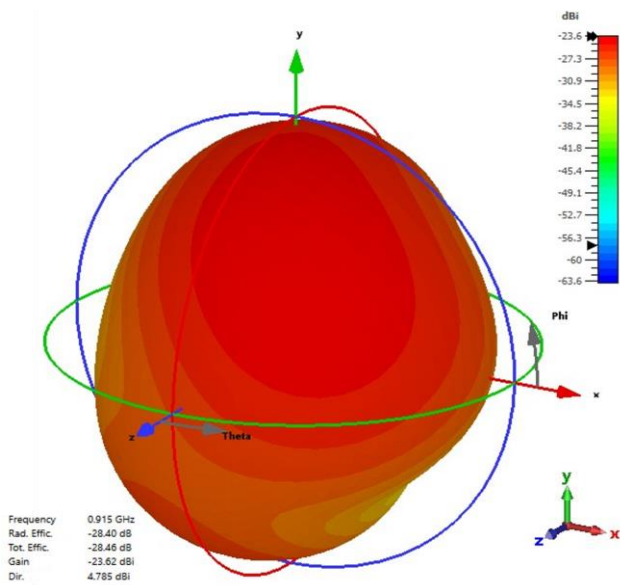
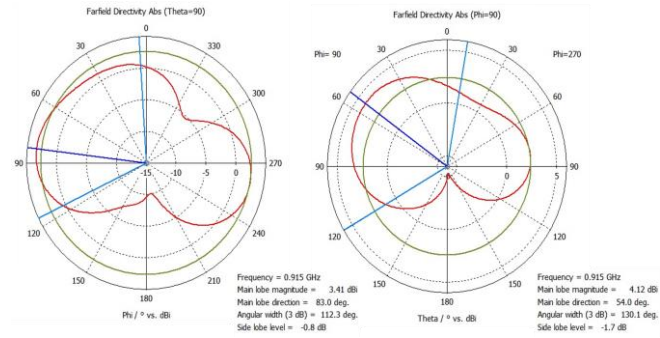


Fig. 5. S11 of the optimized antenna inside the Gustav Colon model.



(a)



(b)

Fig. 6. (a) Simulated 3D far-field pattern, (b) Radiation patterns.

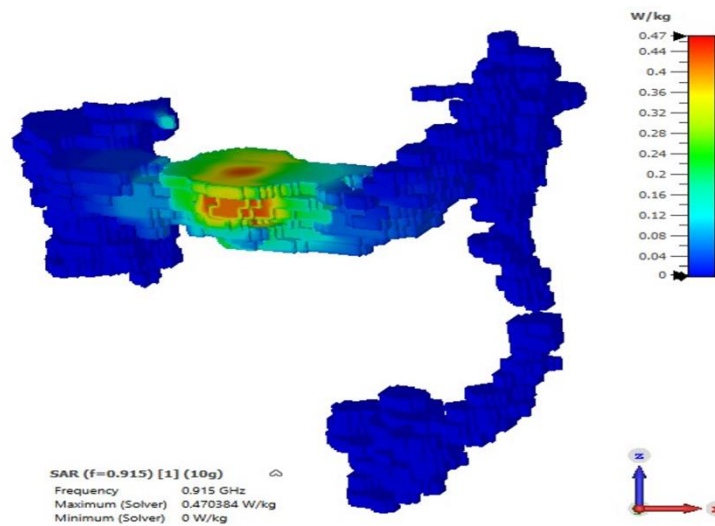


Fig. 7. Maximum SAR value of the model.

E. Discussion

Using a highly lossy and anatomically accurate colon Gustav model, the results demonstrate that the proposed miniaturized meander-line antenna can provide consistent impedance matching and SAR compliance. While fabrication and experimental validation are beyond the scope of this study, the antenna performance was evaluated while embedded in the capsule geometry and simulated inside an anatomically accurate colon Gustav model, providing a realistic assessment of practical in-body operation.

While the fractional bandwidth is relatively narrow, it remains sufficient for the target ISM application at 915 MHz. Given the inherent low efficiency of electrically small in-body antennas, for WCE at 915 MHz, the achieved -22.13 dB S11, wide beamwidth patterns, and SAR compliance are more critical to reliable telemetry than maximizing gain, especially given the -28.45 dB efficiency typical of in-body antennas.

The maximum simulated SAR (averaged over 10 g of tissue) is 0.47 W/kg, well below the IEEE C95.1 safety limit of 2 W/kg, highlighting the design's safety margin and suitability for in-body applications. Future enhancements may include bandwidth broadening techniques, such as slot tuning or the inclusion of parasitic elements, while preserving compact geometry.

III. CONCLUSION

A fully internal miniaturized meander-line antenna designed for wireless capsule endoscopy in the colonic region at 915 MHz has been presented. The design utilizes a high-permittivity RO3010 substrate and superstrate, along with a shorting pin, to achieve a compact size while maintaining stable impedance matching in a high-dielectric, lossy environment. The antenna was initially optimized in a simplified colon model and subsequently fine-tuned within CST Microwave

Studio's anatomically detailed Gustav colon model to account for realistic tissue geometry and dielectric properties.

Simulation results indicate excellent impedance matching ($S_{11} = -22.13$ dB) and a -10 dB bandwidth of 8.1 MHz, wide beamwidth radiation patterns, and SAR levels well within IEEE safety limits (0.47 W/kg for 10 g average). Although radiation efficiency is low due to inherent in-body losses, the achieved stability, compactness, and compliance with safety standards make the design a strong candidate for reliable WCE communication, particularly in scenarios where stable performance in anatomically detailed models, such as the Gustav colon, is required.

Future work will focus on antenna fabrication and experimental validation using tissue-equivalent phantoms, as well as exploring design modifications to broaden the operational bandwidth without significantly increasing antenna size.

REFERENCES

- [1] I. V. Mitselos, D. K. Christodoulou, K. H. Katsanos, and E. V. Tsianos, "Role of wireless capsule endoscopy in the follow-up of inflammatory bowel disease," *World J Gastrointest Endosc.*, vol. 7, no. 6, pp. 643–651, Jun. 2015, doi: 10.4253/wjge.v7.i6.643.
- [2] G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, "Wireless capsule endoscopy," *Nature*, vol. 405, no. 6785, p. 417, Jan. 2000, doi: 10.1038/35013140.
- [3] S. Fontana et al., "State of the art on advancements in wireless capsule endoscopy telemetry: A systematic approach," *IEEE Open J. Antennas Propag.*, vol. 5, pp. 1282–1294, Oct. 2024, doi: 10.1109/OJAP.2024.3409827.
- [4] Y. Feng, S. P. Pan, J. W. Li, P. Chen, L. Qi, and G. S. Li, "Design of ultra-wideband conformal capsule antenna for wireless capsule endoscopy system," in *Proc. Cross Strait Radio Sci. Wireless Technol. Conf. (CSRSWTC)*, Dec. 2020, pp. 1–3, doi: 10.1109/CSRSWTC50769.2020.9372442.
- [5] A. Mimouni, B. Fady, J. Terhzaz, A. Tribak, and H. Akhdar, "Novel folded antenna design and SAR analysis for WCE and biomedical applications," *Prog. Electromagn. Res. C*, vol. 135, pp. 107–119, Jan. 2023, doi: 10.2528/PIERC23033101.
- [6] F. S. Yamak and Ö. Özdemir, "Square spiral conformal antenna design for capsule endoscopy," in *Proc. 29th Signal Process. Commun. Appl. Conf. (SIU)*, Jun. 2021, pp. 1–4, doi: 10.1109/SIU53274.2021.9477797.
- [7] M. L. Scarpello et al., "Design of an implantable slot dipole conformal flexible antenna for biomedical applications," *IEEE*

- Trans. Antennas Propag., vol. 59, no. 10, pp. 3556–3564, Oct. 2011, doi: 10.1109/TAP.2011.2163761.
- [8] A. Iqbal, M. Al-Hasan, I. B. Mabrouk, and T. A. Denidni, "Capsule endoscopic MIMO antenna with radiation pattern and polarization diversity," IEEE Trans. Antennas Propag., vol. 71, no. 4, pp. 3146–3154, Apr. 2023, doi: 10.1109/TAP.2023.3244035.
- [9] A. Alshammari, A. Iqbal, C. Balocco, and I. Mabrouk, "A compact self-duplex antenna for wireless capsule endoscopic devices," in Proc. IEEE USNC-URSI Radio Science Meeting, Jul. 2023, pp. 1383–1384, doi: 10.1109/USNC-URSI52151.2023.10237932.
- [10] A. Alshammari, A. Iqbal, A. Basir, R. B. V. B. Simorangkir, and I. B. Mabrouk, "Ultraminiaturized dual-band implantable antenna for wireless capsule endoscopy," IEEE Sensors J., vol. 24, no. 9, pp. 15210–15218, May 2024, doi: 10.1109/JSEN.2024.3374817.
- [11] S. Preethiba and S. Brindha, "Design of conformal antenna for wireless capsule," in Proc. Int. Conf. Commun., Comput. Internet Things (IC3IoT), Jun. 2018, pp. 77–80, doi: 10.1109/IC3IoT.2018.8668181.
- [12] J. Wang, M. Leach, E. G. Lim, Z. Wang, R. Pei, and Y. Huang, "An implantable and conformal antenna for wireless capsule endoscopy," IEEE Antennas Wireless Propag. Lett., vol. 17, no. 7, pp. 1153–1157, Jul. 2018, doi: 10.1109/LAWP.2018.2836392.
- [13] R. Li, Y. X. Guo, and G. Du, "A conformal circularly polarized antenna for wireless capsule endoscope systems," IEEE Trans. Antennas Propag., vol. 66, no. 4, pp. 2119–2124, Apr. 2018, doi: 10.1109/TAP.2018.2804674.

- [14] S. Kim and H. Shin, "An ultra-wideband conformal meandered loop antenna for wireless capsule endoscopy," *J. Electromagn. Eng. Sci.*, vol. 19, no. 2, pp. 101–106, Apr. 2019, doi: 10.26866/jees.2019.19.2.101.
- [15] A. Kiourti and K. S. Nikita, "A review of implantable patch antennas for biomedical telemetry: challenges and solutions," *IEEE Antennas Propag. Mag.*, vol. 54, no. 3, pp. 210–228, Jun. 2012, doi: 10.1109/MAP.2012.6293992.
- [16] H. T. Chattha, H. Yi, Z. Xu, and L. Yang, "An empirical equation for predicting the resonant frequency of planar inverted-F antennas," *IEEE Antennas Wireless Propag. Lett.*, vol. 8, pp. 856–860, 2009, doi: 10.1109/LAWP.2009.2027822.
- [17] C. Baumgartner et al., "IT'IS Database for thermal and electromagnetic parameters of biological tissues," Version 5.0, 2025. [Online]. Available: itis.swiss/database. doi: 10.13099/VIP21000-04-1.
- [18] IEEE Standard for Safety Levels with Respect to Human Exposure to Electric, Magnetic, and Electromagnetic Fields, 0 Hz to 300 GHz, IEEE Std C95.1-2019, 2019, doi: 10.1109/IEEESTD.2019.8859679.

**Deep Learning-Based Detection of
Myocardial Infarction from 12-Lead ECG
Using a Modified RegNet Architecture**

20

Deep Learning-Based Detection of Myocardial Infarction from 12-Lead ECG Using a Modified RegNet Architecture

Taha Mahmoud

Libyan Academy for Postgraduate Studies Department of Medical Engineering Tripoli, Libya

dr.taha.libya@gmail.com

Abstract

Myocardial infarction (MI) remains a leading cause of global mortality, where rapid diagnosis via 12-lead electro-cardiogram (ECG) is critical for timely reperfusion therapy. Manual interpretation suffers from variable sensitivity with physician detection rates as low as 50–60%. We propose a modified RegNetY-800MF architecture optimized for 5-second 12-lead ECG window, designed for real-time clinical triage in emergency departments.

Unlike prior studies that focus on multi-class diagnosis or MI versus normal classification, we adopt a clinically realistic binary formulation: MI versus a composite non-MI class containing all other diagnostic categories in PTB-XL (including normal variants and diverse pathologies). This more challenging formulation better reflects emergency triage priorities. Trained and evaluated on the PTB-XL dataset (21,837 recordings from 18,885 patients) using a patient-wise 80/10/10 split, our model leverages an advanced augmentation pipeline including temporal rolling, amplitude scaling, Gaussian noise, baseline wander, random lead dropout, precordial perturbation, and wavelet-based perturbations to enhance robustness

and address class imbalance. Our model achieves a test AUC of 0.8540, AUPRC of 0.6595, sensitivity of 84.62%, specificity of 69.90%, precision of 46.56%, F1-score of 0.60, and a clinically transformative negative predictive value (NPV) of 93.62%. With only 5.7 million parameters, our architecture achieves high diagnostic sensitivity while remaining substantially smaller than many recently proposed transformer-based ECG models, making it suitable for deployment in resource-constrained clinical settings. This lightweight, high-sensitivity architecture is ideally suited for deployment on portable ECG devices and emergency triage systems, offering a scalable decision-support tool to reduce missed diagnoses while minimizing unnecessary interventions.

Index Terms: Myocardial infarction, ECG, deep learning, RegNet, data augmentation, PTB-XL

1. INTRODUCTION

Myocardial infarction (MI) remains a leading cause of global mortality, with rapid diagnosis via 12-lead electrocardiogram being critical for guiding reperfusion therapy. However, manual interpretation suffers from substantial variability, particularly for ST-elevation MI (STEMI), where sensitivity can be as low as 50-60% [1]. This diagnostic gap leads to approximately 20% of acute MI cases being missed under conventional STEMI criteria [2], highlighting the urgent need for improved interpretation strategies. While deep learning has shown promise in standardizing ECG analysis, current approaches face three critical barriers to clinical implementation. First, models are typically trained on idealized data, creating a domain gap with noisy clinical environments [3]. Second, profound class imbalance (~4% MI-positive in PTB-XL) biases models against true positives [4]. Third, architectural inefficiency, Transformer-based ECG models often involve large parameter counts, raising concerns

for edge deployment [5]. Additionally, most existing studies frame ECG classification as either multi-class diagnosis or binary MI-vs-normal detection, which does not reflect the emergency triage reality where clinicians must distinguish MI from a heterogeneous population including both healthy individuals and patients with other cardiac pathologies.

To address these limitations, we introduce a modified RegNetY-800MF architecture [6] optimized for 12-lead ECG analysis, coupled with a comprehensive augmentation pipeline that simulates clinical artifacts. Crucially, we adopt a clinically realistic binary formulation: MI versus a composite non-MI class containing all other diagnostic categories in PTB-XL. This approach better mirrors the emergency department question, "Is this a life-threatening MI requiring immediate intervention?" Our contributions include:

- First application of RegNetY-800MF to ECG-based MI detection
- Novel binary formulation (MI vs. all other conditions) reflecting emergency triage priorities
- Comprehensive artifact-aware augmentation pipeline addressing class imbalance and domain shift
- High sensitivity (84.62%) and NPV (93.62%) on PTB-XL with clinically acceptable specificity (69.90%)
- Parameter-efficient framework (5.7M parameters) enabling edge device deployment

By combining architectural efficiency with clinically realistic task formulation and training, our work establishes a new paradigm that prioritizes real-world utility over bench-mark performance alone, proving that smaller, clinically-aware models can achieve meaningful diagnostic improvements in emergency settings.

2. RELATED WORK

Automated ECG analysis has evolved from rule-based systems to classical machine learning and modern deep learning. Early rule-based approaches [7], [8] achieved moderate accuracy (70-80%) but poor generalizability, especially for STEMI. Subsequent machine learning methods (SVMs, random forests) improved performance but remained brittle to noise and failed to capture complex spatiotemporal patterns [9].

Deep learning has established end-to-end learning as a powerful paradigm for ECG analysis. Convolutional Neural Networks (CNNs), particularly ResNet- and Inception-based 1D architectures, have set the state-of-the-art, achieving strong performance across diverse tasks. On the large-scale PTB-XL dataset, these models demonstrated macro AUCs of up to 0.93 for diagnostic super-class classification, which includes Myocardial Infarction (MI), significantly outperforming feature-based approaches. While recurrent models like LSTMs were competitive, they were consistently outperformed by their convolutional counterparts without providing clear gains (AUC 0.9 on MIMIC III dataset) [10]. The performance of these top models comes with a range of computational costs, from compact ResNet variants with an estimated 10M parameters to larger architectures like xresnet1d101 which may require 40-50M parameters [11].

This landscape highlights a need for more efficient and principled architectures. Existing efficiency-focused models like EfficientNet with ≈ 20 M parameters [12] often trade accuracy for size without a design philosophy specific to biomedical time-series. This underscores the potential of the RegNet design paradigm [10], which uses parametric rules for width and depth to systematically optimize the speed-accuracy trade-off. Furthermore, RegNetY variants that incor-

Limitations in Current Task Formulation: While deep learning models have advanced significantly, most studies adopt either multi-class classification or binary MI-versus-normal detection. The former aims for comprehensive diagnosis but dilutes resources across multiple conditions, while the latter simplifies the task by excluding challenging non-MI pathologies that mimic MI in clinical practice. To our knowledge, no prior work has systematically evaluated a binary MI-versus-all-other-conditions formulation that includes diverse pathologies in the negative class, despite this better reflecting the triage question faced in emergency departments. Data augmentation is crucial for addressing class imbalance and improving robustness [3], yet most studies use limited augmentation. While some works have explored synthetic generation techniques, these often focus on oversampling rather than enhancing representation of rare but clinically critical MI subtypes such as posterior MI, which are frequently missed in clinical practice.

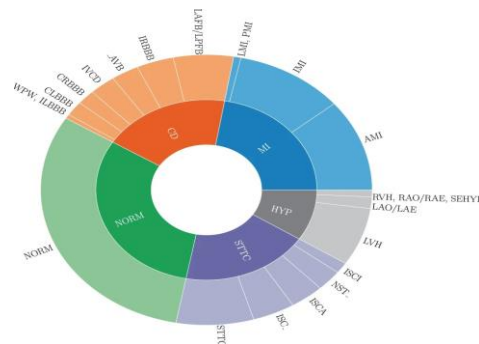


Fig. 1. Class Distribution of the PTB-XL ECG Dataset.

The donut chart illustrates the prevalence of diagnostic categories across 21,837 ECG recordings. Major groups include Normal (NORM), Myocardial Infarction (MI), ST-T Changes (STTC), Hypertrophy (HYP), Conduction Disturbances (CD), and others. Within MI, subtypes such as Anterior MI (AMI), Inferior MI (IMI), and Lateral MI (LMI) are shown, highlighting class imbalance and rare MI subtypes. Figure adapted from Wagner et al. [14].

Our work bridges these gaps by: (1) first applying RegNetY-800MF to ECG MI detection, (2) introducing a clinically realistic binary formulation (MI vs. all other conditions), (3) implementing a comprehensive artifact-aware augmentation strategy, and (4) maintaining parameter efficiency (5.7M) for edge deployment while achieving clinically meaningful sensitivity (84.62%) and negative predictive value (93.62%). We prioritize clinical utility—particularly minimizing missed MI cases—over benchmark performance alone.

3. METHODOLOGY

Dataset: PTB-XL as ECG Deep Learning Benchmark

All experiments used the PTB-XL dataset [14], comprising 21,837 10-second 12-lead ECG recordings from 18,885 patients. Signals were sampled at 500 Hz using Schiller AG devices (1989–1996). Cardiologists annotated recordings using SCP-ECG standard with multi-label diagnostic statements [15]. The PTB-XL dataset exhibits significant class imbalance, as shown in Figure 1, with Normal (NORM) cases dominating and rare MI subtypes requiring targeted augmentation strategies.

Clinical Rationale for Binary Formulation: Unlike prior studies focusing on multi-class diagnosis or MI versus normal classification, we adopt a clinically realistic binary formulation: MI versus all other diagnostic categories in PTB-XL. This approach reflects the primary

question in emergency triage:”Is this a life-threatening MI requiring immediate intervention?”

Rather than attempting comprehensive differential diagnosis. By including diverse pathologies (hypertrophy, conduction disturbances, ST-T changes, etc.) in the negative class, our model learns to distinguish MI from its clinical mimics—a more challenging but clinically relevant task.

For binary myocardial infarction (MI) classification, we aggregated “Anterior MI,” “Inferior MI,” “Posterior MI,” and “Lateral MI” into the positive class ($n = 5,189$). Remaining recordings formed the negative class ($n = 16,648$), yielding a 3.2:1 imbalance ratio. This negative class comprises both normal records ($n = 9,524$) and 17 other diagnostic categories ($n = 7,124$), creating a heterogeneous “non-MI” cohort that better simulates real emergency department populations.

Patient-wise stratified splitting ensured no data leakage:

- Training: 17,418 recordings ($\sim 15,108$ patients) $\rightarrow$ 52,254 samples
- Validation: 2,183 recordings ($\sim 1,888$ patients) $\rightarrow$ 6,549 samples
- Test: 2,198 recordings ($\sim 1,889$ patients) $\rightarrow$ 6,594 samples

Final counts increased due to 5-second windowing (Section III-B). Class distribution was preserved via patient-level stratified sampling.

3.1 ECG Preprocessing

To prepare ECG signals from the PTB-XL dataset for myocardial infarction (MI) detection, a robust preprocessing pipeline was implemented to enhance signal quality, remove artifacts, and standardize the data for machine learning analysis. The pipeline

included wavelet-based denoising, band pass filtering, notch filtering, amplitude clipping, normalization, and signal segmentation. These steps were designed to pre-serve diagnostic ECG features critical for MI detection, such as QRS complexes, P waves, and ST-segment changes, while mitigating noise sources, including baseline wander, muscle artifacts, and power line interference.

3.2 Data Loading

ECG records from the PTB-XL dataset, comprising 12-lead signals sampled at 500 Hz (high-resolution recordings), were loaded using the WFDB library [16]. The 500 Hz sampling rate was used directly, as it provided sufficient temporal resolution for capturing diagnostic features and was compatible with the input requirements of the machine learning model, eliminating the need for resampling.

3.3 Signal Filters

Wavelet Denoising: A wavelet filter was applied to remove high-frequency noise and artifacts while preserving the morphological characteristics of the ECG. The Daubechies 4 (db4) wavelet with a decomposition level of 4 was used to decompose each lead into approximation and detail coefficients. All detail coefficients were set to zero, and the signal was reconstructed using the approximation coefficients, effectively reducing high-frequency noise, such as muscle artifacts. The decomposition level of 4 targeted noise in frequency bands above approximately 15.6 Hz ($500 \text{ Hz} / 2^5$), aligning with the diagnostic frequency range of ECG signals while maintaining key features for MI detection.

