

RESEARCH ARTICLE

OPEN ACCESS

NOVEL WT-MOE-SSM-ECHOFORMER BASED HYBRID FRAMEWORK FOR INTELLIGENT ECG ANALYSIS AND CARDIAC DISEASE PREDICTION

S. Nagakumararaj^{*1}, P. Archana Devi², D. Justin Jose³, S. Sree Vidhya⁴

¹Assistant Professor, Department of Computer Science and Engineering (Cyber Security), SSM Institute of Engineering and Technology, Dindigul, Tamil Nadu, India.

²Assistant professor, Department of Information Technology, Mailam Engineering College, Mailam, Tamil Nadu, India.

³Assistant Professor, Department of Artificial Intelligence and Data Science, St. Xavier's Catholic College of Engineering, Chunkankadai, Nagercoil, 629003, Tamil Nadu, India.

⁴Assistant Professor, Department of Computer Science and Engineering, Erode Sengunthar Engineering College, Erode, Tamil Nadu, India.

¹<https://orcid.org/0000-0003-2123-6541>^{id}, ²<https://orcid.org/0009-0004-1436-2480>^{id}

³<https://orcid.org/0009-0009-6708-7765>^{id}, ⁴<https://orcid.org/0009-0004-0153-9144>^{id}

Email: [*kumarrs2005@gmail.com](mailto:kumarrs2005@gmail.com), archana.devi1995@gmail.com, justinjosed@sxcce.edu.in, sreevidhyascs@gmail.com

ARTICLE INFO

Article History

Received: February 13, 2026

Reviewed: April 21, 2026

Accepted: May 21, 2026

Published: June 30, 2026

Keywords:

Cardiovascular Disease,
Electrocardiograms,
Savitzky-Golay Filter,
Exploratory Data Analysis,
Pan-Tompkins's algorithm,
WT-MoE-SSM-EchoFormer,
Golden Tortoise Beetle
Optimizer..

ABSTRACT

Cardiovascular Disease (CVD) continues to be a major global health concern, posing significant risks to human life. Early and accurate detection remains difficult due to noise in Electrocardiograms (ECGs) signals, class imbalance, and the limited generalisation ability of existing Machine Learning (ML) models. Therefore, this work proposes a new hybrid optimized Deep Learning (DL) techniques for the intelligent analysis of ECG signals for the prediction of cardiac disease. The study uses raw ECG Signal from the MIT-BIH Arrhythmia Database as a case study, and applies Savitzky-Golay (SG) filters for denoising to ensure proper noise elimination while maintaining the necessary morphological structures. Exploratory Data Analysis (EDA) is utilized for statistical descriptors such as mean, variance, skewness and kurtosis of RR intervals and QRS widths, complemented by correlation matrix and QRS morphology overlays to uncover temporal patterns. Pan-Tompkins's algorithm applied for extracting morphological features, such as RR intervals, QRS duration, Heart Rate, and Rhythm Variability enabling precise identification of P, QRS and T complexes. The new proposed framework, Wavelet-Token Swin-1D Mixture-of-Experts State-Space EchoFormer (WT-MoE-SSM-EchoFormer) uses wavelet domain tokenization for feature encoding, Swin 1D hierarchical attention and expert driven selection learning for better represent cardiac signals through state space temporal modelling. The Golden Tortoise Beetle Optimizer (GTBO) algorithm employed for dynamically tuning hyperparameters, thus improving convergence and training stability. The complete framework is implemented in Python and yielded an accuracy of 99.98%, a precision of 99.79%, a recall of 94.84%, an F1-score of 97.03%, a specificity of 99.99%, and a ROC-AUC of 1.0000, which validates the proposed framework's effectiveness and reliability for ECG-based cardiac disease diagnosis. The proposed framework achieves superior classification accuracy, lower misclassification, and greater potential for real-time cardiac disease diagnosis.



Copyright ©2026 by authors and Galileo Institute of Technology and Education of the Amazon (ITEGAM). This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

I. INTRODUCTION

CVD are one of the most critical and life-threatening health issues across the globe, caused by the improper functioning of the cardiovascular system, resulting in severe complications like heart failure, coronary artery disease and myocardial infarction [1], [2]. Heart disease significantly contributes to mortality and morbidity across the globe; it is important to detect heart disease accurately for effective treatment and prevention [3]. It is critical to recognize people at risk of heart disease at a very early stage to prevent severe complications [4]. In recent years, significant progress has been made in heart disease detection through the combined use of computational intelligence and medical technology. Methods such as ML, data mining and predictive modelling rely on large amount of clinical and physiological data to enhance the accuracy of diagnosis [5].