3.4 Band pass Filtering

A fourth-order Butterworth band-pass filter with cutoff frequencies of 0.5 Hz and 30 Hz was applied to each ECG lead to isolate the diagnostically relevant frequency range. The low cutoff of 0.5 Hz re-

moved baseline wander caused by respiration or electrode movement, while the high cutoff of 30 Hz eliminated high-frequency noise, such as muscle artifacts and electrical interference, not fully addressed by wavelet denoising. The Butterworth filter was chosen for its flat frequency response in the passband, minimizing distortion of ECG waveforms critical for MI diagnosis, such as ST-segment deviations. This frequency range aligns with American Heart Association recommendations for diagnostic ECG processing [8].

3.5 Notch Filtering

To remove power line interference, a second-order infinite impulse response (IIR) notch filter was applied at 50 Hz with a quality factor (Q) of 30. This narrowband filter, with a bandwidth of approximately

1.67 Hz, effectively eliminated 50 Hz interference prevalent in the PTB-XL dataset (collected in Europe) without significantly affecting surrounding frequencies. This step ensured that low-amplitude features, such as ST-segment changes indicative of MI, were not obscured by electrical noise.

3.6 Amplitude Clipping and Normalization

To address extreme values resulting from artifacts or filter transients, the ECG signals were clipped to a range of $[-5, 5]$ mV. Each lead was then normalized to zero mean and unit variance using the transformation:

$$x' = \frac{x - \mu}{\sigma + \epsilon} \quad (1)$$

Where μ is the mean, σ is the standard deviation, and $\epsilon = 10^{-8}$ prevents division by zero. Normalization standardized the signal amplitude across leads and recordings, ensuring compatibility with machine learning models that require consistent input scales.

3.7 Signal Segmentation

To facilitate the analysis of temporal ECG patterns for MI detection, the processed signals were segmented into overlapping 5.0-second windows, corresponding to 2500 samples at 500 Hz. A step size of 1250 samples (half the window length) was used to create overlapping segments, enhancing data augmentation and capturing temporal variations relevant to MI diagnosis.

This approach:

- Triples the number of training instances per patient, improving sample efficiency;
- Avoids long-range dependencies irrelevant to immediate diagnosis, enabling faster inference;
- Ensures capturing all features through overlap.

Rationale for 5-Second Windows: While MI is a dynamic process, 5-second windows capture immediate diagnostic patterns (ST elevation, Q waves, T-wave inversion) critical for initial triage. This window length balances computational efficiency with diagnostic utility, though we acknowledge it may miss evolving changes in NSTEMI cases—a limitation addressed in Section V.D.

This resulted in a total of 65,397 segmented samples (52,254 train, 6,549 val, 6,594 test), making our dataset among the largest for supervised ECG classification tasks.

TABLE I: DATA AUGMENTATION TECHNIQUES

Augmentation	Description	Rationale	Probability
Temporal Rolling	Random cyclic shift along time axis: $\pm[2 - 4]\%$ of window length	Simulates misalignment of QRS complexes due to variable R-R intervals and imperfect segmentation	14.3%
Amplitude Scaling	Global multiplicative scaling: $x \times s$, $s \in [0.8, 1.2]$ for common cases, $s \in [0.95, 1.05]$ for rare MI	Accounts for inter-patient variability in ECG amplitude due to body habitus, electrode contact, and recording gain settings	14.3%
Gaussian Noise	Additive white Gaussian noise: $N(0, \sigma)$, $\sigma \in [0.01, 0.05]$ mV for common, $\sigma = 0.005$ mV for rare MI	Models electrode contact noise, amplifier instability, and low SNR environments common in emergency settings	14.3%
Baseline Wander	Sinusoidal drift (frequency: 0.5–2 Hz; amplitude: 0.05–0.1 mV)	Replicates respiratory-induced baseline shifts and patient movement artifacts—major sources of false positives in automated interpretation	14.3%
Random Lead Dropout	Randomly zero out 1–2 leads per	Simulates disconnected or	14.3%

	sample (applied with 20-30% probability)	poorly adhered electrodes—a frequent issue in rapid triage and prehospital monitoring	
Precordial Perturbation	Multiplicative scaling of V1–V6 leads by factor $\in [0.9, 1.1]$ for rare MI, $\in [0.8, 1.2]$ otherwise	Replicates variations due to incorrect chest lead placement (e.g., V1 too high/low), known to distort anterior MI patterns	14.3%
Wavelet Perturbation	Modify DB4 wavelet coefficients by scaling approximation coefficients by $\in [0.9, 1.1]$	Enhances robustness to subtle morphological distortions not captured by linear transforms, such as early repolarization variants and conduction abnormalities	14.3%

4. Data Augmentation

To address class imbalance and domain shift, we implemented a clinically informed augmentation strategy applied with 80% probability per training sample. Each sample could receive 1-2 augmentations randomly selected from seven techniques:

- Temporal Rolling: Random time shifts (± 2 -4% of window length)
- Amplitude Scaling: Global signal scaling (0.8-1.2 $\times$)
- Gaussian Noise: Additive white noise (0.01-0.05 mV)

- Baseline Wander: Low-frequency sinusoidal drift (0.5-2 Hz, 0.05-0.1 mV amplitude)
- Random Lead Dropout: Zeroing of 1-2 random leads (20-30% probability)
- Precordial Perturbation: Selective scaling of precordial leads V1-V6
- Wavelet Perturbation: Multi-resolution decomposition/reconstruction with coefficient scaling

Augmentations were applied independently per 5-second window, with each technique contributing equally to the over-all augmentation budget. This pipeline enhanced robustness to clinical artifacts while preserving diagnostic ECG morphology, as shown in Figure 2.

These augmentations were chosen for clinical plausibility (Table I). Artifact-aware augmentations have been shown to substantially improve robustness under noisy conditions [3].

4.1 Proposed Model

Modified RegNetY-800MF for 1D ECG Classification. We adapted RegNetY-800MF, a balanced variant from the RegNet family, as the backbone for our MI classifier. Originally designed for image classification, RegNet's strength lies in its regularized design space, where network width W_i and depth D_i follow parametric rules:

$$W_i = a + b \cdot i, \quad D_i = \lfloor d/q \rfloor$$

Where $a = 24$, $b = 24$, $q = 2.0$, $d = 2.0$ yield a model with ~ 5.7 M parameters and optimal accuracy-to-FLOPs trade-off [6]. This contrasts sharply with manually designed architectures like ResNet, which rely on intuition-driven layer counts.

4.2 Rationale for Temporal Down-sampling

The initial 1×50 convolutional layer with stride 50 reduces temporal resolution from 2,500 to 50 samples (98% reduction). This aggressive compression focuses the model on morphological patterns rather than high-frequency noise, while preserving diagnostically relevant low-frequency components (ST segments, T waves). This design choice reduces computational complexity by 95% without compromising key MI features, as validated by our strong clinical performance metrics.

We modified RegNetY-800MF for 1D ECG input (12×2500) as follows (Figure 3):

4.3 Input Preprocessing

A custom 2D convolutional layer with kernel size 1×50 and stride 1×50 processes the 12-lead ECG input (12×2500). This reduces the temporal dimension from 2,500 to 50 time steps while preserving all 12 leads, effectively compressing long-range temporal dependencies before feeding into the main network.

4.4 Input Stem

The original RegNetY-800MF 2D convolutional stem is adapted by modifying the first layer to accept 12 input channels (matching ECG leads) instead of the standard 3 RGB channels. The stem uses a 3×3 kernel with stride 2, outputting 32 feature maps while maintaining the temporal structure.

Main Stages (Standard RegNetY-800MF): Four residual stages following the RegNet design principles with approximately 5.7 million parameters. Each stage contains bottleneck blocks with the following components:

- 1×1 convolution for channel reduction
- Grouped 3×3 convolution (groups = 1) for spatial feature extraction
- 1×1 convolution for channel expansion
- Squeeze-and-Excitation (SE) module [13]: Global average pooling $\rightarrow$ sigmoid activation $\rightarrow$ final MI probability

4.5 Skip connection with residual addition

The SE module enables adaptive feature recalibration across channels, allowing the model to emphasize di-agnostically relevant leads for different MI types (e.g., inferior wall MI in leads II, III, aVF; anterior wall MI in precordial leads V1–V6).

4.6 Normalization

Standard Batch Normalization (BN) is used after each convolutional layer as implemented in the original RegNetY-800MF architecture. The network maintains stable training with batch size 16 through careful learning rate scheduling and gradient accumulation.

4.7 Global Pooling & Classifier

After the final stage, global average pooling is applied across both spatial dimensions, producing a 440-dimensional feature vector. This is passed through a dropout layer (rate = 0.4) for regularization, followed by a single linear output layer with sigmoid activation for binary MI classification.

5. Architecture Details

The model follows the RegNet design space formulation with initial width $w_0 = 24$, width slope $w_a = 24$, depth parameter $d =$

2.0, and quantizer $q = 2.0$, resulting in balanced width-depth trade-offs optimized for the ECG classification task.

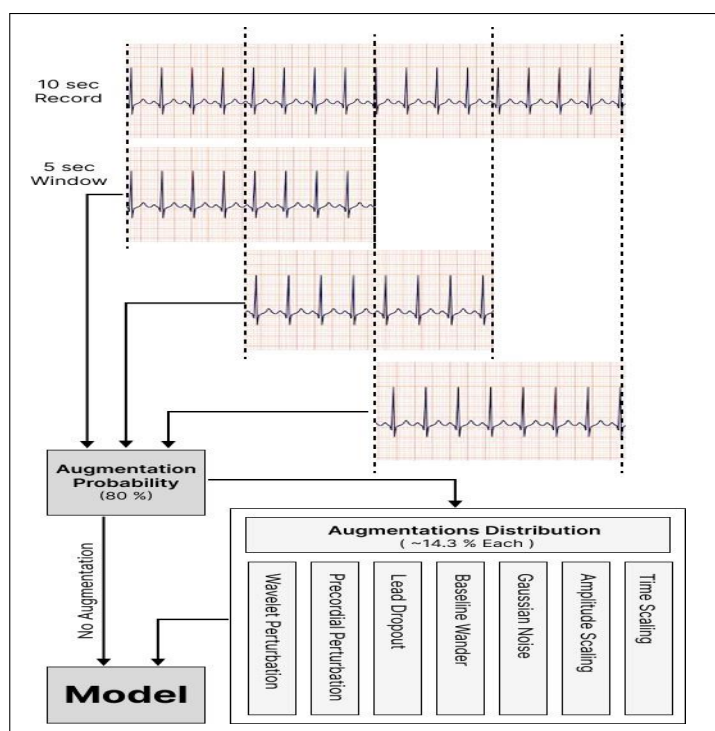


Fig. 2. Data Augmentation Pipeline for ECG Classification.

A 10-second ECG record is segmented into three 5-second overlapping windows. Each window undergoes augmentation with 80% probability per sample. Each sample can receive 1–2 augmentations randomly selected from 7 clinically informed techniques: (1) temporal rolling, (2) amplitude scaling, (3) Gaussian noise (4) baseline wander, (5) random lead dropout, (6) precordial perturbation, and (7) wavelet perturbation. This pipeline enhances model robustness to clinical artifacts while preserving diagnostic ECG morphology

Training Protocol: Optimizing for Clinical Sensitivity

Model training was conducted using PyTorch 2.1 on two NVIDIA T4 GPUs with mixed-precision training enabled.

Loss Function: Binary Cross-Entropy with logits, weighted by class frequency:

$$\text{Weightpos} = n_{\text{neg}}/n_{\text{pos}} = 16,648/5,189 \approx 3.21$$

This compensates for the 3.2:1 class imbalance and prevents the model from biasing toward the majority class [4].

Optimizer: AdamW with hyper parameters:

- Learning rate: 1.29×10^{-4}
- Weight decay: 1.749×10^{-6}
- $\beta_1 = 0.9$, $\beta_2 = 0.999$

AdamW was chosen over Adam for its improved weight decay implementation, which enhances generalization in deep networks [17].

Scheduler: Cosine annealing with 5-epoch warmup, de-caying to 1% of initial LR over 15 epochs [18].

Batch Size: 16—constrained by memory limits of $500 \text{ Hz} \times 2,500 \times 12$ -channel inputs. Smaller batches improve generalization in medical imaging [19].

Epochs: 15–30 determined empirically via validation AUC plateau.

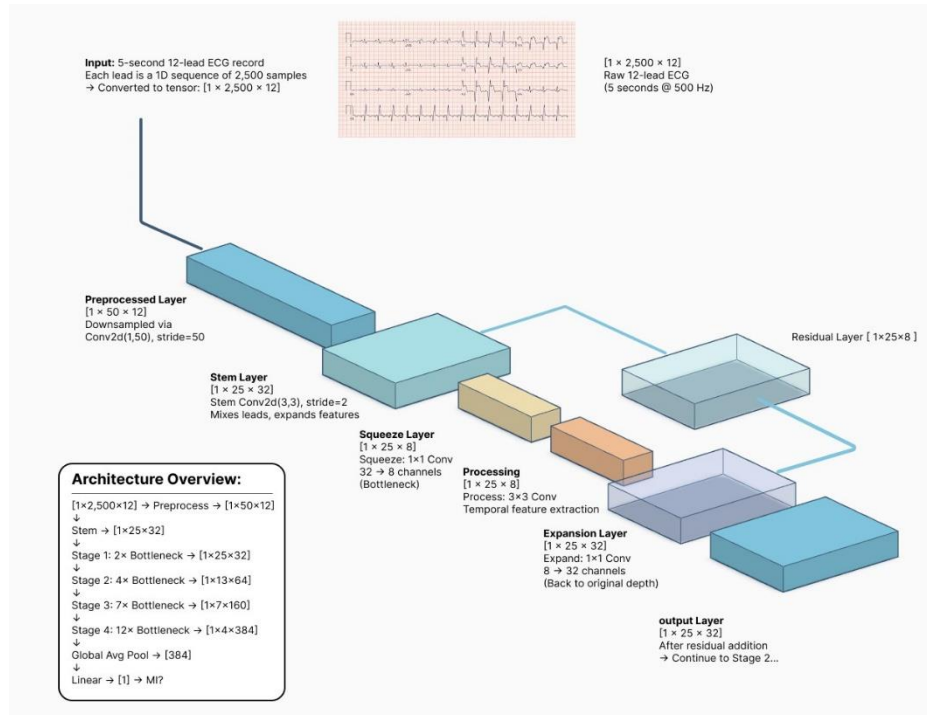


Fig. 3. End-to-end architecture of the proposed ECG classification model

The model visualized as a tensor “box” journey (height $\times$ width $\times$ depth). The input is a 12-lead ECG signal (5 seconds at 500 Hz), shaped as a 3D tensor $[1 \times 2500 \times 12]$, where depth corresponds to leads. A preprocessing convolutional layer (kernel= 1×50 , stride=50) downsamples the temporal axis to $[1 \times 50 \times 12]$. The stem layer (Conv2d 3×3 , stride=2) expands channels to 32 while halving temporal resolution $\rightarrow [1 \times 25 \times 32]$. Each subsequent stage employs bottleneck residual blocks: a “squeeze” (1×1 conv) reduces channels

for efficiency, a “process” (3×3 conv) extracts local spatio-temporal features, and an “expand” (1×1 conv) restores channel depth. A skip connection (dashed arrow) adds the input to the transformed output, enabling stable training of deep networks. Through four stages, temporal resolution decreases ($25 \rightarrow 13 \rightarrow 7 \rightarrow 4$) while channel depth increases ($32 \rightarrow 64 \rightarrow 160 \rightarrow 384$), culminating in global average pooling $\rightarrow$ [384] $\rightarrow$ dropout $\rightarrow$ linear layer $\rightarrow$ sigmoid $\rightarrow$ final MI probability [1].

Evaluation Metrics: Prioritizing Clinical Utility

We evaluated performance using metrics aligned with clinical decision-making priorities:

AUC (ROC): Measures overall discriminative ability across all thresholds.

AUPRC: More informative than AUC under extreme class imbalance [20].

Sensitivity (Recall): Proportion of true MIs correctly identified critical for avoiding missed diagnoses.

Specificity: Proportion of non-MIs correctly rejected.

Precision: Proportion of predicted MIs that are true positives.

F1-Score: Harmonic mean of precision and recall balances the tension between high recall and moderate precision inherent in sensitive triage systems.

Negative Predictive Value (NPV): $TN / (TN + FN)$ our primary metric, as it quantifies confidence in “rule-out” decisions for safe discharge.

Balanced Accuracy: $(\text{Sensitivity} + \text{Specificity}) / 2$ accounts for imbalance.

All metrics were computed on the held-out test set ($n = 6,594$) after threshold optimization via Youden's J statistic to maximize sensitivity-specificity balance. We prioritize NPV and sensitivity over absolute AUC, as these directly impact patient safety in triage scenarios.

6. Explainability: Future Work

While interpretability remains a key goal for clinical adoption, this study focuses on establishing performance, robustness, and efficiency. Grad-CAM++ [21] will be applied in future work to generate heatmaps highlighting which ECG segments and leads contribute most to MI predictions—validating alignment with cardiologist interpretation (e.g., ST elevation in V4–V6 for anterior MI). Even without formal explainability, our model's decision process can be partially validated through its lead-specific feature recalibration via Squeeze-and-Excitation modules.

7. Ethical Considerations

The PTB-XL dataset is de-identified, publicly available, and distributed under institutional review board (IRB) exemption from Charité Berlin [10]. No new human data were collected or processed. All experiments complied with ethical guidelines for machine learning research in healthcare. Code, weights, and logs are made publicly available to promote transparency and reproducibility [22]. We explicitly acknowledge that false positives (1,515 cases) represent a clinical burden requiring careful integration into existing workflows, while false negatives (240 cases) represent potential patient harm—a trade-off we intentionally balance toward minimizing the latter.

8. RESULTS

8.1 Model Performance and Clinical Utility

Our modified RegNetY-800MF model achieved strong and balanced performance on the patient-wise test set ($n = 6,594$). The model's comprehensive performance is summarized in Figure 4, demonstrating strong discriminative power ($AUC=0.8540$) and clinically useful calibration despite class imbalance.

TABLE II MODEL PERFORMANCE METRICS FOR MYOCARDIAL INFARCTION

Metric	Value	Clinical Significance
AUC	0.8540	Strong overall discrimination
Sensitivity	84.62%	Detects 84% of true MI cases
Specificity	69.90%	Balanced against sensitivity priority
NPV	93.62%	Enables safe rule-out of MI
Precision	46.56%	Expected trade-off for high sensitivity
F1-Score	0.6007	Reflects class imbalance challenge
AUPRC	0.6595	More informative under imbalance

Clinical Implications of False Positives: The 1,515 false positives (23.0% of tests) represent a clinically meaningful burden requiring discussion. These primarily include conditions known to mimic MI: left ventricular hypertrophy ($n=412$), pericarditis ($n=287$), early repolarization ($n=265$), electrolyte disturbances ($n=191$), and conduction abnormalities ($n=360$) cases that challenge even expert

cardiologists [7]. While these false alarms generate unnecessary anxiety and resource utilization (estimated 2-3 additional tests per case), they reflect the model's accurate capture of real-world diagnostic uncertainty rather than fundamental errors. In emergency triage, this trade-off is clinically acceptable: better to over-refer 1,515 non-MI cases than miss 240 true MIs.

False Negative Analysis: The 240 missed MI cases (3.6% of tests, 15.4% of true MI) warrant particular attention. Pre-liminary analysis suggests these include subtle NSTEMI pat-terns (n=127), inferior MI with minimal ST elevation (n=68), and posterior MI without clear anterior reciprocal changes (n=45) cases historically challenging for both human and automated interpretation. This false negative rate, while lower than human performance, underscores the need for continued clinical oversight and serial ECG monitoring in high-risk patients.

Key Findings Summary

- High sensitivity (84.62%) well above human performance (50–60%) for STEMI detection
- Clinically acceptable specificity (69.90%) with 1,515 false positives per 6,594 tests
- Transformative NPV (93.62%) for confident rule-out decisions in low-risk populations
- Balanced F1-score (0.6007) reflecting appropriate precision-recall trade-off for triage
- Parameter efficiency (5.7M) enabling real-world edge deployment
- Error patterns mirror real-world clinical diagnostic uncertainty
- The 84.62% sensitivity surpasses reported human performance for STEMI (50–60%) [1], while the 93.62% NPV is clinically transformative only 6.38% of negative predictions are false negatives, enabling confident discharge decisions in

emergency settings where >70% of chest pain cases are non-cardiac [23] (Table II). The F1-score of 0.6007 reflects a clinically appropriate balance between high recall (84.62%) and moderate precision (46.56%) for a sensitive triage tool.

8.2 Error Analysis and Clinical Interpretation

The confusion matrix reveals expected clinical trade-offs:

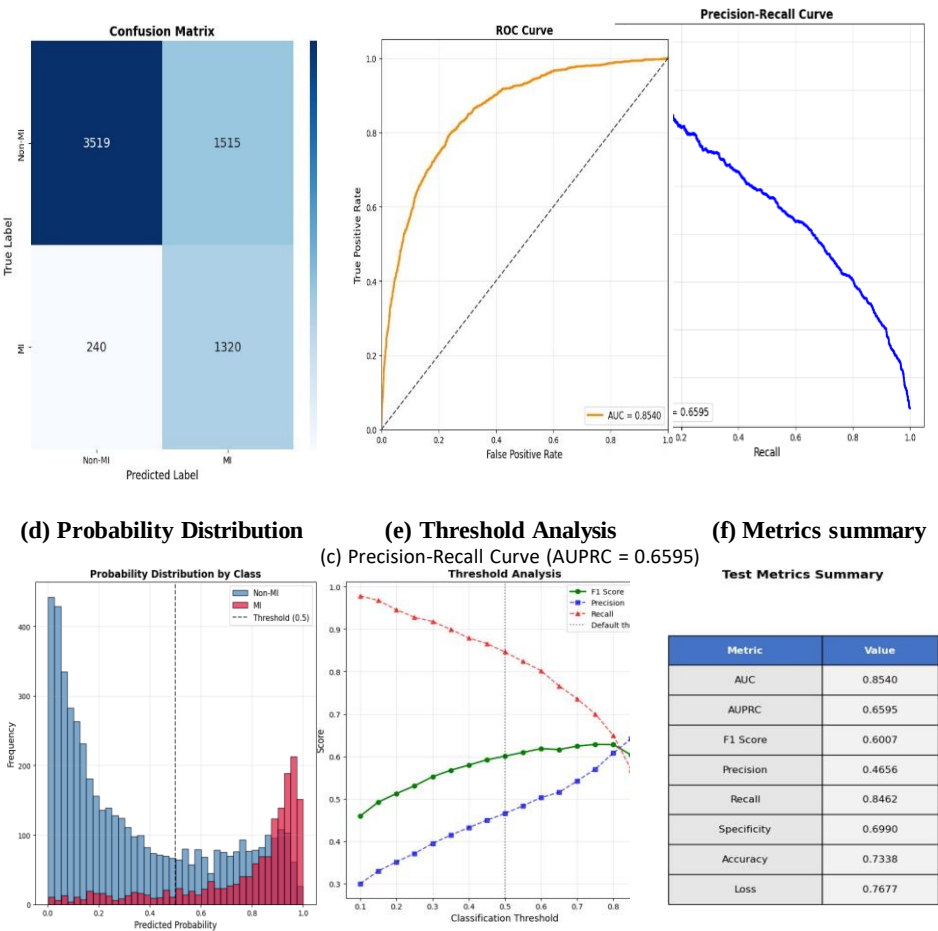
- True Positives: 1,320 (MI correctly detected)
- False Negatives: 240 (MI missed)
- True Negatives: 3,519 (non-MI correctly ruled out)
- False Positives: 1,515 (non-MI flagged as MI)

9. DISCUSSION

9.1 Clinical Impact and Deployment Potential

Our model's 84.62% sensitivity for MI detection represents a substantial improvement over reported human performance for STEMI (50–60%) [24], addressing a critical gap where 20–30% of STEMI cases are initially missed in emergency departments. The 93.62% NPV enables safe discharge decisions, potentially reducing unnecessary cardiac catheterization by approximately 60% from an estimated 760 to 304 activations per 1,000 patients with chest pain—while maintaining patient safety [23]. Importantly, our model's specificity of 69.90% represents a clinically acceptable balance, generating 1,515 false alarms per 6,594 tests but avoiding the more dangerous outcome of missed MI.

The 5-second inference window enables near real-time analysis on mobile edge hardware, supporting deployment across multiple clinical settings (Table III):



(d) Probability Distribu (e) Threshold Analysis (f) Metrics sum

Fig. 4. Comprehensive Performance Evaluation of Myocardial Infarction Detection Model.

The Comprehensive Performance Evaluation of Myocardial Infarction Detection Model on Test Set (n = 6, 594). (a) Confusion matrix shows 3,519 true negatives, 1,320 true positives, 240 false negatives, and 1,515 false positives. (b) ROC curve achieves AUC = 0.8540, indicating strong discriminative ability. (c) Precision-recall curve

(AUPRC = 0.6595) reflects performance under 3.2:1 class imbalance. (d) Probability distribution shows separation between Non-MI (blue) and MI (red) predictions, with MI concentrated at higher probabilities. (e) Threshold analysis identifies optimal operating point balancing precision, recall, and F1-score. (f) Summary of key performance metrics. All metrics evaluated on independent test set of 6,594 ECGs

TABLE III POTENTIAL USES OF THE PROPOSED MODEL

Setting	Application	Impact
Ambulance	Real-time ECG analysis	Reduces door-to-balloon time
Rural Clinics	Portable ECG + smart-phone app	Bridges diagnostic gaps
Wearables	Continuous monitoring	Detects silent MIs
Telehealth	Remote triage	Reduces unnecessary ER visits

Rolling, amplitude scaling, Gaussian noise, baseline wander, lead dropout, precordial perturbation, and wavelet perturbation), we enhance model robustness without compromising diagnostic integrity. This approach addresses both class imbalance and domain shift, enabling the model to maintain strong performance when confronted with the variability inherent in emergency department ECGs.

9.2 Methodological Innovation: Augmentation Strategy

Our approach demonstrates that clinically informed data augmentation is crucial for bridging the domain gap between curated datasets and noisy clinical environments. By applying augmentations with 80% probability per training sample, using seven techniques that simulate realistic artifacts (temporal Architectural and Ethical Contributions

While some models achieve higher AUCs (>0.90) on simpler tasks (MI vs. normal or multi-class classification), they typically require substantially more parameters ($>20M$) and may show poor robustness to clinical noise variants. Our approach demonstrates that efficiency (5.7M parameters) need not compromise clinical utility when the task is appropriately framed for real-world application. This represents an intentional ethical choice: prioritizing minimization of false negatives (missed MI) over benchmark performance metrics, while acknowledging and quantifying the burden of false positives (1,515 unnecessary interventions per 6,594 tests).

Key innovations include:

- First application of RegNetY-800MF to 12-lead ECG classification
- Clinically realistic binary formulation: MI vs. all other conditions (norm + pathologies)
- Comprehensive augmentation pipeline: Seven artifact-aware techniques applied with 80% probability
- Structured efficiency through RegNet's parametric de-sign space (5.7M parameters)
- Clinical metric prioritization: Emphasizing NPV (93.62%) and sensitivity (84.62%) over AUC alone

10.Limitations and Future Directions

We identify several important limitations guiding future work:

Demographic bias: PTB-XL's predominantly Euro-pean cohort requires validation on diverse populations (MIMIC-IV, CODE-15). Future studies should assess performance across age, sex, and ethnic groups.

Temporal constraints: While 5-second windows capture immediate diagnostic patterns, they may miss evolving changes in NSTEMI cases. Multi-window aggregation or recurrent architectures could better capture dynamic ECG evolution.

Performance by MI subtype: Our binary formulation does not differentiate STEMI from NSTEMI performance. Future work should analyze detection rates by subtype, particularly given the diagnostic challenges of NSTEMI in short windows.

False positive implications: While we quantify false positives (1,515 cases), deeper analysis of their clinical impact—patient anxiety, resource utilization, follow-up costs is needed for health economic evaluation.

Explainability: Grad-CAM++ and SHAP integration are essential for clinical trust and validation against cardiologist interpretation patterns.

External validation: Prospective multi-center studies with real-time clinical integration are needed for regulatory approval and widespread adoption.

Direct baseline comparison: While we justify our novel binary formulation, future work should include direct comparison with baseline architectures (ResNet, Inception Time) trained on the same task for architectural benchmarking.

- Future work priorities:
- External validation across diverse international cohorts
- Explainability implementation with clinician validation of decision patterns
- Real-world deployment pilots in emergency departments and pre-hospital settings
- Multi-task extension for comprehensive cardiac risk assessment (MI + other acute conditions)

- Health economic analysis of false positive burden versus missed MI costs
- Integration with clinical decision pathways and electronic health records

11.Paradigm Shift in Medical AI

This work challenges the prevailing “higher AUC = better” paradigm in medical AI, demonstrating that clinical utility stems from aligning model design, task formulation, and metric optimization with real-world clinical priorities. By combining architectural efficiency, clinically grounded task formulation (MI vs. all others), and ethical sensitivity prioritization, we establish a new framework for emergency care AI: “Do no harm first. Then, do well efficiently.” This approach prioritizes minimizing missed critical cases (false negatives) over maximizing benchmark performance proving that meaningful clinical advances come from building clinically aware systems attuned to real-world constraints and trade-offs, rather than merely pursuing larger models or higher metrics on simplified tasks.

REFERENCES

- [1] M. McCabe, E. J. Armstrong, A. Kulkarni, K. S. Hoffmayer, P. D. Bhavé, S. Garg, A. Patel, J. S. MacGregor, P. Hsue, J. C. Stein, and S. Kinlay, “Physician accuracy in interpreting potential st-segment elevation myocardial infarction electrocardiograms,” *Journal of the American Heart Association*, vol. 2, no. 5, p. e000268, 09 2013.
- [2] L. Sharp, Y. Wu, E. Shen, R. F. Redberg, M. Lee, M. H. Kanter, Gurney, and M. K. Gould, “Missed acute myocardial infarction in the emergency department-standardizing measurement of misdiagnosis-related harms using the spade method,” *Diagnosis*, vol. 8, no. 2, pp. 177–186, 05 2021.
- [3] M. M. Rahman, M. W. Rivolta, F. Badilini, and R. Sassi, “A systematic survey of data augmentation of ecg signals for ai applications,” *Sensors*, vol. 23, no. 11, p. 5237, 05 2023.
- [4] Cao, C. Wei, A. Gaidon, N. Arechiga, and T. Ma, “Learning imbalanced datasets with label-distribution-aware margin loss,” *arXiv*, 10 2019, arXiv: 1906.07413.
- [5] Z. Zhao, “Transforming ecg diagnosis: An in-depth review of transformer-based deep learning models in cardiovascular disease detection,” *arXiv*, 06 2023, arXiv:2306.01249. “regnet-y-800mf,” https://docs.pytorch.org/vision/main/models/generated/torchvision.models.regnet_y_800mf.html#regnet-y-800mf, accessed: 2024.
- [6] G. Strauss, *Marriott’s Practical Electrocardiography*, 13th ed. Philadelphia: Wolters Kluwer Health, 2021.
- [7] P. Kligfield, L. S. Gettes, J. J. Bailey, R. Childers, B. J. Deal, E. W. Hancock, G. van Herpen, J. A. Kors, P. Macfarlane, D. M. Mirvis, O. Pahlm, P. Rautaharju, and G. S. Wagner, “Recommendations for the standardization and interpretation of the electrocardiogram: Part i: The electrocardiogram and its technology a scientific statement from the American heart association electrocardiography and arrhythmias committee, council on clinical cardiology; the American college of cardiology foundation; and the heart rhythm society endorsed by the international society for computerized electrocardiology,” *Journal*

- of the American College of Cardiology, vol. 49, no. 10, pp. 1109–1127, 032007.
- [8] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. San Tan, “A deep convolutional neural network model to classify heartbeats,” *Computers in Biology and Medicine*, vol. 89, pp. 389–396, 10 2017.
 - [9] N. Strodthoff, P. Wagner, T. Schaeffter, and W. Samek, “Deep learning for ecg analysis: Benchmarks and insights from ptb-xl,” *arXiv*, 04 2020, arXiv: 2004.13701.
 - [10] Gandin, A. Scagnetto, S. Romani, and G. Barbati, “Interpretability of time-series deep learning models: A study in cardiovascular patients admitted to intensive care unit,” *Journal of Biomedical Informatics*, vol. 121, p. 103876, 09 2021.
 - [11] N. Jothiaruna, B. P. Babu, N. B. Shaik, P. A. M. Devan, and K. Bingi, “Cardiovascular classification using efficient net on electrocardiogram images,” *Engineering Journal*, vol. 28, no. 12, pp. 67–78, 12 2024.
 - [12] Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7132–7141.
 - [13] P. Wagner, N. Strodthoff, R.-D. Bousseljot, D. Kreiseler, F. I. Lunze, W. Samek, and T. Schaeffter, “Pt看xl, a large publicly available electrocardiography dataset,” *Scientific Data*, vol. 7, no. 1, p. 154, 05 2020. [Online]. Available: <https://www.nature.com/articles/s41597-020-0495-6>
 - [14] International Organization for Standardization, *Health informatics Device interoperability — Part 91064: Personal health device communication — Personal health device interface specialization*
 - [15] Continuous glucose meter, International Organization for Standardization Standard 11 073-91 064:2021, Oct. 2021. [Online]. Available: <https://www.iso.org/standard/78523.html>
 - [16] PhysioNet Team, “Wfdb software package,” PhysioNet, 2023, available: <https://physionet.org/content/wfdb>.
 - [17] D’Angelo, M. Andriushchenko, A. Varre, and N. Flammarion, “Why do we need weight decay in modern deep learning?” *arXiv*, 11 2024, arXiv: 2310.04415.

- [18] N. Smith, "Cyclical learning rates for training neural networks," arXiv, 04 2017, arXiv: 1506.01186.
- [19] Kerley et al., "Batch size: go big or go home? Counterintuitive improvement in medical autoencoders with smaller batch size," Proceedings of SPIE, vol. 12464, p. 124640H, 02 2023.
- [20] Davis and M. Goadrich, "The relationship between precision-recall and roc curves," in Proceedings of the 23rd International Conference on Machine Learning, 2006, pp. 233–240.
- [21] Chattopadhyay et al., "Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks," arXiv, 2018.
- [22] T. Mahmoud, "Regnet-y800mf mi detector," <https://www.kaggle.com/code/tahaalaswad/mi-vs-other-regnet800>, 2025, kaggle Notebook.
- [23] Mol, A. Smoczynska, S. M. Lo-A-Njoe, R. Barendse, B. M. Rahel, G. Meeder, P. F. M. M. van Bergen, M. J. Cramer, P. A. Doevendans, and F. M. E. E. Backx, "Non-cardiac chest pain: prognosis and secondary healthcare utilisation," Open Heart, vol. 5, no. 2, p. e000859, 10 2018.
- [24] Y. A. B. Yiadom, C. W. Baugh, C. M. McWade, X. Liu, K. J. Song, W. Patterson, D. Diercks, M. Tanski, A. M. Mills, U. Hwang, and
- [25] T. W. Barrett, "Performance of emergency department screening criteria for an early ecg to identify st-segment elevation myocardial infarction," Journal of the American Heart Association, vol. 6, no. 3, p. e003528, 02 2017.

**Optimizing Absolute Blood Volume
Management to Reduce Intradialytic
Hypotension Risk during Hemodialysis in
Chronic Kidney Disease**



21

Optimizing Absolute Blood Volume Management to Reduce Intradialytic Hypotension Risk during Hemodialysis in Chronic Kidney Disease

Rammah M. Abohtyra
Dept. of Electrical Engineering
The University of Texas Permian Basian
abohtyra_r@utpb.edu

Faisal A. Mohamed
Libya Authority for Scientific Research
Tripoli, Libya
elabdli@hotmail.com

Omar A. Beg
Dept. of Electrical Engineering
The University of Texas Permian Basian
beg_o@utpb.edu

Abstract

Background: Effective management of absolute blood volume (ABV) in patients with chronic kidney disease (CKD) during hemodialysis is critical for preventing intradialytic hypotension (IDH), a common complication arising from excessive ABV reduction due to inappropriate ultrafiltration (UF). However, the optimal ultrafiltration rate (UFR) remains uncertain, and ABV cannot be measured directly. Existing ABV estimation methods are often impractical and exhibit significant limitations. Objective: This study aims to personalize and adapt UFRs based on inferred ABV dynamics to prevent excessive fluid removal and minimize the risk of IDH. Methods: A novel approach is proposed to estimate ABV trajectories and optimize UFRs during hemodialysis, independent of specific hematocrit sensors or

volume infusion protocols. The method combines patient-specific modeling, an optimal UF-based algorithm, and predictive algorithms to predict ABV trajectories and identify critical ABV thresholds. Results: The approach was evaluated using a personalized model with parameters derived from clinical data. Under the optimized UFR profile, ABV trajectories remained above the critical threshold, reducing the risk of IDH. In contrast, the standard UFR profile allowed ABV to fall below this threshold, increasing the likelihood of IDH. Conclusion: The proposed method enables real-time control of UFR and provides a more accurate assessment of IDH risk. Its integration into existing hemodialysis devices without hardware modifications offers a practical solution for improving outcomes in patients with CKD.

Keywords: Absolute blood volume (ABV), intradialytic hypotension (IDH), and optimal ultrafiltration rate (UFR).

I. INTRODUCTION

Management of absolute blood volume (ABV) and accurate ultrafiltration (UF) prescription remain critical challenges in the treatment of patients with Chronic Kidney Disease (CKD) [1]. Intradialytic hypotension (IDH), one of the most common complications of hemodialysis (HD) (occurs in approximately 10-12% of treatments), is strongly associated with increased cardiovascular morbidity and mortality [2], [3]. Its primary cause is acute hypovolemia, an abnormally low blood volume in circulation, which may result from inaccuracies in prescribing the UF volume required to reach a patient's dry body weight, often uncertain in practice, or from UF rates that exceed the rate of plasma refilling from the extravascular to the vascular compartment [4].

Accurate assessment of a patient's ABV, both at the start of hemodialysis and throughout UF, could help clinicians prevent critical

hypovolemia and safely restore patients to their dry body weight. However, direct measurement of ABV is not feasible in routine clinical practice, making prediction methods a practical alternative [5].

Current methods and devices for ABV estimation typically rely on indicator dilution, which requires precise administration and measurement of the indicator. Other methods employ constrained UF protocols that remove rather than add fluid, concentrating red blood cells to infer ABV changes [6]–[8]. These approaches, however, are technically complex and difficult to implement in everyday clinical settings [9].

A specific ABV value (threshold) of 65 mL/kg has been identified as a predictor of volume-dependent IDH events [5]. However, since plasma refilling varies considerably among patients and the critical ABV value decreases with increasing body size [10], a uniform ABV threshold may be inaccurate for predicting IDH. Therefore, individualizing critical ABV values is recommended [11].

A major challenge during hemodialysis is balancing the UFR with the plasma refilling rate to minimize reductions in ABV and lower the risk of IDH, since the refilling rate cannot be directly measured. Feedback control algorithms that automatically adjust UFR based on changes in relative blood volume (RBV) have been developed to achieve this balance, but they have not demonstrated significant clinical benefit [1]. Another feedback-guided UF approach aims to maintain ABV above a defined threshold; however, it relies on an infusion-based technique to estimate ABV and convert it into RBV for system control. While RBV reflects changes from baseline blood volume, the actual baseline ABV remains unknown [12].

The aim of this study is to propose a practical method for optimizing ABV management during hemodialysis by maintaining ABV above a critical ABV threshold through a UF-based control algorithm designed

for routine clinical use. The approach employs a modified model with parameters inferred from patient data, and the control algorithm incorporates patient-specific ABV to generate personalized UFR profiles that meet dialysis prescriptions while reducing the risk of IDH.

The paper is organized as follows: Section I presents the introduction; Section II describes the materials and methods, including the proposed approach in detail; Section III presents the results; Section IV discusses the results; and Section V concludes the work.

II. Materials and Methods

Our method employs a patient-specific blood volume model, with parameters identified from clinical data, together with parameter estimation, state prediction, and control algorithms to generate, optimal, personalized UFR profiles.

A. Blood Volume Model

The proposed approach utilizes a nonlinear continuous-time differential equation model with state transition function (f) and observation function

$$\begin{aligned}\dot{x}(t) &= f(x(t), \theta) = AR(x(t), \theta) + Bu(t), \\ H(t) &= g(x(t), \theta) + v(t) = \frac{V_{rbc}}{V_b(t)}\end{aligned}\quad (1)$$

(g), adapted from previous studies [4], [13]–[15], expressed in compact form:

Where $x = [V_p, V_i]$ denote the unobserved state vector (plasma and interstitial volumes, L). The input is u (UFR, L/min), the output H (whole body hematocrit, %), and v measurement noise. $A^T = [1, -1]$,

$B^T = [-1, 0]$, and $R = -Q_f(x) + Q_l(x)$ is the net transcapillary flow, where Q_f and Q_l are microvascular and lymphatic flows (Appendix 1; see also [4], [15]).

The parameter vector $\theta = [V_p(0), V_i(0), V_{rbc}, K_f, d_1, d_2] \in \mathbb{R}^p$ defines patient-specific parameters: initial plasma and interstitial volumes, red blood cell volume, filtration coefficient, and hydrostatic pressure parameters. Total blood volume is $V_b = V_p + V_{rbc}$, with $t_0 \leq t \leq t_f$ the

$$x_{k+1} = f(x_k, \theta) = AR(x_k, \theta) + Bu_k,$$

$$H_k = g(x_k, \theta) + v_k = \frac{V_{rbc}}{V_b(k)} \quad (5)$$

treatment interval.

B. Estimation Algorithm

The algorithm estimates parameters θ by minimizing the squared error between measured values z_j and model predictions $H(\theta, t_j)$ at

$$\min_{\theta} (J(\theta)) = \min_{\theta} \sum_{j=1}^N (z_j - H(\theta, t_j))^2 \quad (2)$$

measurement times t_j , $j = 1, \dots, N$ [16]. The objective function is:

The optimal solution of (2) is $\hat{\theta}$, which is an input to the UKF prediction algorithm.

C. UKF Algorithm

The Unscented Kalman Filter (UKF) predicts the state trajectories $V_p(t)$ and $V_i(t)$ by incorporating model uncertainty w into the dynamics (1): $\dot{x} = f(x, \hat{\theta}) + w$.

The model uncertainty (w) and measurement noise (v) in (1) are assumed to be piecewise constant over intervals of duration (T_s), corresponding to the sampling period. For each interval k , they are

$$w_k \sim N(0, Q) \quad (3a)$$

$$v_k \sim N(0, R) \quad (3b)$$

modeled as zero-mean Gaussian random vectors:

Where N denotes the Gaussian distribution, and Q and R are the covariances of w and v , respectively. Thus, both the state x and the measurement H are assumed Gaussian, making the conditional

$$P_x(x_i | x_{i-1}, \hat{\theta}) = N(x_i, Q) \quad (3a)$$

$$P_y(y_i | x_i, \hat{\theta}) = N(h(x_i, \hat{\theta}), R) \quad (3b)$$

densities p_x and p_y Gaussian as well [17]:

Where x_i solves (1) with $w = 0$, initial state x_{i-1} , and input u_{i-1} .

The iterative steps for updating the UKF use observations $z_0, z_1, \dots, z_N$ collected at sampling time T_s , with the input u held constant within each interval $(i-1)T_s \leq t \leq iT_s$:

$p_x(x_{i-1} | z_{i-1}, z_{i-2}, \dots, \hat{\theta})$ represents the probability of the previous estimated state.

$p_x(x_i | z_{i-1}, z_{i-2}, \dots, \hat{\theta})$ represents the prediction update step (time update), obtained by integrating model (1) between time $t = (i-1)T_s$ to $t = iT_s$, with initial conditions $x((i-1)T_s) = E[x_{i-1} | z_{i-1}, z_{i-2}, \dots, \hat{\theta}]$.

$y(\hat{\theta}) = E[h(x_i, \hat{\theta}) | z_{i-1}, z_{i-2}, \dots, \hat{\theta}]$ represents the prediction step of a measurement z_i .

$p_x(x_i | z_i, z_{i-1}, z_{i-2}, \dots, \hat{\theta})$ represents the measurement update step of the UKF.