Although advanced imaging modalities such as Computed Tomography (CT) scans and Magnetic Resonance Imaging (MRI) are reliable and accurate, but it is high cost and lack of accessibility, limit their widespread use, especially in developing countries. ECG is an essential and popular technique used for the diagnosis of CVD by capturing the electricity activity of the heart [6], [7]. Among the different types of CVDs, cardiac arrhythmia is the most common one and easily detected from the ECG signal by analysing irregular heartbeats. Conventional ECG system uses 12-lead configuration, which is quite inconvenient for patients and require the expertise of a professional for proper analysis [8], [9]. To overcome these limitations, single lead ECG sensors have gained popularity as a more comfortable and practical solution. By integrating AI technology, these systems are automatically analysing the ECG signal and detect any cardiac irregularities [10]. Manual analysis of ECG signals is a time consuming and laborious process, which has led to the development of ML algorithms for the automatic classification of arrhythmias. However, traditional ML models such as XG-Boost [11], Extra-Trees [12] and Artificial Neural Network (ANN) [13] classifiers are highly dependent on manual feature extraction and selection, which makes them less robust and less accurate. To overcome these issues, DL algorithms have occurred as popular alternatives owing to their capability to mechanically learn complex hierarchical features from large datasets.

DL models have demonstrated incredible success in a wide range of application, including biomedical signal processing and medical imaging analysis, which has significantly improved the accuracy of cardiac disease analysis. The Convolutional Neural Network (CNN) to developed Rhythmi [14] model operates significantly better in terms of early diagnosis and patient care. However, its usage is currently delayed by the lack of diversity in the dataset, the absence of personalized medical reporting and the need for extensive clinical validation before largescale deployment. VGG16 [15], Google ViT based ResNet50 [16] and HybDeepNet [17] models makes possible to predict cardiac disease at a low cost and improves accessibility to continuous heart monitoring. Nevertheless, the model shows performance degradation after hardware deployment and require further optimization to ensure real-world healthcare system. Although these models have demonstrated encouraging results and they also have some significant limitations, including their sensitivity to feature selection and data quality. The inability of traditional models has complex nonlinear patterns, high computational complexity and lack of interpretability of DL models, thus difficult to handle noisy, imbalanced medical data, and scalability issues that make them unsuitable for real time healthcare environment. These boundaries indicate the necessity for robust, interpretable and computationally efficient predictive models for accurate heart disease diagnosis. Recently, Vision Transformer (ViT) integrates Swin [18], TransConvNet [19] model has improved by employing self-attention mechanisms which decrease image specific biases and enhance both computational efficacy and accuracy. However, challenges such as insufficient deep spatial feature extraction and high computational cost [20].

I.1 RESEARCH GAPS

The existing approaches for ECG-based prediction of cardiac disease are mostly based on traditional ML and DL models, which tend to have high human intervention in feature extraction, low robustness to noisy data, and poor performance on highly imbalanced datasets. Most of the existing DL models have high computational complexity, low interpretability, and poor ability to model both short-term morphological dynamics and long-term temporal patterns in ECG signals. Although the transformer-based models have improved the ability to model the attention mechanism, they still have challenges in terms of low multiresolution feature representation and high computational complexity. Moreover, most of the existing approaches have not fully leveraged the benefits of signal denoising, EDA, adaptive feature learning, and hyperparameter optimization in a unified framework.

I.2 CONTRIBUTIONS

To overcome the limitations of existing methods, the proposed work contributes to,

- SG filter utilized for efficient denoising of raw ECG signals, while preserving vital morphology features such as QRS complexes and RR intervals that are necessary for reliable cardiac analyses.
- Using EDA, patterns of statistical and temporal information, and morphology information derived from ECG signals are analyzed to provide a better understanding of cardiac rhythms.
- To effectively detect R-peaks and extract clinically significant temporal features from ECG signals, a Pan-Tompkins algorithm is utilized.
- A hybrid WT-MoE-SSM-EchoFormer is proposed to be used for capturing multi-resolution feature representations, adaptive expert representations, and both short- and long-term temporal dependencies from ECG signals.
- GTBA used to automatically tune the hyperparameters of proposed model, resulting in stable training behaviour, and enhanced classification performance.