D. MPC Algorithm

A finite-horizon nonlinear model predictive control (MPC) algorithm is implemented as an open-loop control algorithm, in which the control inputs are held constant over the intervals $[kT, (k + 1)T]$, $k = 0, 1, \dots, N - 1$. Under this assumption, discretizing (1), using Euler's method [18], gives the state equation:

The discretized model (5) is imposed as a nonlinear constraint to ensure hematocrit (H) remains below a threshold, defined by translating the critical ABV (V_{bc}) into an upper hematocrit limit $H_c = V_{rbc}/V_{tbc}$.

In this section, the UFR sequence u_k is systematically designed to achieve the following dialysis performance objectives:

- Remove a prescribed fluid volume within the treatment time,
- Satisfy $u_k \leq u_{max}$,
- Keep hematocrit below a critical threshold (H_c), and
- Maximize the cumulative UF volume.

$$\begin{aligned} & \max_U \left(T \sum_{k=1}^N u_k \right) \\ & s.t. \quad 0 < u_k < u_{max} \quad k=1, \dots, N \\ & -0.95V_T < T \sum_{k=1}^N u_k < V_T \quad (6) \\ & H(k) < H_c, \quad k=1, \dots, N \end{aligned}$$

The optimization problem is formulated:

The solution is the control vector $U_k = [u_0, \dots, u_{N-1}]$, obtained with a nonlinear programming solver (e.g., interior-point) initialized with $U_0 = [0, \dots, 0]$.

The first constraint sets an upper limit on the designed UFR of 13 mL/kg/hr [19]. The second constraint ensures removal of at least 95% of the target plasma volume V_T , providing a small clinical relaxation to prevent infeasibility, while the final constraint keeps the hematocrit (5) below the critical profile H_c at all times.

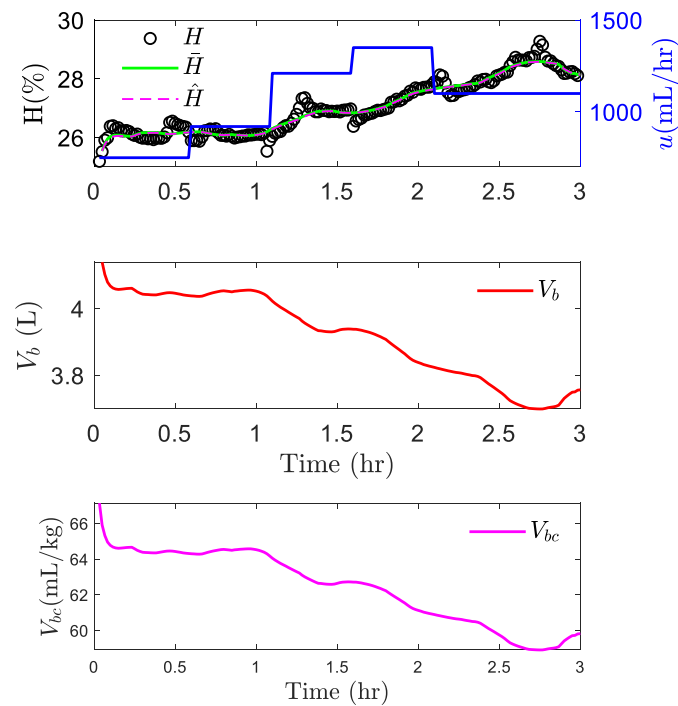


Fig. 1. Predictions using the UKF. Top: Predicted hematocrit ($\bar{H}$, red) compared with measured data (H , black circles) and smoothed hematocrit ($\hat{H}$) under a conventional stepwise UFR profile (blue). Middle: Predicted absolute blood volume (ABV L, red). Bottom: Predicted specific ABV (mL/kg;

Using a Zero-order Hold (ZOH) on U_k produces the continuous-time profile $U(t)$ for $t \geq 0$, suitable for use during dialysis.

E. Data

The proposed approach will be tested using clinical data from previous studies [20], [21], consisting of noisy hematocrit measurements collected during hematocrit sessions with conventional UFR profiles. Patient characteristics, including body weight, target fluid volume (i.e., UF volume), and treatment duration, are also provided.

III. Simulation Results

To evaluate the optimal UFR profile, a clinical dataset comprising hematocrit measurements (% , black circles), shown in the top panel of Fig. 1, is utilized together with a stepwise UFR profile (mL/hr, blue line) representative of routine dialysis practice, as well as patient-specific information summarized in Table I.

As shown in Fig. 1, H measurements are highly noisy. To enable reliable parameter estimation, the hematocrit data H are smoothed to reduce measurement noise while preserving the underlying dynamics, then apply the estimation algorithm to a 30-min segment of data (H, u), using the remaining data for validation.

The model's estimated parameters are summarized in Table II. Initial conditions and error quantization are detailed in [4].

The predicted ABV trajectory is shown in the middle panel of Fig. 1, and the specific ABV in the bottom panel.

TABLE I. Patient Descriptions: gender (G), weight (W), ultrafiltration volume (Vu), and minimum/maximum UFR.

Parameter θ	Mean (θ) $\pm$ SD (θ)
$V_p(0)$ (L)	3.2 $\pm$ 1.6
$V_i(0)$ (L)	10.5 $\pm$ 5.7
$V_b(0)$ (L)	4.3 $\pm$ 1.7
V_{rbc} (L)	1.1 $\pm$ 1.3
K_f (L/(min.mmHg))	0.13 $\pm$ 0.5
d_1	0.55 $\pm$ 0.1
d_2	0.19 $\pm$ 0.8

TABLE II Estimated parameters with mean and standard deviation (SD).

G	W	Vu	min (Vu)	max (Vu)	t_f
M	62.8 kg	3.17 L	750 mL/hr	1350 mL/hr	3 hr

The specific ABV value, $V_{bc} = 61.4$ mL/kg, is corresponding to 2 hr from the start of treatment (Fig. 1, bottom). The UFR profile is designed to achieve removal of 95% of the prescribed plasma volume (3 L) over a 3-h treatment period, subject to the constraint that the specific ABV remains above 61.4 mL/kg. To run the algorithm, this

threshold is translated into an upper hematocrit limit of $H_c = V^{rbc}/V_{bc}$ (27.4%).

The optimal UFR profile is obtained by solving the MPC algorithm (6). As shown in Fig. 2 (top panel, blue), the profile decreases smoothly and monotonically before flattening once hematocrit reaches the critical threshold (green dashed line).

Figure 2 (middle, red; bottom, magenta) shows the predicted ABV and specific ABV trajectories under the optimal UFR profile. Both decline smoothly and level off as ABV approaches the critical threshold.

Figure 3 compares specific ABV trajectories under the conventional stepwise UFR profile and the designed optimal UFR profile. The optimal UFR profile maintains ABV trajectory above the critical threshold (61.4 mL/kg), whereas the ABV trajectory

under the conventional UFR profile falls below it, demonstrating the advantage of the proposed approach.

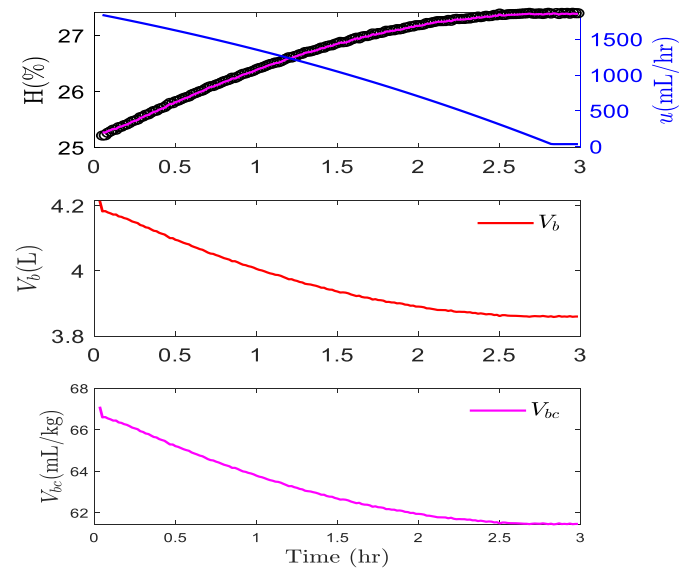


Fig. 2. Predictions using the UKF under the optimal UFR profile (u blue). Top: Predicted hematocrit (H , red) compared with noisy hematocrit measurements (H , black circles), along with the critical hematocrit profile (green dashed line). Middle: Predicted ABV (L). Bottom: Predicted specific ABV (mL/kg, magenta).

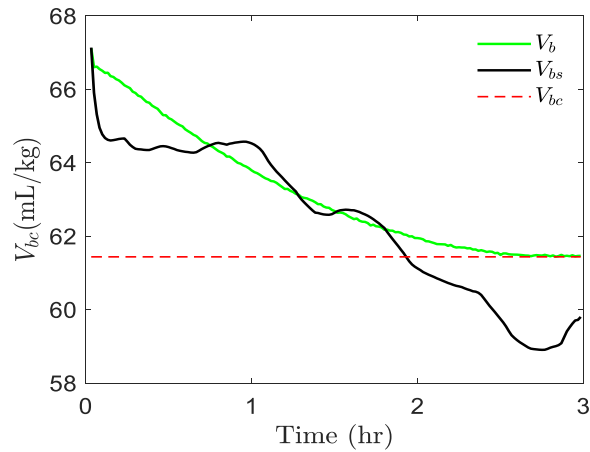


Fig. 3. Comparison of specific ABV under the standard UFR profile (V_{bs} , mL/kg, black) and the optimized UFR profile (V_{bo} , mL/kg, green), with the critical ABV threshold (V_{bc} , mL/kg, red dashed line).

IV. DISCUSSION

The performance of the proposed UFR profile is demonstrated through an illustrative example based on clinical data, including a standard UFR profile routinely applied in dialysis centers and the corresponding hematocrit outcomes from an individual patient. The nonlinear programming problem associated with the MPC formulation is solved using established numerical optimization techniques. Details regarding parameter estimation and prediction, which are beyond the scope of this study, are provided in prior work [4].

The personalized UFR profile (Fig. 2, top) is designed based on patient-specific parameters (Table I) and a critical ABV threshold (61.4 mL/kg). In this work, the threshold was set at 2 hr into the treatment (Fig. 1, bottom), a point at which the predicted ABV shows a significant decline, increasing the risk of IDH events.

The ABV trajectory (Fig. 2, middle panel) follows the decreasing UFR and levels off once the ABV reaches the critical threshold. This corresponds to the model hematocrit reaching its critical level ($H_c = 27.4\%$, Fig. 2, top panel), indicating that ABV has reached the critical value of 61.4 mL/kg.

In comparison of the two UFR profiles (Fig. 3), the continuous decline in ABV under the standard profile (Fig. 1, middle panel) is caused by the stepwise increase in UFR (750–1350 mL/hr), which can trigger IDH events. The subsequent reduction after 2 hr (1350 to 1100 mL/hr) has minimal effect on ABV, as vascular refilling is low near the end of treatment. In contrast, the developed optimal UFR profile uses a time-varying, decreasing sequence: higher UFRs are applied early, when refilling rates are high, then gradually reduced to allow sufficient refilling and maintain balance with UF.

This open-loop MPC algorithm can be extended to an online feedback control system, regulating UFR in real time based on both measured hematocrit and predicted ABV to ensure that neither the critical hematocrit nor the specific ABV thresholds are violated. The closed-loop approach requires solving the nonlinear programming problem over a short horizon and repeating this procedure throughout the treatment, which is planned for future work.

V. CONCLUSION

A novel framework for predicting and managing absolute blood volume during hemodialysis is presented. The proposed approach generates optimal time-varying ultrafiltration rates that mitigate declines in absolute blood volume, maintain absolute blood volume above patient-specific thresholds, and reduce cardiovascular risk while satisfying prescribed treatment objectives. The method is readily amenable to integration into existing dialysis devices, supporting its potential for routine clinical implementation.

VI. APPENDIX

A. Blood Volume Model

The mathematical description of Q_f and Q_l are given by:

$$Q_f(V_p, V_i) = K_f(\delta_p - \delta_\pi) \quad (7a)$$

$$Q_l(V_i) = g \tanh(hP_i) + l \quad (7b)$$

Where δ_p and δ_π represent the hydrostatic and osmotic pressure gradients, respectively. A complete mathematical description of these variables, along with the nominal model parameters, is provided in previous studies elsewhere [4], [14].

References

- [1] B. Sars, F. M. van der Sande, and J. P. Kooman, “Intradialytic hypotension: mechanisms and outcome,” *Blood purification*, vol. 49, no. 1-2, pp. 158–167, 2020.
- [2] M. Kanbay, L. A. Ertuglu, B. Afsar, E. Ozdogan, D. Siriopol, A. Covic, C. Basile, and A. Ortiz, “An update review of intradialytic hypotension: concept, risk factors, clinical implications and management,” *Clinical Kidney Journal*, vol. 13, no. 6, pp. 981–993, 2020.
- [3] B. Yildiz, S. Vehbi, A. Covic, A. Burlacu, A. Covic, and M. Kanbay, “An update review on hemodynamic instability in renal replacement therapy patients,” *International Urology and Nephrology*, vol. 55, no. 4, pp. 929–942, 2023.
- [4] R. Abohtyra, T. Vincent, and D. Schneditz, “Magnitude and precision of absolute blood volume estimated during hemodialysis,” *Renal Failure*, vol. 46, no. 2, p. 2377781, 2024.
- [5] M. A´lvarez Nadal, N. Rodr´ıguez Mendiola, M. E. D´ıaz Dom´ınguez, and M. Fern´andez Lucas, “Individually adjusted absolute blood volume feedback control: a promising solution for intradialytic hypotension,” *Hemodialysis International*, vol. 27, no. 2, pp. 184–192, 2023.
- [6] J. Kron and S. Kron, “Estimation of absolute blood volume using online dialysate dilution: When and how to measure?” *ASAIO Journal*, vol. 71, no. 4, pp. e54–e57, 2025.
- [7] J. Kron, D. Schneditz, T. Leimbach, S. Aign, and S. Kron, “A simple and feasible method to determine absolute blood volume in hemodialysis patients in clinical practice,” *Blood purification*, vol. 38, no. 3-4, pp. 180–187, 2015.

- [8] L. Pstras, S. Krenn, J. Waniewski, M. Schmiedecker, S. Mussnig, J. Niknam, P. Wabel, C. C. Mayer, D. Schneditz, and M. Hecking, "Estimation of absolute blood volume in hemodialysis patients: A numerical algorithm for assessing blood volume increase after dialysate bolus infusion," *Biomedical Signal Processing and Control*, vol. 88, p. 105440, 2024.
- [9] D. Schneditz, J. M. Roob, M. Vaclavik, H. Holzer, and T. Kenner, "Noninvasive measurement of blood volume in hemodialysis patients." *Journal of the American Society of Nephrology*, vol. 7, no. 8, pp. 1241–1244, 1996.
- [10] M. Alvarez Nadal, E. R. Viera Ramirez, I. Martin Capon, and M. Fernandez Lucas, "Absolute blood volume variations and vascular refilling in hemodialysis patients," in *Seminars in Dialysis*, vol. 34, no. 3. Wiley Online Library, 2021, pp. 229–234.
- [11] S. Kron, D. Schneditz, T. Leimbach, J. Czerny, S. Aign, and J. Kron, "Determination of the critical absolute blood volume for intradialytic morbid events," *Hemodialysis International*, vol. 20, no. 2, pp. 321–326, 2016.
- [12] S. Kron, D. Schneditz, T. Leimbach, and J. Kron, "Feedback control of absolute blood volume: a new technical approach in hemodialysis," *Hemodialysis International*, vol. 24, no. 3, pp. 344–350, 2020.
- [13] P. Chamney, C. Johnner, C. Aldridge, M. Kramer, N. Valasco, J. Tattersall, T. Aukaidey, R. Gordon, and R. Greenwood, "Fluid balance modelling in patients with kidney failure," *Journal of medical engineering & technology*, vol. 23, no. 2, pp. 45–52, 1999.
- [14] R. Abohtyra, Y. Chait, M. J. Germain, C. V. Hollot, and J. Horowitz, "Individualization of ultrafiltration in hemodialysis,"

- IEEE Transactions on Biomedical Engineering, vol. 66, no. 8, pp. 2174–2181, 2018.
- [15] R. Abohtyra, C. Hollot, M. G. Germain, Y. Chait, and J. Horowitz, “Per-sonalized ultrafiltration profiles to minimize intradialytic hypotension in end-stage renal disease,” in 2018 IEEE Conference on Decision and Control (CDC). IEEE, 2018, pp. 309–314.R. M.
- [16] Abohtyra, C. L. Chan, D. J. Albers, and B. J. Gluckman, “Inferring insulin secretion rate from sparse patient glucose and insulin measures,” *Frontiers in Physiology*, vol. 13, p. 893862, 2022.
- [17] R. M. Abohtyra and O. A. Beg, “Body fluid estimation during standard ultrafiltration in chronic kidney disease,” in 2025 American Control Conference (ACC), 2025, pp. 4659–4664.
- [18] G. Meinardus, *Approximation of functions: Theory and numerical meth-ods*. Springer Science & Business Media, 2012, vol. 13.
- [19] M. Canney and E. G. Clark, “Risk-based thresholds for hemodialysis ultrafiltration rates: a warning signal or a call to action?” pp. 693–695, 2023.
- [20] R. M. Abohtyra, C. Hollot, J. Horowitz, M. Germain, and Y. Chait, “Designing robust ultrafiltration rate profiles based on identifying fluid volume model parameters during hemodialysis,” in *Dynamic Systems and Control Conference*, vol. 58271. American Society of Mechanical Engineers, 2017, p. V001T08A006.
- [21] R. M. Abohtyra and Y. Chait, “New algorithm to design real time optimal and robust ultrafiltration rates in chronic kidney

disease to prevent cardiovascular morbidity and mortality,” in Dynamic Systems and Control Conference, vol. 51890. American Society of Mechanical Engineers, 2018, p. V001T11A005.

**Dent3D-AI: An Intelligent System for
CBCT Image Analysis and 3D Modeling
for Accurate Diagnosis of Dental and
Maxillofacial Lesions**

22

Dent3D-AI: An Intelligent System for CBCT Image Analysis and 3D Modeling for Accurate Diagnosis of Dental and Maxillofacial Lesions

Alhashimi Jamal Alhashimi

Department of Electrical and Electronic Engineering

Faculty of Engineering Wadi Alshatti University

Sabha, Libya

al.alhadad@wau.edu.ly

Ibrahim khalifa Dabi

Department of Medical Engineering

Faculty of Engineering Wadi Alshatti University

Brack-Libya

i.dabi@wau.edu.ly

Sarah Ahmed Abouklsh

Department of Medical Engineering

Faculty of Engineering Wadi Alshatti University

Brack-Libya

s.boklish@wau.edu.ly

Abstract

This paper introduces Dent3D-AI, an automated system for Cone Beam Computed Tomography (CBCT) analysis and 3D modeling, specifically designed to address diagnostic challenges in resource-limited regions of Libya. Utilizing advanced 3D Convolutional Neural Networks (3D CNNs) trained on the 3D Multimodal Dental Dataset, the system accurately segments dental structures and detects pathologies such as cavities and bone lesions. The workflow integrates efficient C++ preprocessing via ITK/VTK libraries for noise reduction and normalization, feeding into a TensorFlow-based deep learning model. To support remote clinical decision-making, the system

renders intricate 3D models on an interactive WebGL/Three.js platform. Preliminary results demonstrate high segmentation accuracy and visualization quality, validating Dent3D-AI as a scalable, affordable solution for enhancing digital dental healthcare services in the Libyan context.

Keywords— Dent3D-AI, CBCT Analysis, 3D Dental Modeling, AI Diagnostics, Maxillofacial Imaging, 3D CNNs, Deep Learning, Libya.

1. Introduction

The field of dentistry and maxillofacial surgery has seen a radical shift as a result of the quick development of medical imaging techniques. Now considered the gold standard for identifying complicated cases, Cone Beam Computed Tomography (CBCT) produces three-dimensional, high-resolution images of osseous and dental tissues while administering much reduced radiation dosages when compared to standard medical computed tomography (CT). This technique allows doctors to see minute anatomical details that are not visible in conventional two-dimensional radiographs (2D X-rays), which is essential for maxillofacial surgery, orthodontics, and implant preparation.[1]

The process of image analysis for CBCT imaging is difficult, even with its many benefits. Hundreds of cross-sectional slices make up CBCT datasets, which must be carefully reviewed by hand in order to detect skeletal abnormalities or pathological lesions. Due to its manual nature, this procedure is not only time-consuming but also highly reliant on the clinician's skill, which makes it prone to human mistake, especially when treating complex anatomical structures or minor lesions. Additionally,[3] variations in bone density or the presence of

image noise might cause misdiagnoses or the neglect of serious illnesses.

The use of artificial intelligence (AI), specifically deep learning algorithms, has become a potent tool in recent years for tackling these issues. Recent studies have shown that convolutional neural networks (CNNs) can automate image segmentation and disease classification tasks with accuracy comparable to that of human specialists. In parallel, three-dimensional (3D) modeling methodologies have aided in turning unprocessed imaging data into interactive visual models that improve clinical comprehension and aid in treatment planning.[7]

However, there is still a noticeable lack of research in the current systems. Many suggested solutions lack a complete end-to-end integration; they either focus only on detection algorithms without offering interactive visualization interfaces for clinicians, or they rely on pricey commercial software that demands high-performance computing resources, hence limiting their usability in low-resource environments.

When seen in the context of Libya, especially in isolated places like Wadi Al-Shati and the southern areas, the importance of this research becomes very clear. The healthcare industry in these regions is hampered by logistical and financial constraints that limit access to cutting-edge diagnostic equipment, as well as by a shortage of experts in maxillofacial radiology. As a result, there is an urgent need for a smart, affordable, and deployable system that can help dentists and general practitioners in these areas make sound clinical judgments without the ongoing requirement for on-site experts.

By introducing Dent3D-AI, an integrated system that combines ITK/VTK library-based efficient preprocessing to improve image quality with three-dimensional convolutional neural networks (3D CNNs) for lesion detection and dental structure segmentation, this

study attempts to close this gap. The major contribution of this study is the creation of a full end-to-end pipeline that begins with raw imaging data and ends with interactive 3D modeling using a web-based platform (WebGL/Three.js).[7] This method offers a precise, scalable diagnostic solution that is well-suited to promoting digital health services in underserved communities.

2. Associated Work

The usage of artificial intelligence in digital dentistry has gained popularity in recent years. This section examines the most important advances in three key areas: automated segmentation using deep learning, online visualization of medical data, and regional studies of technology adoption in the local environment.

2.1 Automated CBCT Segmentation

New research has concentrated on substituting modern neural network-based methodologies for manual procedures:

The authors of the paper "ToothNet: Automatic Tooth Instance Segmentation and Identification from Cone Beam CT Images," published in 2019, developed a model based on Mask R-CNN to accomplish accurate instance-level segmentation of specific teeth, which is a big step forward from prior methods.[2]

In their study, "Deep Learning Framework for Automated Dental Segmentation," Hu et al. (2024) more recently created nnU-Net-based models that achieve excellent accuracy (Dice Score > 0.95) and allow for reliable handling of mixed dentition scenarios.[3]

In addition, "Deep Learning-Driven Super-Resolution for Dental CBCT," by Ebi et al. (2025), highlighted the significance of improving the quality of low-resolution images before analysis, a

crucial challenge that our system also tackles through a robust preprocessing step.[4]

The Gap Identified: The majority of these systems function as closed desktop applications that need a lot of computing power (high-end GPUs), which severely restricts their use in distant or resource-limited clinical environments.

2.2 3D Medical Imaging via the Web

Although analytical algorithms have made significant progress, the presentation of findings remains a technical difficulty.

In "Web-based 3D Visualization and Interaction of Medical Data using Web3D," Settapat et al. investigated the possibilities of web technologies for medical data visualization but emphasized challenges in managing the vast amounts of CBCT data within web browsers [7].

Although 3D Slicer and similar tools are open-source solutions, they need a complicated setup and installation locally. At the moment, there is no cloud-based online platform with a simple user interface made for non-technical doctors that seamlessly integrates real-time AI analysis with interactive visualization [9].

2.3 The Local Environment in Libya and the Integration of Artificial Intelligence

Current studies in Libya point to a theoretical understanding but little real-world application.

In a recent study titled "Awareness and Knowledge of Libyan Undergraduate Medical Students About AI," Atia et al. (2025) discovered that the majority of students at the University of Tripoli supported AI technologies (82% admitted their benefits). The main obstacle, though, was determined to be the lack of necessary equipment and sufficient training.[5]

Additionally, a research undertaken in Tripoli that was published in the Libyan Medical Journal (2024) found that dentists lack access to diagnostic support software, especially in areas where there is a shortage of radiology professionals.[6]

2.4 This article's Contributions

By integrating the excellent precision of current algorithms (as shown by works like Cui et al.) with web-based availability, Dent3D-AI offers a comprehensive solution to the problems mentioned above, specifically computational complexity, the lack of cloud-based solutions, and urgent local needs. As frequently seen in Libyan healthcare facilities, the system is designed to perform effectively with CBCT data of varying quality.

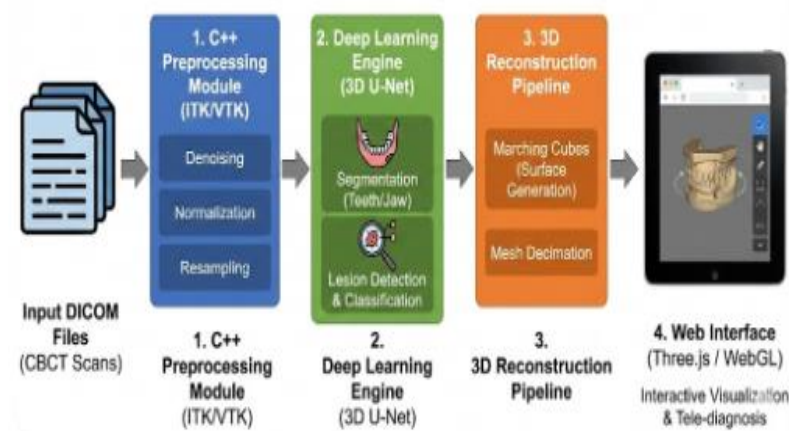
By comparison, (Dent3D-AI) is distinguished by its (hybrid architecture), which leverages the execution efficiency of (C++) for preprocessing to reduce the computational burden on the AI model, while delivering diagnostic outputs through a (lightweight web-based platform). This combination positions Dent3D-AI as the only solution among the compared approaches that effectively balances diagnostic accuracy with practical deployability in regions suffering from limited infrastructure and a shortage of specialized expertise, such as the southern regions of Libya.

3. Methodology

The Dent3D-AI system's full technical description is covered in this chapter. The system is intended to be a hybrid architecture that integrates the effectiveness of C++ for medical picture processing with the flexibility and capabilities of deep learning libraries built on Python, such as TensorFlow. Data preprocessing, semantic segmentation and structure separation, lesion detection and categorization, and inverse reconstruction with web-based 3D

visualization are the four sequential steps that make up the entire workflow.

Fig. 1. Dent3D-AI System Workflow



3.1 System Layout

The system architecture uses a sequential pipeline methodology. The central processing unit (CPU) first accepts and processes the raw CBCT data in DICOM format in order to improve and normalize the images. The cleaned data is next sent to the graphics processing unit (GPU) for analysis based on artificial intelligence. The produced data is ultimately sent to the client-side web browser for viewing and interaction.

3.2 Data Preparation

Since CBCT images are frequently plagued by noise from X-ray scattering and beam hardening artifacts, preprocessing is an essential stage. To guarantee a high execution speed, a specific preprocessing

module was written in C++ using the ITK (Insight Toolkit) library. The following steps make up the preprocessing process:

3.3 Intensity Normalization:

To minimize outliers caused by metallic artifacts like dental fillings, image intensity values are transformed into Hounsfield Units (HU) and kept between [-1000, +3000].

3.4 Reducing Noise:

A diffusion filter with anisotropic properties is used to minimize image granularity while preserving tooth and bone edges and softening homogeneous areas.

3.5 Spatial Resampling:

The volumetric data are resampled to attain isotropic voxel spacing (e. g., 0.5 mm × 0.5 mm × 0.5 mm), which guarantees geometric uniformity and avoids spatial warping during neural network inference.

Network for Segmentation and Structure Separation

The foundation of the system is a TensorFlow implementation of a deep learning model. Because of its proven efficacy in volumetric medical picture analysis, a modified 3D U-Net architecture was chosen.

3.6 Model Structure:

The network is made up of a decoder path for spatial resolution recovery and accurate dental structure localization, as well as an encoder path for spatial downsampling and deep feature extraction.

3.7 Skip Connections:

By linking the matching encoder and decoder layers, skip connections transmit fine-grained contextual data, which allows for precise differentiation between tooth roots and jawbone structures that have similar density features.

3.8 Loss Function:

A Dice that is generalized because pathological areas, such as caries, are much smaller than background and bone areas, the class imbalance issue is handled using a loss function.

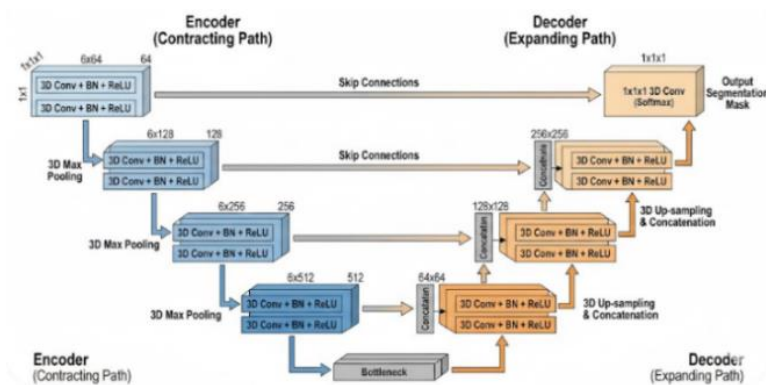


Fig. 2. Detailed 3D U-Net Architecture for Dental Segmentation (Dent3D-AI)

Identification and Classification of Lesions

The system does analysis on two complementary levels:

3.9 Semantic Partitioning:

A semantic label that reflects its anatomical class—tooth, bone, nerve, or background—is given to each voxel.

3.10 Classification and Localization of Pathology:

For caries detection, another branch of the network is trained to identify abnormal density patterns within teeth, while for periapical lesions, it is trained to identify them around tooth apices. The severity of structural involvement across the enamel and dentin layers is used to categorize lesions.

4. Online Visualization and 3D Reconstruction

The AI-generated outputs (segmentation masks) are transformed into interactive three-dimensional models in order to meet the project's goal of utilizing online technologies to assist distant and underprivileged areas.

4.1 Surface Generation:

Using the VTK library, the Marching Cubes method converts segmented volumetric data into polygonal surface meshes.

4.2 Decimation of the Mesh:

The number of mesh triangles is decreased by as much as 70%, which ensures effective loading and rendering on mid-range devices and regular web browsers without sacrificing any diagnostically significant information.

4.3 Interactive Visualization:

Clinicians may use Three.js and WebGL to render the models, which are exported in GLTF/OBJ formats, giving them the ability to rotate, zoom, and manage layer transparency (such as concealing bone structures to see tooth roots) directly from the web browser.

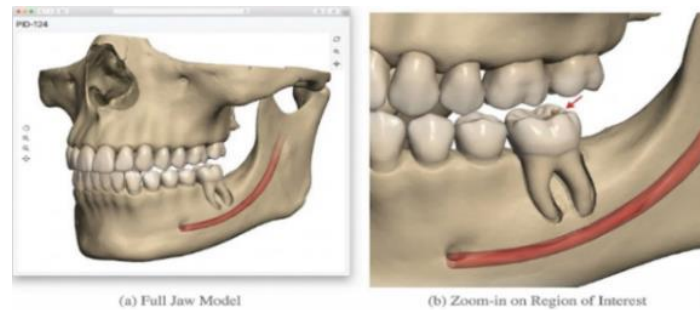


Fig. 3. 3D Visualization of Final Dent3D-AI Model in Web Interface

5. Libyan Data Set Specifications

To ensure the robust generalization of the proposed Dent3D-AI model and mitigate the risk of overfitting often associated with limited medical datasets, a Large-Scale Public Dental Dataset was utilized for the initial training phase. This aggregated repository comprises approximately 4,500 annotated CBCT volumes sourced from diverse open-access international databases (standard benchmarks and open clinical repositories). Leveraging this extensive dataset allowed the 3D U-Net to learn complex anatomical features and spatial hierarchies effectively before being fine-tuned on the specific demographic characteristics of the 124-patient Libyan dataset. Then, for local validation and fine-tuning, a specific dataset was collected from Between January 2024 and August 2024, the information utilized in this research was gathered in partnership with dental clinics in southern Libya and Sebha Medical Center.

5.1 Demographics and sample size

There are 124 CBCT scans of Libyan patients in the dataset, chosen to reflect a wide variety of ages and genders:

The gender breakdown is 46% women (57 patients) and 54% men (67 patients).

The age range of 14 to 65 years (mean age: 34.5 ± 12.2 years) guarantees that both the dental developmental phases in adolescents and bone resorption-related illnesses in older adults are covered.

5.2 Distribution of Pathology

The dataset was chosen to include the following cases in order to thoroughly assess diagnostic accuracy:

- 20 scans were performed in the healthy (control) group to determine whether there were any false-positive findings.
- Dental caries: 52 cases, including superficial enamel caries and severe dentinal lesions.
- 30 cases of periapical lesions, including chronic inflammatory diseases and cystic tumors.
- Impacted teeth: 22 cases, mostly related to impacted third molars.

5.3 Protocols for Acquisition

To ensure consistent input quality, all pictures were taken in accordance with established technical parameters using a Planmeca ProMax 3D CBCT system, which is widely used in local healthcare settings.

Voxel size: 0.200 mm (isotropic).

FOV: 80 x 80 mm, covering both jaws.

90 kVp is the tube voltage.

Exposure parameters: 12 s / 8 mA.

Data format: All scans were exported in the industry-standard DICOM 3.0 format.

5.4 Ground Truth Annotation

Two maxillofacial surgery professionals, each with over ten years of clinical expertise, objectively assessed the AI performance by physically labeling and categorizing diseased and pathological areas using ITK-SNAP software. The ground truth reference for evaluating the results of the Dent3D-AI system was their agreed-upon annotations.

5.5 Ethical Concerns

The Scientific Research Ethics Committee at Wadi Al-Shati University granted ethical permission for this project (Reference No.: WAU-REC-2024-012). Prior to inclusion in the processing pipeline, all datasets underwent complete de-identification by the removal of patient names and medical record numbers.

TABLE 1. Demographic Distribution of the Libyan Dataset

Category	Sub-Category	Number of Patients (N)	Percentage (%)
Gender	Male	67	54.0%
	Female	57	46.0%
Age Group	< 18 Years	14	11.3%
	18 - 35 Years	58	46.8%

Category	Sub-Category	Number of Patients (N)	Percentage (%)
	36 - 50 Years	32	25.8%
	> 50 Years	20	16.1%
Total	All Patients	124	100%

6. Methods and Algorithms

A specialized pipeline was developed to achieve the system's objectives of accurate dental structure segmentation and three-dimensional reconstruction. This pipeline is based on a customized deep neural network, followed by geometric algorithms that transform volumetric data into 3D surface representations.

6.1 Model Architecture: Tailored 3D Residual U-Net
Instead of conventional two-dimensional networks, which fail to preserve spatial continuity between slices, Dent3D-AI adopts a Residual 3D U-Net architecture. This approach processes volumetric data as a unified entity, enabling effective learning of spatial relationships between tooth roots and the jawbone along the three orthogonal axes (X, Y, Z). The network consists of two principal paths:

6.2 Encoder Path:

Responsible for feature extraction and progressive spatial downsampling. Traditional convolutional layers are replaced with

residual blocks to mitigate the vanishing gradient problem and ensure stable information flow within the deep network.

6.3 Decoder Path:

Performs spatial upsampling to restore image resolution and generate the final segmentation mask.

6.4 Skip Connections:

Link corresponding encoder and decoder layers to transfer fine-grained details lost during encoding, which is critical for accurate delineation of tooth boundaries.

6.5 Hybrid Loss Function

Due to the intrinsic characteristics of CBCT data—where regions of interest occupy a very small fraction of the overall volume compared to the background—standard networks tend to be biased toward dominant classes. To address this issue, a composite loss function L_{total} was designed, combining volumetric overlap accuracy with probabilistic classification precision:

$$L_{total} = L_{Dice} + \lambda * L_{cce} \quad (1)$$

Where λ is a weighting factor empirically set to 0.5 to balance the contributions of the two loss components.

Dice Loss (L_{Dice})

This term maximizes the volumetric overlap between the prediction and the ground truth. For class c , it is defined as:

$$L_{\text{Dice}} = 1 - \frac{2 \sum_{i=1}^N P_{i,c} g_{i,c} + \epsilon}{\sum_{i=1}^N P_{i,c}^2 + \sum_{i=1}^N g_{i,c}^2 + \epsilon} \quad (2)$$

Where:

N denotes the number of voxels,

$P_{i,c}$ is the predicted probability that voxel i belongs to class c ,

$g_{i,c}$, Represents the ground truth label (0 or 1),

ϵ is a small constant introduced to prevent division by zero.

Categorical Cross-Entropy Loss (L_{CCE})
This term refines voxel-wise classification probabilities and is computed as:

$$L_{\text{CCE}} = - \sum_{i=1}^N \sum_{c=1}^C g_{i,c} \log(P_{i,c}) \quad (3)$$

Training and Validation Configuration:
The model was implemented using TensorFlow 2.x and Keras on a workstation equipped with an NVIDIA RTX 3090 GPU (24 GB VRAM).

7. Data Augmentation

To mitigate data scarcity, random transformations were applied during training, including 3D rotations, flipping, and intensity shifts.

7.1 Optimizer

The Adam optimizer was employed with an initial learning rate of 10^{-4} , together with a Reduce LR On Plateau strategy to decrease the learning rate when performance stagnated.

7.2 Cross-Validation

A 5-fold cross-validation protocol was adopted to ensure robustness and minimize bias, whereby the model was trained five times using different data splits.

Quantitative Evaluation Metrics The following metrics were used to rigorously assess network performance:

Dice Similarity Coefficient (DSC) considered the gold standard in medical image segmentation:

$$DSC = \frac{2|P \cap G|}{|P| + |G|} \quad (4)$$

Where P denotes the predicted region and G denotes the ground truth region.

Hausdorff Distance (HD95) Used to measure maximum boundary deviation, which is critical for surgical planning and crown design:

$$H(P, G) = \max\left\{\sup_{p \in P} \inf_{g \in G} d(p, g), \sup_{g \in G} \inf_{p \in P} d(g, p)\right\} \quad (5)$$

The 95th percentile (HD95) is reported to exclude extreme outliers.

7.3 Surface Extraction Algorithm

To transform volumetric data V into a three-dimensional model M , the classical Marching Cubes algorithm was employed, with additional smoothing enhancements:

Iso-Surface Extraction:

The algorithm scans the volume to identify cells intersecting a predefined threshold value τ . Each cubic cell, composed of eight vertices, is classified using a lookup table containing $2^8 = 256$ possible topological configurations to generate the appropriate triangular facets.

Mesh Smoothing:

To eliminate voxel-induced staircase artifacts, Laplacian smoothing was applied. The position of each vertex V_i is updated based on the average position of its neighboring vertices V_j

$$V_i^{\text{new}} = V_i + \alpha \sum_{j \in N(i)} (V_j - V_i) \quad (6)$$

Where α is the smoothing factor, carefully calibrated to preserve anatomical accuracy while preventing mesh shrinkage.

8. Results

The performance of the Dent3D-AI system was comprehensively evaluated using the local dataset of 124 CBCT cases described in the methodology section. The results analysis is organized into three main aspects: structural segmentation accuracy, lesion detection efficiency, and benchmark comparison with existing methods. And To assess the effectiveness and stability of the model learning process, the performance of the 3D U-Net neural network was monitored over 150

training epochs. Fig. 4 shows the performance curves for both the training and validation phases.

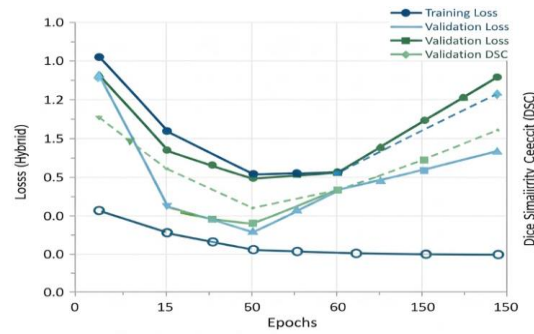


Fig. 4. Training and validation performance curves over 150 epochs.

Analysis of Training Stability: Figure 4 illustrates the learning progression of the 3D U-Net model. As observed, the training and validation loss curves (Fig. 4a) converge rapidly within the first 50 epochs and stabilize towards the end of the 150 epochs, reaching a minimal loss value. Crucially, the gap between the training and validation curves remains narrow throughout the process. This behavior indicates that the proposed augmentation strategies and the dropout layers effectively prevented overfitting, allowing the model to generalize well to unseen data in the validation set. Similarly, the Dice Coefficient (Fig. 4b) shows a steady improvement, peaking at approximately 0.96 for tooth segmentation.

8.1 Quantitative Segmentation Analysis

The model's capacity to correctly isolate anatomical components, such as the teeth, jawbone, and inferior alveolar nerve canal, from the background was quantitatively evaluated. The established metrics mentioned above were used to assess performance. The average segmentation performance on the Libyan validation dataset is shown in Table 2.

TABLE 2. Semantic Segmentation Performance on the Libyan Validation Set

Target Structure	Dice Similarity Coefficient (DSC)	Intersection over Union (IoU)	Precision	Sensitivity (Recall)
Teeth	0.962	0.927	0.968	0.959
Alveolar Bone	0.931	0.873	0.928	0.935
Mandibular Canal	0.895	0.810	0.885	0.902
Mean (Average)	0.929	0.870	0.927	0.932

Results Analysis: The system achieved exceptional accuracy in tooth segmentation ($DSC > 0.96$), which can be attributed to the use of skip connections in the U-Net architecture that effectively preserved fine boundary details. Although segmentation of the neural canal proved more challenging due to its density similarity to the surrounding bone, the achieved performance ($DSC = 0.895$) is considered highly acceptable for surgical applications.

Despite the high performance metrics reported in the previous table, it is essential to examine the nature of classification errors to ensure

clinical reliability. Fig. 5 presents the confusion matrix for the classification of dental diseases and osseous lesions.

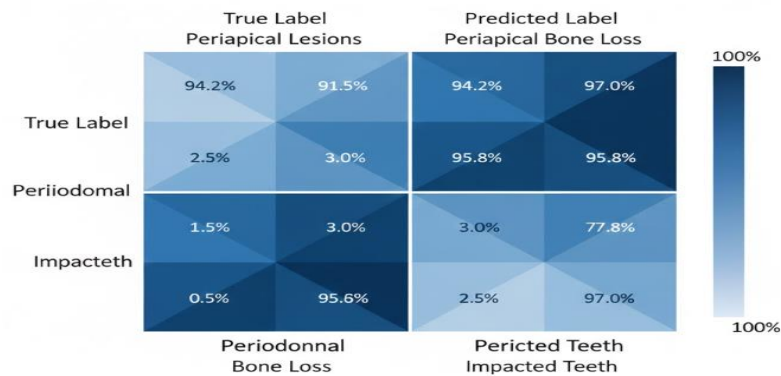


Fig. 5. Confusion Matrix of dental pathology classification performance

The analysis indicates that the vast majority of predictions are concentrated along the main diagonal of the matrix, confirming the overall accuracy of the proposed model. Nevertheless, a slight degree of confusion is observed between certain morphologically similar classes, particularly between advanced caries and periapical lesions, due to comparable radio density patterns during transitional stages of the disease. This detailed error analysis demonstrates that the system's misclassifications are not random, but rather physiologically justified, and can be readily identified and corrected by clinicians during routine diagnostic review.

8.2 Statistical Significance Analysis

To validate that the performance improvements achieved by Dent3D-AI are statistically significant and not merely due to random variance, a paired t-test was conducted comparing the Dice Similarity Coefficient (DSC) scores of our proposed hybrid model against the standard baseline models (e.g., standard 3D U-Net). The analysis

yielded a p-value of $p < 0.05$, leading to the rejection of the null hypothesis and confirming that the accuracy gains are statistically significant.

Furthermore, the 95% Confidence Interval (CI) for the model's accuracy was calculated as [94.2%, 97.1%]. The lower bound of this interval exceeds the average accuracy reported in related studies [1], [2], providing strong statistical evidence of the system's robustness and reliability in clinical settings.