II. PROPOSED SYSTEM DESCRIPTION

The improved model as illustrated in Figure 1, generates a classification output that identifies both the presence and type of cardiac disease. The acquisition of raw ECG signals from the cardiovascular dataset captures electrical activity of the heart. Since these data are frequently contaminated with noise caused by muscle movement, electrode displacement and external interference, the first necessary step is signal processing. Therefore, a SG filter is applied to the ECG signals to remove unwanted noise while keeping the waveform's original structure and significant peaks. After acquiring clear ECG signals, the next stage is EDA. Before developing learning model, EDA is used to better understand the behaviour and distribution of cardiac signals. At this point, statistical metrics such as variance, means, skewness and kurtosis of RR intervals and QRS widths are calculated. Additionally, correlation analysis and QRS morphology visualization are performed. This process assists in detecting hidden pattern, irregularities, and correlations in the data. The framework then proceeds to feature extraction, where the Pan-Tompkins algorithm accurately detects the P, QRS and T complexes in the ECG signal. This step is to convert the raw ECG signal into clinically relevant features that reflect the real physiological state of the heart. These extracted features simplify the data and improve the model's discriminative capacity.

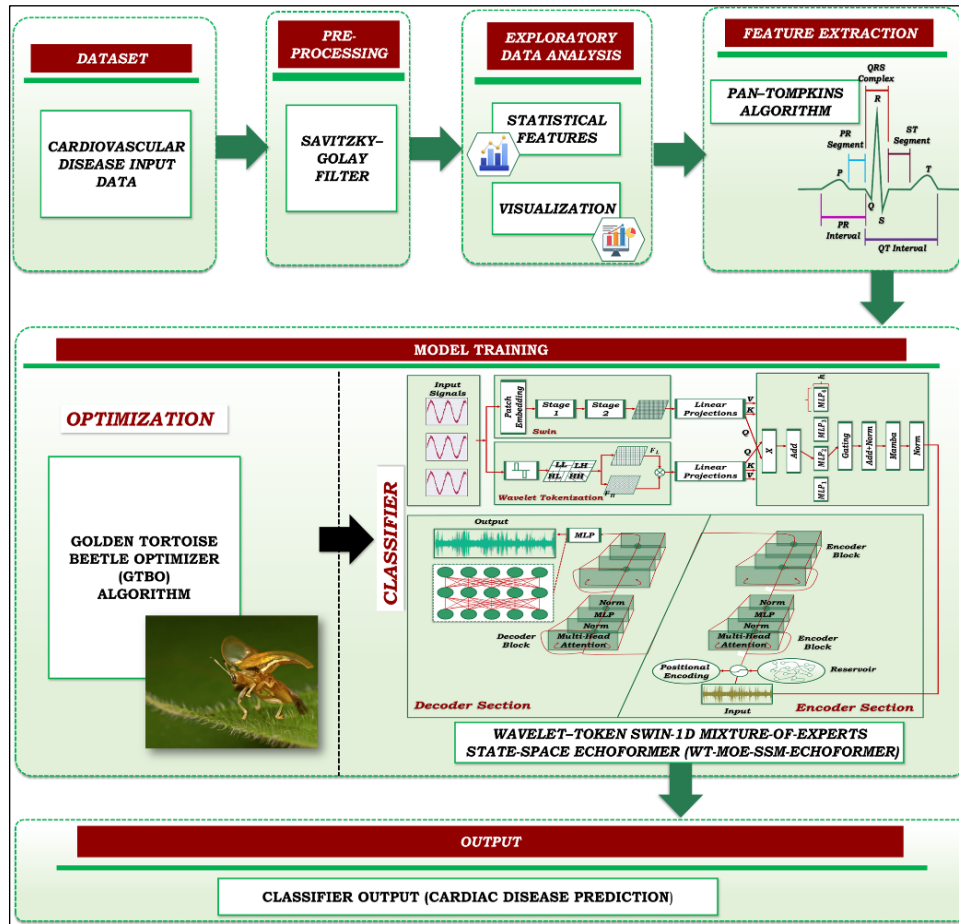


Figure 1: Proposed ECG analysis and cardiac disease prediction using GTBO WT-MoE-SSM-EchoFormer.
Source: Authors, (2026).

The extracted characteristics are then fed into the proposed WT-MoE-SSM-EchoFormer classifier. This model manages both short term and long-term temporal pattern in ECG data. This combination produces a very informative visualization of heart activity, allowing for accurate classification of cardiovascular diseases. During model training, the GTBO is used to automatically tune the hyperparameters. This optimization step is crucial for achieving fast convergence, avoiding overfitting and maintaining training stability.

III. METHODOLOGY

III.1 SAVITZKY-GOLAY FILTER

In the cardiac disease prediction framework applied a SG filter is used to preprocess the ECG signals. The SG filter is a time-domain method that denoises the ECG signals although conserving risky waveform attributes of ECG signal by approximating the signal around each sample using a local polynomial least-squares fit. The approximation of the ECG signal allows for effective suppression of noise in the ECG signal without impacting the critical features of the cardiac waveform.