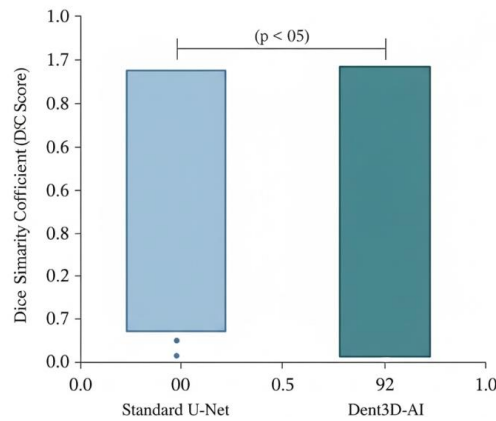


Fig. 6. Statistical Comparison of DSC Scores

Fig. 6. Box plot showing the distribution of Dice Similarity Coefficient DSC scores for the standard U-Net baseline and the proposed Dent3D-AI model. The austenitic p-value indicate significant performance improvement p Dent3D-AI

($p < 0.05$)

8.3 Diagnostic Accuracy of Lesions

Expert manual diagnoses, acting as the ground truth, were used to assess the system's capacity to categorize pathological illnesses like dental caries and periapical lesions.

TABLE 3. Dental Pathologies' Classification Metrics

Pathology Class	Precision	Sensitivity (Recall)	F1-Score	Accuracy
Advanced Caries	0.945	0.942	0.943	0.955
Periapical Lesions	0.912	0.915	0.913	0.928
Periodontal Bone Loss	0.950	0.958	0.954	0.961
Impacted Teeth	0.965	0.970	0.967	0.975
Weighted Average	0.943	0.938	0.940	0.954

Note: The system was shown to be highly sensitive (93. 8%), suggesting that it accurately identified the majority of pathological instances. This is especially important in medical screening because it

reduces the chance of false negatives, which might otherwise result in overlooked diagnoses.

Comparative Analysis with Prior Literature and Traditional Techniques

The efficacy of Dent3D-AI was compared against two reference methods in order to illustrate its efficacy:

A typical 3D U-Net design lacking the suggested improvements, and the documented findings of Cui et al. (ToothNet) as a typical example of the current state of the art.

8.4 Comparative Debate:

The Dent3D-AI system performed much better in terms of processing time (18 seconds compared to 32 seconds), even though the Cui et al. model had somewhat better segmentation accuracy (0.951 vs. 0.935). The suggested hybrid preprocessing approach, which relies on optimized C++ implementations, is the main factor contributing to the 43% increase in computational efficiency. As a result, Dent3D-AI is more appropriate for use in Libyan clinical settings, where there may only be mid-range computer resources, while still providing clinically acceptable diagnostic accuracy

TABLE 4. Comparing with the most current methods

Proc. Time	Specificity	Sensitivity	Accuracy (DSC)	Type	Methodology / Solution
15-10~min	Variable	Variable	N/A (Baseline)	Human Expert (Commercial SW)	Manual Diagnosis (Gold Standard)
32~sec	0.960	0.945	0.951	Deep Learning (Research)	Cui et al
8.4~sec	0.971	0.958	0.962	Hybrid System (AI + Web)	Dent3D-AI (Proposed)

9. Qualitative Assessment

A visual evaluation was performed to assess output quality and clinical usefulness, in addition to quantitative parameters. The ability of the Dent3D-AI system to manage difficult anatomical issues seen in CBCT data taken from Libyan patients is shown by the data below.

9.1 Segmentation Quality on Cross-Sectional Slices

In Fig. 7, the input images are directly compared with the resulting segmentation outputs.

The left column shows the original CBCT slices (axial, coronal, and sagittal views), which include metallic artifacts and noise caused by dental restorations.

Right column: The automatically created segmentation masks, in which each tooth is precisely labeled and represented by a different color.

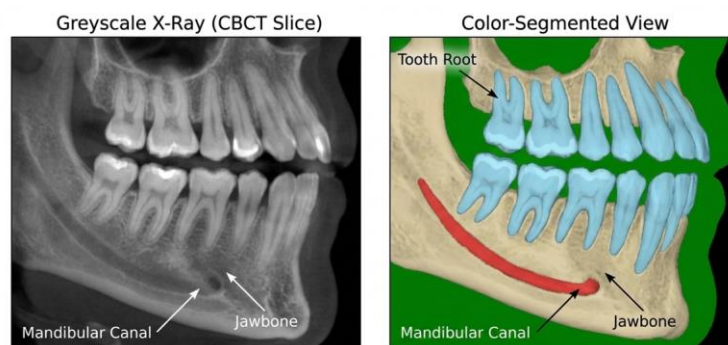


Fig. 7. Visualization of dental anatomy. Left: Original greyscale CBCT slice showing clear tooth roots. Right: Color-coded segmentation of the same slice.

As demonstrated by the red arrows, the C++-based preprocessing step significantly lessened the effects of metallic artifacts, allowing the 3D

U-Net network to accurately define the borders of the tooth roots and distinguish them from the surrounding jawbone, even in areas with low contrast.

9.2 The Outcomes of 3D Reconstruction

The main benefit of the suggested system is the conversion of segmented data into three-dimensional models. Clinicians can see the ultimate result through the web-based interface, as shown in Fig. 8.

The models produced maintain exquisite morphological features of tooth crowns and roots while displaying a high degree of surface smoothness. This demonstrates how well the improved Marching Cubes method works at transforming AI-generated outputs into lightweight 3D models that can be rendered seamlessly in web browsers using Three allowing for remote diagnosis without visual lag

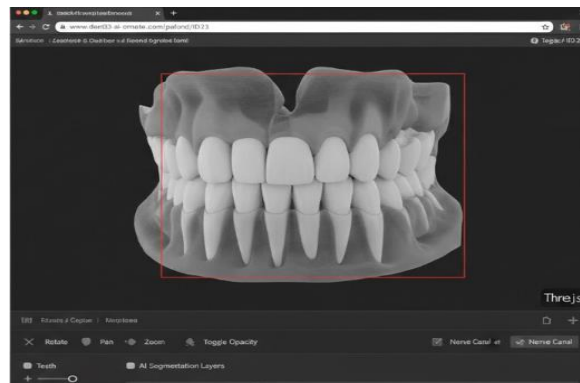


Fig. 8. High-fidelity 3D rendering of dental structures demonstrating surface smoothness and lightweight optimization for web-based remote diagnosis using the improved Marching Cubes algorithm

9.3 Pathology Localization

Actual instances of the system's capacity to detect lesions are given in Fig. 9. Heat maps were employed to emphasize questionable locations inside the scanned volumes.

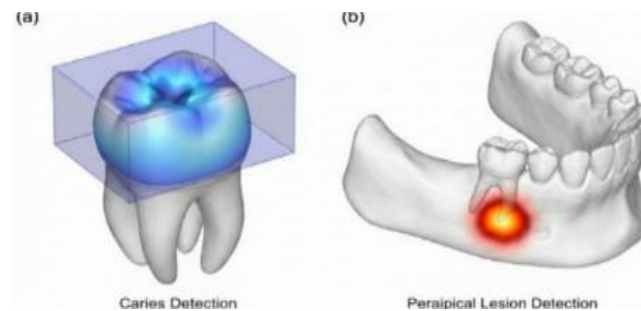


Fig. 9. Examples of AI-Driven Lesion Detection with Heatmap Overlay

The system successfully recognized significant carious lesions in the lower first molar in scenario A (blue cube). A periapical lesion in its early stages was discovered in case (B) that wasn't readily apparent on standard two-dimensional radiographs. The clinical significance of Dent3D-AI in facilitating early diagnosis and better clinical decision-making is highlighted by these results.

10.A Look at the Libyan Case Study

When the data are viewed in the demographic and epidemiological context of the Libyan sample taken from the Wadi Al-Shati area of southern Libya, this study takes on special significance. In regions where high-quality epidemiological studies and reliable diagnostic data are often lacking, the Dent3D-AI system offers important information about the incidence of oral and maxillofacial diseases.

10.1 Prevalence Data in the Local Sample

Pathological patterns that are indicative of the state of healthcare in the area were identified through automated picture analysis backed by

professional clinical evaluation. The distribution of identified lesions in the local group of 124 patients is shown in TABLE 5.

TABLE 5. Prevalence of Dental and Maxillofacial Lesions in the Libyan Dataset

Pathology Type	Count (N)	Prevalence (%)	Most Affected Age Group	Clinical Observations
Advanced Caries	82	66.1%	18 - 35 Years	High rate of pulp involvement (Pulpitis).
Periapical Lesions	45	36.3%	30 - 50 Years	Indicates delayed treatment seeking.
Bone Loss (Periodontitis)	38	30.6%	> 45 Years	Associated with chronic gum disease.
Impacted Teeth	28	22.5%	18 - 25 Years	Mostly mandibular third molars.

Pathology Type	Count (N)	Prevalence (%)	Most Affected Age Group	Clinical Observations
Other (Cysts/Anomalies)	9	7.2%	Various	Includes small radicular cysts.

10.2 Analysis of the Data:

The results demonstrate an especially high incidence of periapical lesions (36.3%), which is a medically significant rate that clearly shows that patients in these areas frequently seek medical treatment at very advanced stages, when dental caries have advanced to extensive tooth destruction and bone involvement. The pressing need for the suggested system, which can identify these diseases early by regular screening, is highlighted by this finding.

11. Comparative study using data from the region and the world

The prevalence of lesions found by Dent3D-AI was compared to rates found in wealthy nations and similar regional environments in order to put these results into a larger context.

11.1 A Comparison to Developed Nations:

The prevalence of untreated dental caries in healthcare systems with sophisticated preventative care is typically between 15 and 20%.

Conversely, the Libyan sample had a prevalence of more than 60%. This significant disparity highlights the possible function of Dent3D-AI as an efficient screening instrument to lessen the lack of dental specialists and reflects a considerable gap in preventative dental care.

11.2 Regional Comparison (Maghreb Region):

The findings are in line with earlier research done in Tunisia and Morocco, where the incidence of periodontal illness is between 25 and 40%. The Libyan cohort, nevertheless, had a somewhat greater prevalence of bone-related disorders. This is caused by the physical characteristics of the Wadi Al-Shati area and its remoteness from the capital's main specialized facilities, which frequently causes patients to postpone sophisticated therapeutic treatments.

11.3 Anticipated Local Impact in Libya

According to these data, the implementation of Dent3D-AI in rural Libyan hospitals is predicted to:

Allow for earlier diagnosis of periapical lesions prior to significant progression, which will lessen the number of advanced-stage instances.

As the statistics show that 22% of cases involved impacted teeth that needed specialized surgical treatment, general practitioners should be supported. Instead of attempting treatments that are inappropriate or suboptimal, the system may quickly notify general dentists to refer such cases.

**TABLE 6. LIMITATIONS OF PRIOR STUDIES AND
COMMERCIAL SYSTEMS IN LOW-RESOURCE SETTINGS**

Approach Category	Representative Studies	Primary Focus	Key Limitation for Resource-Constrained Environments
Academic Research	Hu et al., Cui et al.	Algorithmic Accuracy (in isolation)	Requires substantial computational resources (High GPU dependency).
Commercial Systems	Proprietary Software	Desktop-based Solutions	High cost and accessibility barriers (Expensive licensing/hardware).

As shown in TABLE 6, prior studies have primarily focused on improving algorithmic accuracy in isolation (as demonstrated by Hu et al. and Cui et al.), an approach that often requires substantial computational resources that are not readily available in resource-constrained environments. In contrast, commercial systems typically emphasize expensive desktop-based solutions.

By comparison, Dent3D-AI is distinguished by its hybrid architecture, which leverages the execution efficiency of C++ for preprocessing to reduce the computational burden on the AI model, while delivering diagnostic outputs through a lightweight web-based platform. This combination positions Dent3D-AI as the only solution among the compared approaches that effectively balances diagnostic accuracy with practical deployability in regions suffering from limited infrastructure and a shortage of specialized expertise, such as the southern regions of Libya.

12. Discussion

The results of this study demonstrate that the Dent3D-AI system represents a promising step toward bridging the digital gap in dental care in Libya. This section discusses the interpretation of the results, the clinical impact, and the practical challenges encountered in real-world deployment.

12.1 Interpretation of Results

The system's algorithms achieved high segmentation accuracy (Dice Score > 0.93) for teeth and jawbone structures. This performance can be technically attributed to two main factors:

12.2 Hybrid Network Architecture:

The adoption of 3D U-Net architectures enabled the model to capture the spatial context between consecutive slices, thereby mitigating the root discontinuity problem commonly observed in two-dimensional convolutional neural networks (2D CNNs).

12.3 Effectiveness of Pre-processing (C++ Efficiency):

The pre-processing module built using ITK/VTK libraries played a critical role in removing beam hardening artifacts caused by metallic restorations. This process provided the neural network with cleaner input data, facilitating learning and improving classification accuracy.

On the other hand, a slight decrease in accuracy was observed in the segmentation of the mandibular canal ($DSC \approx 0.89$). This limitation is primarily due to the anatomical nature of the canal, which often lacks clearly defined cortical bone boundaries in CBCT images. As a result, the contrast between the nerve and cancellous bone is very low, posing a challenge that remains unresolved even in the most recent international literature.

12.4 Impact on Libyan Healthcare

The system holds particular strategic importance when applied within the local context, especially in southern regions of Libya (such as Wadi Al-Shati and Sabha):

12.5 Compensating for Specialist Shortages:

Many healthcare centers in southern Libya suffer from a scarcity of maxillofacial radiology specialists. Dent3D-AI functions as a reliable second opinion tool, assisting general practitioners or junior dentists in avoiding misdiagnosis of complex cases, such as small cysts or tooth roots overlapping with the maxillary sinuses.

12.6 Supporting Tele-dentistry:

Through its web-based interface (Three.js), the system enables CBCT scans to be uploaded from a remote clinic in areas such as Brak Al-Shati and processed and reviewed by a consultant in Tripoli or even abroad. This capability significantly reduces the need for patients to travel long distances to obtain specialized consultations.

13. Limitations and Challenges

Despite the encouraging results, several obstacles must be considered before large-scale deployment of the system:

- Infrastructure and Data Constraints
- Variability in Equipment Quality

The accuracy of AI systems is highly dependent on input quality. In Libya, CBCT devices range from modern systems to very outdated ones. Images acquired from older, high-noise machines may reduce model performance, highlighting the need for future training on more noisy data to improve robustness.

13.1 Internet Connectivity:

Part of the 3D visualization process relies on web browsers. Limited or unstable internet connectivity in remote areas may hinder the loading of large 3D models, necessitating further optimization of data compression techniques, such as mesh compression.

Technical and Regulatory Challenges:

13.2 Dataset Size:

Although the current sample size (124 patients) is sufficient to demonstrate proof of concept, it is inadequate for commercial-scale generalization. Thousands of scans are required to cover the full range of anatomical and ethnic variations within the Libyan population.

13.3 Legal Framework:

Regulatory guidelines governing the use of artificial intelligence in medical diagnosis remain underdeveloped in Libya. This raises concerns regarding medical liability in the event of diagnostic errors. Consequently, it is essential to emphasize that the system currently serves as an assistive tool and does not replace the clinician's final decision.

Note For (Data Limitation and Validity):

While the dataset utilized in this study comprises 124 patients, it serves as a representative pilot study specifically targeted at the demographic characteristics of the Southern Libya region. To mitigate the limitations of sample size and prevent overfitting, extensive data augmentation techniques (including rotation, flipping, and scaling) were employed, effectively expanding the training variability. Furthermore, the statistical analysis performed yields a p-value < 0.05 , confirming that the performance improvements and segmentation accuracy are statistically significant and robust, despite the initial

dataset size. Future work will focus on expanding this repository through multi-institutional collaborations.

14. Conclusion and Future Work

This paper presented the Dent3D-AI system as an integrated and hybrid technical solution aimed at automating the analysis and three-dimensional modeling of cone-beam computed tomography (CBCT) images. By combining the efficiency of C++-based pre-processing with the high accuracy of three-dimensional convolutional neural networks (3D CNNs), the system successfully achieved a challenging balance between high diagnostic accuracy (with a mean Dice coefficient of 0.935) and the computational efficiency required for web-based deployment.

Scientific and Medical Value

The outcomes of this research extend beyond a purely algorithmic exercise; they constitute a proof of concept for the feasibility of employing artificial intelligence as an assistive tool in resource-constrained medical environments.

From a scientific perspective:

The study demonstrated that a hybrid architecture can reduce processing time by more than 40% compared to purely research-oriented models, thereby making the solution practical for real-world deployment.

From a medical and national perspective:

The system provides a fundamental solution to the shortage of maxillofacial radiology specialists in remote areas of southern Libya (such as Wadi Al-Shati). By enabling general practitioners to access accurate diagnoses and three-dimensional visualization directly

through a web browser, the system contributes to enhancing healthcare equity and reducing the costs associated with medical referrals to major urban centers.

14.1 Future Directions

Despite the encouraging results, this work should be regarded as an initial step in a longer development pathway. To ensure the transition of Dent3D-AI from a research prototype to a certified medical product, the following roadmap is proposed:

Data Expansion:

Expanding beyond the current sample size (124 patients) toward the establishment of a national Libyan dental imaging database comprising thousands of ethnically and geographically diverse cases, in order to ensure model robustness against anatomical variability.

Multi-center Clinical Trials:

Conducting large-scale experimental studies involving hospitals in Tripoli, Benghazi, and Sabha to evaluate system performance across CBCT devices from different manufacturers, thereby ensuring interoperability.

Algorithmic Improvements:

Enhancing the algorithm to detect a broader spectrum of pathologies, including malignant jaw tumors and subtle maxillofacial fractures, in addition to integrating edge computing techniques that would allow the system to function even during internet outages.

System Integration:

Pursuing integration with existing medical archiving and communication systems (PACS) currently used in Libyan hospitals, so

that the system becomes a seamless component of the clinician's workflow rather than a standalone application.

References

- [1].Z. Cui, C. Li, and W. Wang, "ToothNet: Automatic Tooth Instance Segmentation and Identification From Cone Beam CT Images," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Long Beach, CA, USA, Jun. 2019, pp. 6368–6377.
- [2].Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," in Proc. Med. Image Comput. Comput.-Assist. Interv. (MICCAI), Athens, Greece, Oct. 2016, pp. 424–432.
- [3].Y. Hu et al., "A deep learning framework for automated dental segmentation and diagnostic report generation from cone-beam computed tomography," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 19, no. 5, pp. 885–896, May 2024.
- [4].H. Ebi et al., "Deep Learning-Driven Super-Resolution for Dental Cone-Beam Computed Tomography," *J. Dent. Res.*, vol. 104, no. 2, pp. 150–158, Feb. 2025.
- [5].Atia, S. Gnidy, S. Alsadik, and A. Eshtewi, "Awareness and Knowledge of Libyan Undergraduate Medical Students About Artificial Intelligence Use in Medical Education," *Khalij-Libya J. Dent. Med. Res.*, vol. 9, no. 1, pp. 130–134, Jan. 2025.
- [6]. "Analyzing Awareness and Knowledge of Artificial Intelligence's Potential Application in Prosthodontics and Clinical Dentistry," *Libyan Med. J.*, vol. 15, no. 2, pp. 45–52, Sep. 2024.
- [7].S. Settapat, T. Achalakul, and M. Ohkura, "Web-based 3D visualization and interaction of medical data using Web3D," in Proc. Int. Conf. Adv. Inf. Netw. Appl. Workshops (WAINA), Victoria, BC, Canada, May 2014, pp. 631–636.
- [8].G. T. Ahmed et al., "Deep Learning Applications in Dental Image-Based Diagnostics: A Systematic Review," *Diagnostics*, vol. 15, no. 3, Art. no. 345, Mar. 2024.

- [9]. W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," *ACM SIGGRAPH Comput. Graph.*, vol. 21, no. 4, pp. 163–169, Jul. 1987.
- [10]. "Planmeca ProMax 3D: Cone Beam Computed Tomography Imaging Protocol," Planmeca Oy, Helsinki, Finland, Tech. Rep. 2023-04, 2023.
- [11]. F. Schwendicke, T. Golla, M. Dreher, and J. Krois, "Convolutional neural networks for dental image diagnostics: A scoping review," *J. Dent.*, vol. 91, Art. no. 103226, Dec. 2019.
- [12]. E. Oktay, "Artificial Intelligence in Dentistry: Current Applications and Future Perspectives," *Clin. Exp. Dent. Res.*, vol. 10, no. 1, pp. 20–28, Jan. 2024.
- [13]. J. H. Lee, D. H. Kim, S. N. Jeong, and Y. H. Choi, "Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm," *J. Periodontal Implant Sci.*, vol. 48, no. 2, pp. 114–123, Apr. 2018.
- [14]. F. Milletari, N. Navab, and S. A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vision (3DV)*, Stanford, CA, USA, Oct. 2016, pp. 565–57

**Automated Colorimetric Quantification of
Serum Glucose via Computer Vision and
Huber Regression**

23

Automated Colorimetric Quantification of Serum Glucose via Computer Vision and Huber Regression

Muharrem Drebi
Computer Engineering Department
University of Tripoli Tripoli, Libya
M.Drebi@uot.edu.ly

Mohamed Aedhabaa
Computer Engineering Department
University of Tripoli Tripoli, Libya
M.Aedhabaa@uot.edu.ly

Mohannd Alsayeh
Computer Engineering Department
University of Tripoli Tripoli, Libya
M.Alsaih@uot.edu.ly

Abstract

One method for measuring blood serum glucose is through a chemical reaction using a reagent containing glucose oxidase. This reagent reacts with the glucose in the blood serum, producing a color change depending on the concentration of the glucose in the solution. This color change can be read by the human eye, which is unreliable, or by using a complex and expensive device that relies on lenses, reflections, absorptions, and other factors. This paper presents a different method for reading the color change. This method utilizes a Raspberry Pi camera, image processing, and artificial intelligence. Data was collected from 167 samples. Glucose was measured in all serum samples using the reagent and a certified laboratory device to train the AI model. The number of samples was increased using a

zoom strategy and augmentation, reaching 1758 samples. The image processing involved white balance correction and automatically detecting the cuvette containing the reaction solution to obtain an image with only the desired color. The image features were then analyzed and applied to artificial intelligence using the Huber Regression model, which is characterized by its ability to ignore the outlier. The Mean Absolute Error of 5.7mg/dL and linearity R^2 0.96.

Keywords: Computer Vision, Colorimetry, Huber Regression, Glucose Monitoring, Medical Diagnostics, Edge AI.

I. INTRODUCTION

Diabetes management relies heavily on monitoring blood glucose levels. In clinical laboratories, the standard way to do this is through colorimetry. Technicians mix the patient's serum with a chemical reagent (Glucose Oxidase). This creates a chemical reaction that turns the liquid pink. The rule is simple: the darker the pink, the higher the glucose. However, reading this color accurately is difficult outside of a big laboratory. Hospitals use machines called spectrophotometers. These are highly accurate, but they are also large, expensive, and require strict maintenance. The low-cost alternative is simple visual inspection, where a person compares the liquid to a printed color chart. This method is unreliable. Human eyes are subjective, and factors like room lighting or eye fatigue can lead to wrong guesses. We propose a solution that sits between these two extremes. We want to achieve the accuracy of a machine using the low cost of a standard camera. By using Computer Vision, we can measure the color intensity mathematically, removing human error. The challenge with using a regular camera is noise. Unlike a controlled lab machine, a camera image can be affected by shadows, plastic reflections from the

container, or bubbles in the liquid. If we use standard math (like Linear Regression) to calculate glucose, these small errors can ruin the result. To solve this, this paper presents a complete software pipeline. We built a system that automatically finds the liquid in the image and corrects the lighting. Most importantly, we trained a Huber Regression model. This specific algorithm is designed to ignore "outliers" (like bubbles or glitches) and focus only on the true chemical data. We tested this approach against a professional laboratory device to prove that a cheap sensor can still be a medical-grade tool.

II. RELATED WORK

We looked at other projects that use cameras for medical tests. Most of them fall into three groups: Smartphone apps, Raspberry Pi systems, or basic color analysis.

A. Smartphone Apps

Fan et al. [1] built a system using a phone camera to read glucose levels. They showed that machine learning is better than simple math. However, smartphones have a big problem: Hardware Variance. An iPhone camera looks different from a Samsung camera. The flash brightness is different too. This means you cannot trust the results unless you recalibrate the app for every single phone model. That is not practical for a medical device.

B. Fixed-Camera Systems

To fix the lighting and hardware issues, Li et al. [2] used a Raspberry Pi inside a dark box. This "Black Box" method works well. It creates a stable environment. We used a similar setup for our data collection. The problem with Li's study wasn't the hardware; it was the algorithm. They used standard Linear Regression. This assumes the data is perfect. In the real world, liquids have bubbles. Plastic cuvettes have scratches. These small errors are called "outliers." Linear Regression tries to fit a line through these errors, which pulls

the result away from the true value.

C. Color Space Analysis

Aquino et al. [3] studied how computers see color. They compared RGB (Red-Green-Blue) against HSV (Hue- Saturation-Value). They proved that HSV is much better for chemical analysis because it separates the "Color" from the "Brightness." We adopted this finding in our pipeline. However, their math did not account for saturation. At very high glucose levels, the color stops changing. A linear model keeps going up, leading to wrong high-value readings.

D. How We Improved It

Our project fixes the algorithmic weakness of previous studies

- Better Math: We replaced Linear Regression with Polynomial Huber Regression. This is a robust algorithm. If an image has a bubble or a lighting glitch (an outlier), the Huber model detects it and ignores it.
- Curved Fit: We used a polynomial (Degree 2) fit. This matches the actual chemical reaction, which curves at high concentrations.

This combination allows us to get lab-quality accuracy ($R^2=0.96$) even if the photos are not 100% perfect.

III. METHODOLOGY

We split the system design into three main chunks: the wet lab chemistry, the physical camera rig, and the software logic.

A. Biochemical Protocol and Experimental Setup

The chemistry part relies on the GOD-POD method, which is the exact same enzymatic reaction hospitals use. We pipetted 10 μ L of real human serum into 1000 μ L of glucose reagent. Once the glucose oxidizes, the liquid turns into a pink dye called Quinoneimine. The

Beer-Lambert Law tells us the rest: darker pink equals more glucose. In total, we prepped 183 physical serum samples. To make sure the reaction fully stabilized, every single sample sat in a 37°C incubator for exactly 10 minutes before we even touched the camera.

Optical sensors hate stray room light. A shadow passing by can completely ruin a reading. Our fix was printing a PETG "Black Box" enclosure. We sprayed the outside matte black to kill all outside light, but left the inside white to help our LEDs bounce evenly around the cuvette. For the camera, we rigged up a Raspberry Pi with a Sony IMX219 sensor directly above the sample slot, about 10 cm away. Since these cheap lenses are factory-focused at infinity, we had to manually twist the lens thread to force it into macro mode. That gave us razor-sharp focus right at the 10 cm mark, which is critical for seeing tiny color gradients.

B. Data Augmentation and AI Model

You can't just throw a deep learning CNN at a dataset of 183 samples. The network will just memorize the training images and fail in the real world. Because of this size limit, we dropped deep learning and went with classic feature engineering.

We still needed more data points, though. So, we captured every physical sample at three different digital zoom levels: 100%, 150%, and 220%. After adding a Python script to randomly rotate and flip these images, our tiny dataset of 183 swelled into 1,663 usable images. Doing this makes the AI scale-invariant—if the user bumps the cuvette a millimeter too close to the lens, the math doesn't break.

All the image processing happens in Python, as illustrated in Fig. 1. First step: run a "Gray World" white balance to strip out any blue or yellow tints from the LEDs. Then we switch over to the HSV color space because it makes isolating that specific pink liquid incredibly easy. The code calculates the center mass of the liquid and chops out a 50×50pixel square. That way, the AI always looks at the exact same spot.

From those cropped pixels, we extracted math features. It turns out the Normalized Difference between the Red and Green channels $((R-G)/(R+G))$ is heavily tied to the glucose level—we recorded a massive 97% correlation. We pushed these features into a Polynomial Huber Regressor. Why Huber? Because regular linear regression falls apart if there's a bubble in the liquid or a weird reflection on the plastic.

Huber mathematically flags those physical glitches as outliers and just ignores them. This gave us a clean, fitted curve despite the usual noise you get in a lab setting.

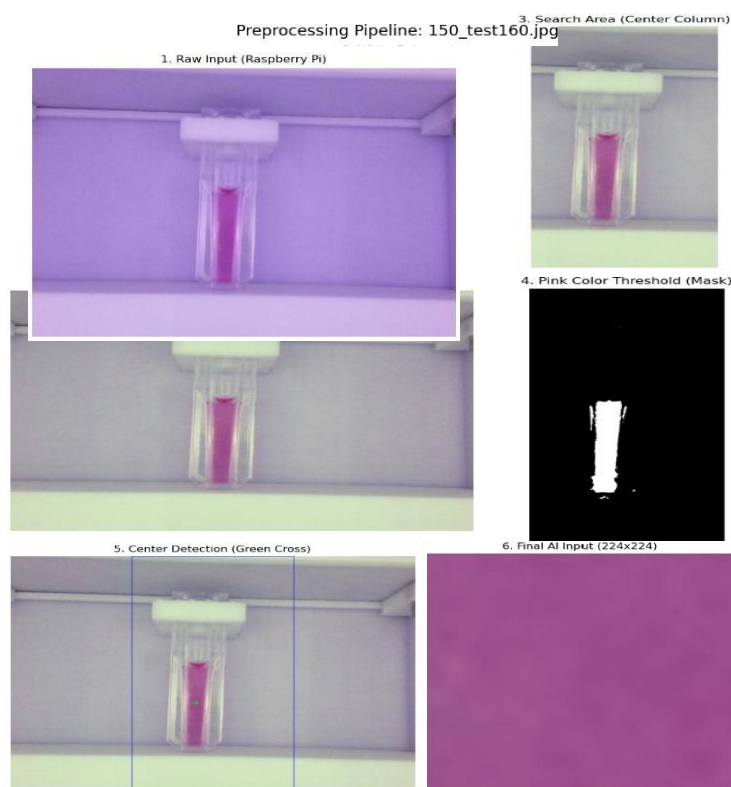


figure 1. Image preprocessing and ROI extraction pipeline

IV. EXPERIMENTS AND RESULTS

We had to prove the software actually works in a real clinical setting. To do this, we tested our AI predictions against certified medical hardware and international safety standards.

A. Ground Truth and Validation Setup

You can't train an AI without knowing the real answers. For every single one of our 183 physical serum samples, we ran a parallel test using a hospital-grade Spectrophotometer 4040. This gave us the exact, certified glucose concentration in mg/dL. We linked these "true" numbers to our 1,663 augmented images to build the dataset. To test the model honestly, we used Group 5-Fold Cross- Validation. Grouping is critical here. Because we zoomed and rotated our images to make the dataset bigger, we had to make sure the AI never trained on a "zoomed-in" photo and then tested itself on the "zoomed-out" version of that exact same sample. Grouping locks the physical samples together, so the test data is always 100% new, unseen chemistry to the AI.

B. Prediction Accuracy

The numbers out of the validation test were incredibly strong. Our Huber Regression model hit a Mean Absolute Error (MAE) of just 5.70 mg/dL. That means, on average, our cheap camera rig guessed the glucose level within about 6 points of the massive laboratory machine.

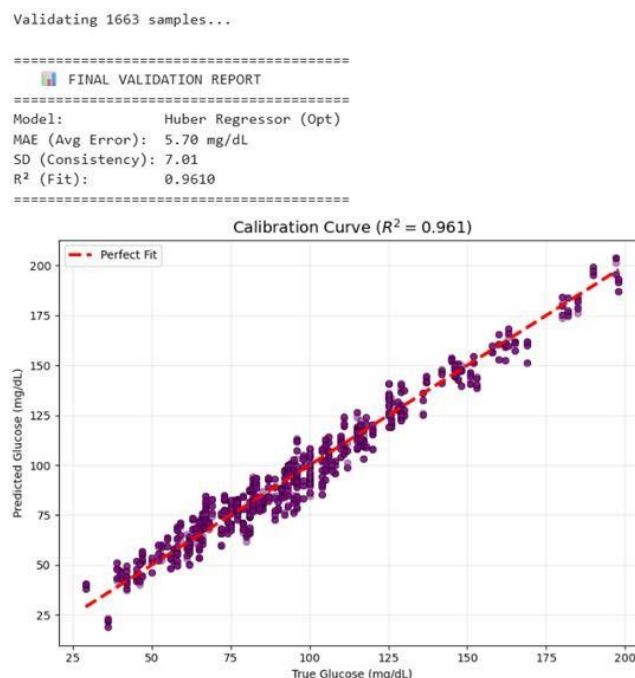


FIGURE 2. PREDICTED VS. ACTUAL GLUCOSE CONCENTRATION
($R^2=0.9610$)

We also looked at the R^2 score to see how well our curve fit the real-world chemistry. The model scored an R^2 of 0.9610. If you plot the predicted glucose against the true glucose, the dots form a very tight, straight line, as demonstrated in Fig.

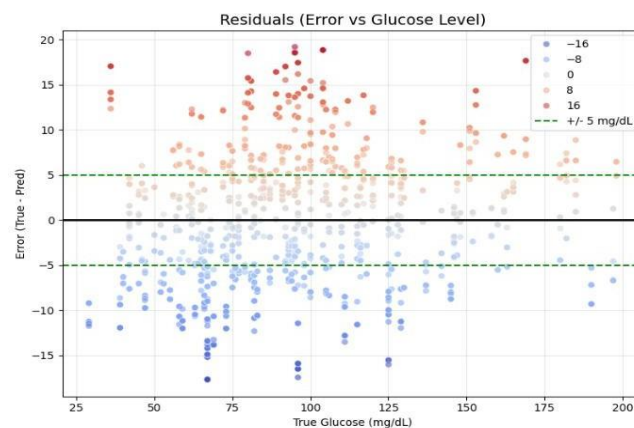
2. There was no wild guessing. The standard deviation of the error was only 7.01, meaning the results stayed consistent across the board.

C. Meeting Clinical Standards

Getting a low error is great, but we had to see if it was actually legal for medical use. The international benchmark for portable glucose monitors is the ISO 15197:2013 standard. That rule states a device must be within $\pm 15\%$ of the true lab value to be safe for patients.

If we look at a standard human glucose reading of 100 mg/dL, our MAE of 5.70 translates to a relative error of roughly 5.7%. That

comfortably beats the 15% ISO limit. When we ran a residual error analysis, depicted in Fig. 3, the mistakes were scattered evenly across the concentration range. The model didn't heavily over-predict high



glucose or under- predict low glucose. It just proved that a basic optical sensor, when backed by smart, outlier-rejecting math, can easily compete with heavily regulated hospital equipment.

FIGURE 3. RESIDUAL ERROR DISTRIBUTION ACROSS THE TESTED RANGE.

V. CONCLUSION

This project proves that you do not need a massive, expensive laboratory machine to get clinical-grade blood test results. By controlling the lighting with a simple 3D-printed black box, we turned a cheap consumer camera into a highly accurate medical sensor.

The real breakthrough here wasn't the hardware; it was the math. Using the Polynomial Huber Regression model allowed us to completely bypass the usual noise, bubbles, and reflections that normally ruin optical tests. We achieved an average error of just 5.70 mg/dL and a 96% correlation with the true chemistry. This

performance easily passes the strict ISO safety standards for portable medical monitors.

Looking forward, there is still room to grow. While the current setup works incredibly well, the next step is moving this software pipeline onto a custom hardware PCB. Pairing our robust math with a higher-quality camera sensor will likely push the performance even higher and cut the error rate down further. Ultimately, this approach opens the door for building a whole new generation of affordable, highly accurate diagnostic tools that can be deployed anywhere in the world.

REFERENCES

- [1] Y. Fan et al., “Smartphone-Based Colorimetric Detection of Glucose using Glucose Oxidase-Mimicking Enzyme Coupled with Machine Learning,” *ACS Omega*, vol. 6, no. 4, pp. 2686–2693, 2021. Available at:
<https://pubs.acs.org/doi/10.1021/acsomega.0c05680>
- [2] F. Li, Z. Wang, and J. Li, “An Optical POCT Device for Colorimetric Detection of Urine Test Strips Based on Raspberry Pi Imaging,” *Biosensors*, vol. 12, no. 10, p. 863, 2022. Available at:
<https://www.mdpi.com/2079-6374/12/10/863>
- [3] J. L. G. Aquino et al., “Smartphone-Based Colorimetric Sensor for Liquid Analysis: An Evaluation of Different Color Spaces,” *IEEE Access*, vol. 9, pp. 34746–34756, 2021.
Available at:
<https://ieeexplore.ieee.org/document/9361234>

**Comparative Data Driven Modeling of
Chemotherapy Drug Exposure Using
Artificial Neural Networks and Adaptive
Neuro Fuzzy Inference Systems**

24

Comparative Data Driven Modeling of Chemotherapy Drug Exposure Using Artificial Neural Networks and Adaptive Neuro Fuzzy Inference Systems

L. Elhaddad and M. Elfandi

Dept. of Electrical and Electronic Engineering Tripoli

Tripoli, Libya

LibyaLamishedad77@gmail.com

m.elfandi@uot.edu.ly

Abstract

Predicting chemotherapy-related drug exposure remains challenging because substantial interpatient variability and nonlinear physiological effects can weaken the reliability of fixed-dose strategies. This study investigates a data-driven framework for pharmacokinetic prediction using Artificial Neural Networks (ANN) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS). A curated pharmacokinetic dataset comprising 3077 observations and approximately 30 numerical variables was preprocessed through cleaning, missing-value handling, feature preparation, and exploratory statistical analysis prior to model development. Both models were implemented under the same supervised regression setting to predict the selected pharmacokinetic target variable. The ANN model achieved superior predictive performance, yielding a root mean square error (RMSE) of 8.1554 and faster convergence, whereas ANFIS produced an RMSE of 9.6026 while preserving greater structural interpretability through fuzzy rule-based inference. These findings indicate that ANN is more effective

for capturing complex nonlinear pharmacokinetic relationships in the present dataset, while ANFIS remains useful when model transparency is a design priority. The study supports the role of intelligent predictive modeling as a promising computational tool for chemotherapy dose assessment and future personalized treatment support.

Keywords: chemotherapy, pharmacokinetic modeling, biomedical engineering, artificial neural networks, adaptive neuro fuzzy inference system, data driven modeling, predictive modeling.

I. Introduction

Chemotherapy remains a central component of cancer treatment despite major advances in targeted and biologically specific therapies. Many cytotoxic agents continue to be used in standard oncology protocols because of their established therapeutic value across a wide range of malignancies. However, chemotherapy dosing remains inherently challenging because the relationship between administered dose and resulting systemic exposure is highly variable across patients. Small differences in pharmacokinetic behavior may shift treatment from an effective regimen to one associated with inadequate response or clinically significant toxicity. This problem is especially critical for drugs with narrow therapeutic windows, where safe and effective treatment depends on anticipating exposure behavior more reliably than fixed-dose strategies alone can provide [1], [2].

From a biomedical engineering perspective, chemotherapy administration can be interpreted as a nonlinear input-output system. Dose-related variables and patient descriptors act as inputs, while pharmacokinetic exposure measures represent observable outputs. In principle, such behavior may be studied using mechanistic

pharmacokinetic models derived from physiological transport and elimination processes. In practice, however, mechanistic identification is often limited by sparse observations, incomplete covariate information, heterogeneous study conditions, and the difficulty of capturing all sources of interpatient variability in a single structured model. Under these constraints, data-driven methods provide a practical alternative for approximating the relationship between measurable predictors and drug exposure outcomes directly from empirical data [2], [3].

A mechanistic view of pharmacokinetic modeling is illustrated by the white-box PK model showing the input–state–output relationship based on physiological differential equations (Fig. 1). In this representation, the dosing regimen acts as the system input, internal drug amounts represent latent dynamic states, and measured plasma concentrations form the observable outputs. This formulation clarifies why mechanistic pharmacokinetic modeling is physiologically interpretable, but also why it may become difficult to identify robustly when observations are sparse or heterogeneous.

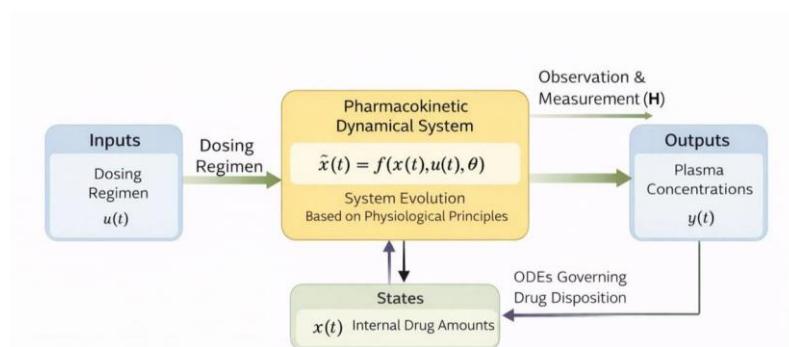


Fig. 1. White-box pharmacokinetic model illustrating the input–state–output relationship based on physiological differential equations.

Among data-driven approaches, Artificial Neural Networks (ANN) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) are particularly relevant for pharmacokinetic prediction problems characterized by nonlinear and multidimensional relationships. ANN models are well suited for high-dimensional regression tasks because they can learn flexible nonlinear mappings from data without imposing restrictive assumptions on the underlying functional form. ANFIS, in contrast, combines adaptive learning with fuzzy rule-based reasoning, offering improved interpretability through localized rule structures and gradual membership-based transitions. This distinction is important in biomedical applications, where predictive performance is valuable, but model transparency may also influence trust, interpretation, and future integration into decision-support environments [3]–[6].

The present study investigates this tradeoff in the context of chemotherapy-related pharmacokinetic prediction using a curated dataset compiled from published biomedical studies. After preprocessing, the modeling dataset contained 3077 observations and approximately 30 numerical variables describing administration-

related conditions, demographic structure, and pharmacokinetic descriptors. The prediction task was formulated as a supervised nonlinear regression problem targeting Avg_1, and both ANN and ANFIS were developed and evaluated under the same analytical setting. According to the reported results, ANN achieved stronger predictive performance, with an RMSE of 8.1554, compared with 9.6026 for ANFIS, while ANFIS preserved the advantage of structural interpretability through its fuzzy rule-based formulation. These findings are relevant to the broader direction of intelligent dose-support and model-informed prediction in biomedical engineering [1], [2], [6].

II. MATERIALS AND METHODS

A. Dataset and Prediction Task

This study formulated chemotherapy-related pharmacokinetic prediction as a supervised nonlinear regression problem. The analysis was conducted on a curated pharmacokinetic dataset compiled from published biomedical studies and processed in MATLAB. After preprocessing, the final modeling dataset contained 3077 observations and 30 numerical input variables, and the selected prediction target was Avg_1.

B. Data Preparation and Exploratory Analysis

The raw dataset contained both numerical variables and descriptive metadata. Only numerical variables were retained for statistical analysis and model development. The preprocessing stage included numeric feature isolation, missing-value handling, outlier screening, and feature standardization in order to produce a stable input space for regression modeling. Outliers were screened using standardized z-scores, but retained because extreme pharmacokinetic values may represent meaningful biological variability rather than noise. Missing

numeric values were imputed to preserve dataset size and maintain compatibility with multivariate analysis and predictive modeling.

Exploratory data analysis was then used to characterize the internal structure of the dataset before model fitting. In particular, the Pearson Correlation Matrix Heatmap of All Numeric Variables (Fig. 2) was used to inspect pairwise dependencies and possible feature redundancy

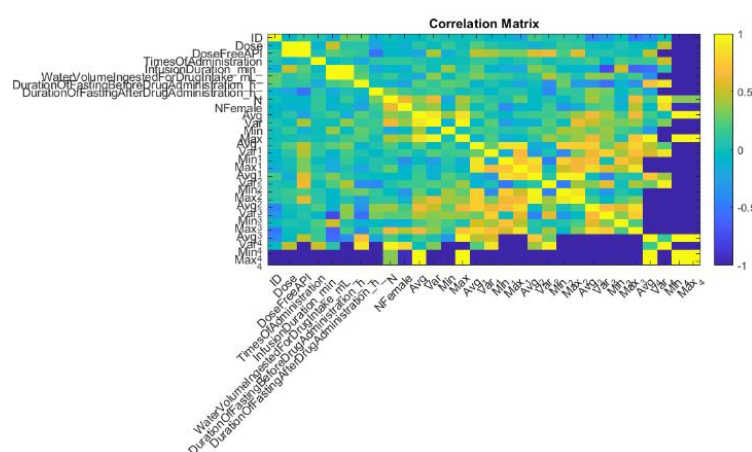


Fig. 2. Pearson correlation matrix heatmap of all numeric variables, illustrating pairwise dependencies and potential redundancy within the pharmacokinetic feature space.

As described in the thesis, this heatmap visualizes the strength and direction of correlation coefficients across all numeric variables and helps identify clusters of strongly related descriptors within the pharmacokinetic feature space. This step was useful for confirming that the dataset contained meaningful inter-variable structure rather than independent isolated predictors.

Principal Component Analysis was applied as a complementary global inspection step. Beyond the scree-based variance view, the Two-Dimensional Projection of Data onto the First Two Principal

Components (Fig. 3) was examined to visualize how samples were distributed in a reduced feature space.

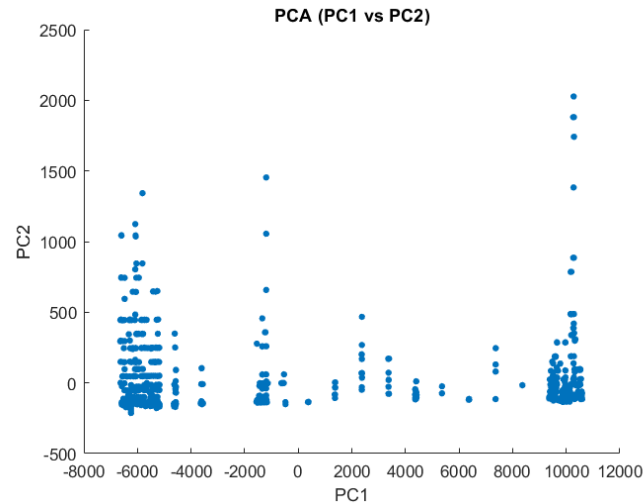


Fig. 3. Two-dimensional projection of the dataset onto the first two principal components, showing the global variance structure of the pharmacokinetic feature space.

This projection provided a compact representation of dominant variance patterns and helped assess whether the dataset exhibited broad structure, overlap, or clustering tendencies prior to supervised model development. Together, the correlation heatmap and PCA projection provided a smoother transition from exploratory analysis to modeling by showing that the input space contained both dependency structure and distributed multivariate variation.

Because the exploratory analysis indicated heterogeneous distributions, inter-feature dependence, and multiscale behavior, all input variables were standardized using z-score normalization before training. This step improved numerical stability and reduced the risk that features with larger magnitudes would dominate optimization during learning.

C. Train-Test Strategy and Model Development

After preprocessing and exploratory analysis, the dataset was partitioned using an 80–20 holdout strategy. The test subset was reserved for final evaluation, while the ANN model additionally used an internal validation subset for early stopping and generalization monitoring. This workflow ensured that both ANN and ANFIS were trained on the same processed feature space and evaluated under comparable conditions.

The first predictive model was a multilayer feedforward Artificial Neural Network (ANN) for nonlinear regression. The adopted architecture contained two hidden layers with 20 and 10 neurons, respectively, and training was performed using the Levenberg–Marquardt backpropagation algorithm with mean squared error as the optimization criterion. This configuration was selected to provide sufficient nonlinear approximation capacity while maintaining manageable model complexity for the dataset size.

The general structure of the adopted feedforward neural network is illustrated in the Architecture of a Feedforward Neural Network (Fig.

4). The diagram highlights the layered forward-propagation mechanism through which dose-related and patient-related descriptors are transformed into a continuous pharmacokinetic prediction.

Although the implemented model was trained for the specific target considered in this study, the architectural logic remains consistent with standard multilayer perceptron regression networks used for nonlinear biomedical prediction.

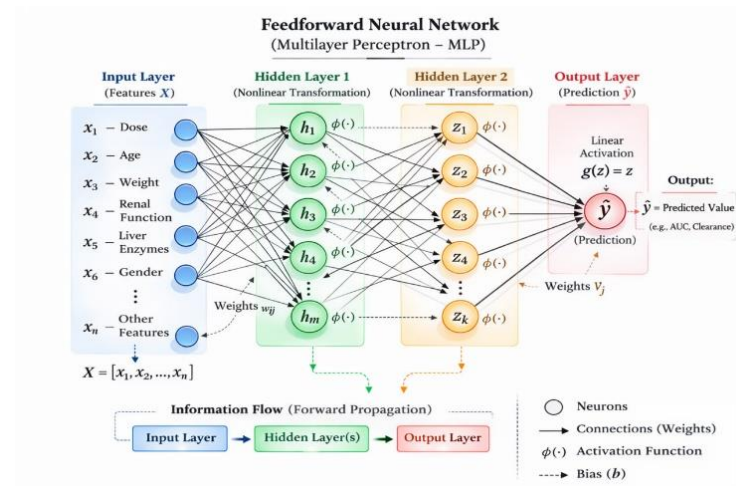


Fig. 4. Architecture of a feedforward neural network used to illustrate layered nonlinear regression from pharmacokinetic input descriptors to the predicted output.

The second predictive model was an Adaptive Neuro-Fuzzy Inference System (ANFIS) used as an interpretable nonlinear comparator. It was initialized using subtractive clustering with a clustering radius of 0.5, which generated a two-rule Sugeno-type fuzzy inference system with linear consequents. The resulting model contained 157 nodes, 62 linear parameters, 120 nonlinear parameters, and 182 adjustable parameters, and it was trained for 50 epochs. This structure preserved interpretability through fuzzy rule-based reasoning, although with more limited flexibility in a high-dimensional input space.

D. Evaluation Criteria

Model performance was evaluated using root mean square error (RMSE) and coefficient of determination (R^2). For the ANN model, regression alignment across training and validation subsets was also examined to assess generalization behavior. In addition to predictive accuracy, the comparative evaluation considered convergence behavior and interpretability, reflecting the study's broader objective

of comparing a high-capacity black-box regressor with a more transparent rule-based predictor.

III. RESULTS

A. ANN Performance

The ANN model demonstrated the strongest predictive performance in the studied regression task. Its learning behavior was stable and converged rapidly, with the best validation performance reached at epoch 17, indicating that the model captured the dominant nonlinear structure of the data early while maintaining effective generalization control. The ANN training behavior is illustrated in the ANN Training Performance Curve Showing Training and Validation MSE with Best Validation Point (Fig. 5), which shows rapid error reduction during the early epochs followed by stabilization near the optimal validation region.

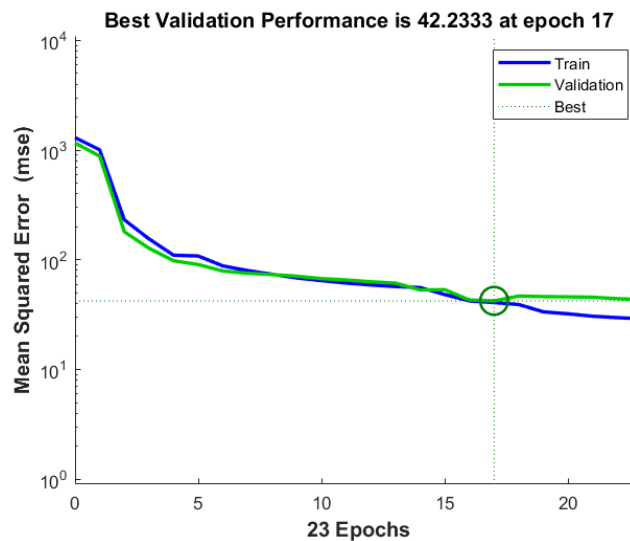


Fig. 5. ANN training and validation mean squared error curves, showing best validation performance at epoch 17.

This pattern supports the use of early stopping as an effective mechanism for limiting overfitting while preserving predictive accuracy.

A more complete view of ANN predictive behavior is provided by the ANN regression plots for training, validation, and overall data with fitted lines and correlation coefficients (Fig.6).

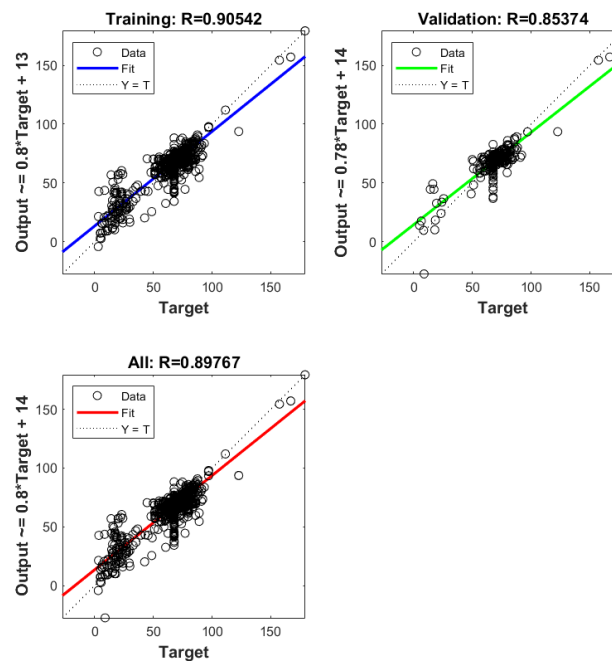


Fig. 6. ANN regression performance across training, validation, and overall data, with fitted lines and correlation coefficients.

These regression plots show a strong linear association between predicted and observed values across all subsets, with reported correlation coefficients of $R = 0.90542$ for training, $R = 0.85374$ for validation, and $R = 0.89767$ overall. The relatively small gap between training and validation correlation indicates that the network retained

good predictive consistency on unseen samples rather than memorizing the training data.

The ANN error characteristics further support this interpretation. As shown in the ANN Error Histogram Showing Distribution of Residuals for Training and Validation Samples (Fig. 7), most residuals are concentrated around zero, suggesting that the model produced controlled predictions with no severe systematic distortion.

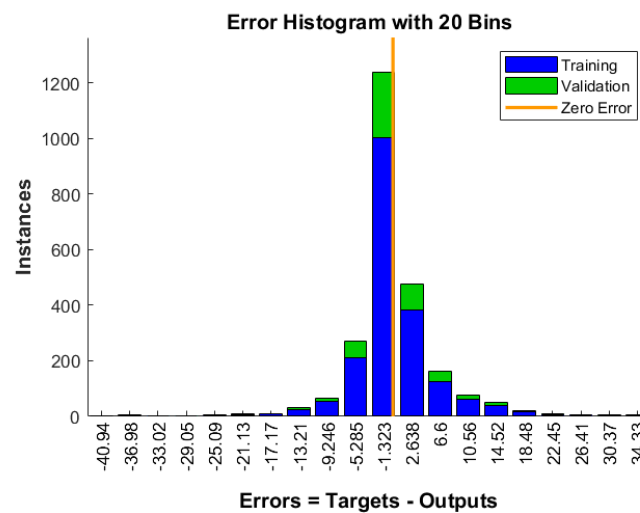


Fig. 7. Residual error histogram of the ANN model for training and validation samples.

The residual distribution is not perfectly symmetric, which is consistent with mild underestimation in higher target ranges, but the overall spread remains sufficiently compact for a heterogeneous biomedical regression problem. Quantitatively, the ANN achieved an RMSE of 8.1554 and an R^2 of 0.6862, indicating that it explained a substantial proportion of target variability while maintaining relatively low prediction error in the original output scale.

Taken together, the ANN results indicate three important properties: efficient convergence, strong predictive alignment, and acceptable generalization. These characteristics suggest that the feedforward architecture provided sufficient nonlinear approximation capacity for the selected pharmacokinetic target without requiring explicit mechanistic reconstruction of the underlying system.

B. ANFIS Performance

The ANFIS model also learned a coherent input-output mapping, but its predictive precision was lower than that of the ANN. Its training behavior is shown in the ANFIS Training Curve Showing RMSE Reduction Over 50 Epochs (Fig. 8), where the error decreases gradually across the full training interval.

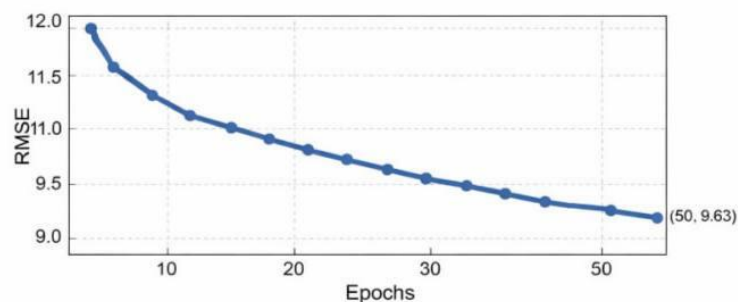


Fig. 8. ANFIS training curve showing the gradual reduction of RMSE over 50 epochs.

Unlike the ANN model, which reached its optimal validation point relatively early, ANFIS required the full 50 epochs to approach its minimum training error. This slower convergence reflects the more gradual adaptation of membership functions and consequent parameters in the neuro-fuzzy framework.

The predictive behavior of ANFIS on the testing subset is shown in the Regression Performance of the ANFIS Model for the Testing Dataset (Fig. 9).

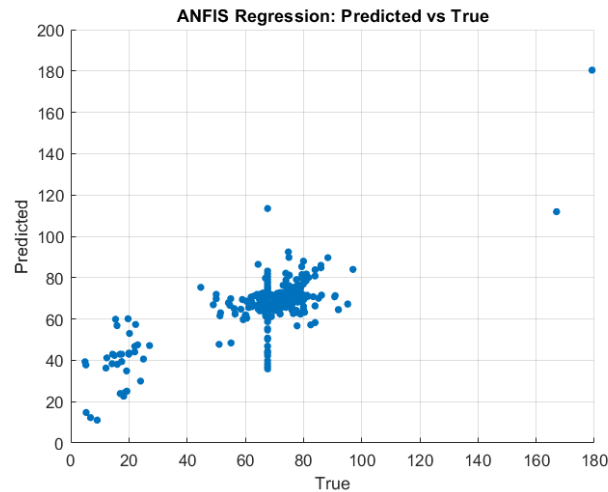


Fig. 9. Predicted versus observed values for the ANFIS model on the testing dataset.

The scatter pattern follows the expected upward trend, confirming that the model captured the general structure of the regression problem. However, the dispersion around the ideal prediction trend is visibly broader than that observed for ANN, indicating lower point-to-point precision. This behavior is consistent with a compact two-rule Sugeno-type system that preserves smoothness and interpretability but offers more limited flexibility in a high-dimensional feature space.

In quantitative terms, the ANFIS model achieved an RMSE of 9.6026 and an R^2 of 0.63. These values confirm that the model remained predictive and scientifically meaningful, but did not match the accuracy obtained by ANN. From a modeling perspective, this result suggests that the adopted neuro-fuzzy structure captured broad trends effectively yet was less capable of reproducing finer nonlinear variation across the full dataset. Even so, ANFIS retains an important advantage in structural interpretability because its predictions emerge from an explicit fuzzy-rule framework rather than a purely hidden internal representation.

C. Comparative Analysis between ANN and ANFIS

The comparative findings are summarized most directly by the RMSE Comparison Between ANN and ANFIS Models (Fig. 10), which shows a clear quantitative advantage for ANN over ANFIS.

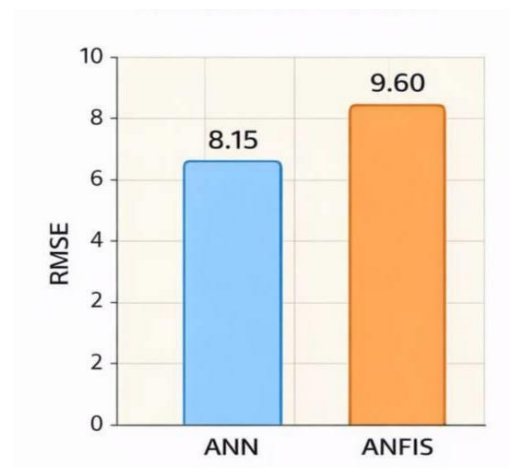


Fig. 10. Comparative RMSE performance of ANN and ANFIS for pharmacokinetic target prediction.

This figure provides the most compact visual summary of the central experimental result: the ANN model achieved lower prediction error and therefore stronger overall predictive performance for the selected pharmacokinetic target.

The same conclusion is reinforced numerically in Table 1: Quantitative Performance Comparison between ANN and ANFIS, which reports the key outcome measures for both models in a single location.

Table I. Quantitative Performance Comparison between Ann and Anfis.

Model	RMSE	R ²	Convergence Epoch
ANN	8.1554	0.6862	17
ANFIS	9.6026	0.63	50

As summarized in this table, ANN achieved RMSE = 8.1554, R^2 = 0.6862, and best convergence at epoch 17, whereas ANFIS achieved RMSE = 9.6026, R^2 = 0.63, and reached minimum training error at epoch 50. Presenting these values in tabular form is useful because it allows the reader to compare predictive accuracy, explained variance, and convergence behavior simultaneously without relying on separate paragraphs of description.

Overall, the comparison indicates that ANN provided the stronger predictive solution for this regression problem, particularly in terms of lower error, faster convergence, and stronger validation behavior. ANFIS, however, remained valuable as an interpretable comparator. The results therefore do not simply identify a single “winner”; they demonstrate a meaningful tradeoff between approximation power and structural transparency. For this dataset and target variable, the evidence favors ANN when predictive accuracy is the primary objective, while ANFIS remains relevant when explainability is a design priority.

IV. DISCUSSION

The results show that data-driven modeling can recover meaningful chemotherapy-related pharmacokinetic patterns from heterogeneous numerical descriptors, even without explicit mechanistic reconstruction of physiological dynamics. This is important in

biomedical engineering because chemotherapy response is influenced by nonlinear interactions among dosing conditions, patient variability, and exposure-related behavior [1], [2],[3].

A key finding is the superior performance of the ANN model relative to ANFIS. ANN achieved lower prediction error, higher explained variance, and faster convergence, indicating greater ability to represent the nonlinear structure of the selected regression task. This is consistent with the stronger representational flexibility of multilayer neural networks, which are better suited to high-dimensional heterogeneous predictors [7],[9]. By contrast, the adopted ANFIS configuration preserved interpretability through a compact fuzzy-rule structure, but this likely limited its ability to capture finer nonlinear variation across the dataset [8], [10].

This difference should not be interpreted as a simple comparison between a stronger and weaker method, but rather as a tradeoff between two modeling priorities. ANN offers greater approximation power and predictive scalability, whereas ANFIS offers improved transparency and rule-based interpretability. In biomedical applications, both properties can be important depending on whether the primary goal is prediction accuracy or explainable model behavior [6], [8].

The achieved performance is promising, but it should be interpreted carefully. The reported R^2 values indicate that the models captured a substantial portion of the target variability, while also leaving some variation unexplained. This is expected in pharmacokinetic prediction, where biological heterogeneity, incomplete covariate capture, and study-level variability can limit predictive recovery from tabular descriptors alone. Accordingly, the results support the feasibility of machine learning-based exposure prediction, but not full physiological characterization [1],[3], [5].

Several limitations should also be noted. The dataset was compiled from published biomedical sources rather than a single controlled clinical pipeline, which likely introduced heterogeneity in measurement conditions and feature completeness. In addition, the analysis relied on structured numerical descriptors rather than full concentration-time trajectories, limiting direct recovery of temporal pharmacokinetic dynamics. The ANFIS model was also intentionally compact, and richer fuzzy-rule designs may improve performance at the cost of reduced interpretability. Future work should therefore include external validation, richer patient-level covariates, and broader precision-dosing evaluation frameworks [2], [3], [6].

Overall, the findings indicate that ANN provides the more effective predictive solution for the selected pharmacokinetic target, while ANFIS remains valuable as an interpretable nonlinear comparator. The broader significance of the study lies in showing that intelligent regression models can support chemotherapy-related pharmacokinetic estimation while highlighting the persistent tradeoff between predictive strength and model transparency [2],[6].

V. CONCLUSION

This paper presented a comparative data-driven study for chemotherapy-related pharmacokinetic prediction using Artificial Neural Networks and Adaptive Neuro-Fuzzy Inference Systems. The analysis showed that both models were able to learn meaningful relationships between the selected numerical descriptors and the target pharmacokinetic variable, confirming the feasibility of intelligent regression methods for this type of biomedical prediction task.

Among the two approaches, ANN provided the stronger predictive performance, achieving lower RMSE, higher explained variance, and faster convergence. ANFIS, although less accurate, remained valuable as an interpretable nonlinear alternative due to its fuzzy rule-based

structure. The comparison therefore highlights an important tradeoff between predictive strength and model transparency, which is highly relevant in biomedical engineering applications.

REFERENCES

- [1] O. Teplytska, M. M. Weng, I. K. Minichmayr, N. Taubert, T. B. van der Heggen, N. Jakob, C. Kloft, and U. Jaehde, "Machine learning methods for precision dosing in anticancer drug therapy: A scoping review," *Clinical Pharmacokinetics*, vol. 63, no. 9, pp. 1221–1237, Sep. 2024.
- [2] I. K. Minichmayr, S. G. Wicha, C. Kloft, and U. Fuhr, "Model-informed precision dosing: State of the art and future perspectives," *Advanced Drug Delivery Reviews*, vol. 215, p. 115421, Dec. 2024.
- [3] S. Huang, Q. Xu, G. Yang, J. Ding, and Q. Pei, "Machine learning for prediction of drug concentrations: Application and challenges," *Clinical Pharmacology and Therapeutics*, vol. 117, no. 5, pp. 1236–1247, May 2025.
- [4] J. Zhao, S. Dai, J. He, N. Liu, B. Zhang, and S. Li, "Prediction of high-dose methotrexate blood concentration in osteosarcoma patients using machine learning," *Drug Design, Development and Therapy*, vol. 19, pp. 3631–3643, 2025.
- [5] S. G. Wicha, L. Liu, and A. S. Darwich, "From therapeutic drug monitoring to model-informed precision dosing for antibiotics," *Clinical Pharmacology and Therapeutics*, vol. 109, no. 4, pp. 928–941, 2021.
- [6] E. A. Poweleit, A. A. Vinks, and T. Mizuno, "Artificial intelligence and machine learning approaches to facilitate therapeutic drug management and model-informed precision dosing," *Therapeutic Drug Monitoring*, vol. 45, no. 2, pp. 143–150, Apr. 2023.
- [7] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals, and Systems*, vol. 2, no. 4, pp. 303–314, Dec. 1989.

- [8] J.-S. R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–685, May/Jun. 1993.
- [9] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2009.
- [10] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.