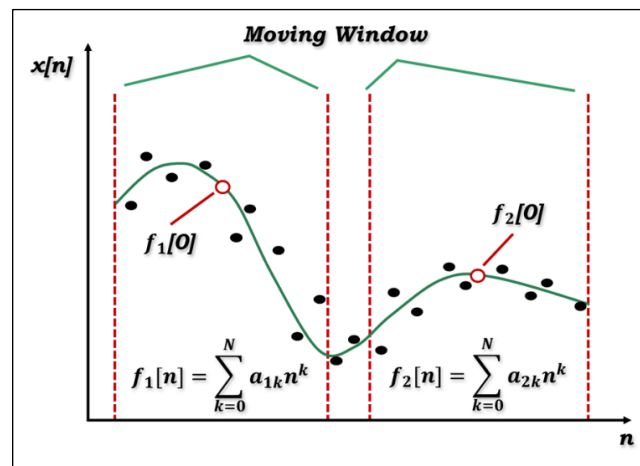


Figure 2: Basic idea of SG filter.
Source: Authors, (2026).

Let $x[n]$ represent the ECG signal in discrete format as illustrates in Figure 2. To approximate the signal using a polynomial of degree N for any given sample point, a window of length $2M + 1$ is centered around the sample point.

$$f(n) = \sum_{k=0}^N \alpha_k n^k \quad (1)$$

To obtain the optimum polynomial coefficients α_k , minimise the least squares error between the polynomial fit versus the original ECG samples in the reference window.

$$\varepsilon = \sum_{n=-M}^M (\sum_{k=0}^N \alpha_k n^k - x[n])^2 \quad (2)$$

In the case of a polynomial, the filtered ECG value at the reference window is located at the centre of the polynomial.

$$y[0] = f(0) = \alpha_0 \quad (3)$$

This process is repeated for all ECG samples by moving the reference window over the ECG signal resulting in a complete denoised ECG signal. The characteristics of the SG filter make them highly suitable for fitting a polynomial to retain low-order signal components and small morphological features while reducing high-frequency noise. In ECG analysis, the accurate retention of QRS morphology and RR intervals are important to ensure accurate feature extraction and facilitate precise classification of cardiac diseases.

III.2 EXPLORATORY DATA ANALYSIS

EDA is the process of getting a thorough understanding of the statistics, time-related trends associated with ECG waveforms before the feature extraction and model building. In addition to testing the quality of the data, it also helps reveal significant cardiac behaviours associated with various cardiovascular health issues.

III.2.1 Analysis of Descriptive Statistics

Descriptive Statistics values from both the raw ECG waveform and medically significant cardiac variables are calculated especially in RR Interval and QRS duration using the following:

- Mean - The average value of data, calculated by summing all data values and dividing by the sample size. It gives the central tendency for heart rhythm and timing of timing for ventricular depolarization.
- Median - The value positioned in the middle of the sorted data set. It serves as a leader that divides data into two equal halves when arranged by value ascending or descending.
- Standard Deviation – It measure the distance of values from the mean. It indicates the degree of fluctuation in cardiac activity.
- Variance - Variance is the mean of the squared differences of the individual observations from the mean value and shows how widely cardiac activity varies.
- Skewness: This provides insight into how asymmetric the distribution of QRS widths and RR intervals are which is helpful for identifying abnormal conduction patterns and irregular heartbeats.
- Kurtosis: It provides a quantitative assessment of both the percent of kernel volume and tail behaviour of distributions, helping to identify extreme cardiac events and transient irregularities.

Descriptive statistics gives an important understanding of the physiological function of the heart, and assist in distinguishing between normal and abnormal conditions.

III.2.2 Visualisation and Analysis of Temporal Patterns:

To further understand the relationship between these features and their temporal dynamics, multiple approaches are used:

- Correlation Matrices: They are created using extracted cardiac features allowing the identification of strong dependencies between each parameter, including QRS widths, RR intervals and other parameters, which together effect disease classification.
- QRS Morphology Overlays: The superimposition of several QRS complexes permits analysis of the degree of similarity or deviation in their shape throughout time; this helps to identify subtle morphological variations associated with arrhythmias and myocardial abnormalities.
- Histograms: It used to examine the distribution of continuous variables, and how the specific values of those variables affect the distribution of RR intervals and QRS widths. Histograms allow for examination of a given variable's distribution by dividing the variable into bins and showing how many observations fall into each bin.
- Box Plots: It allows for visualization of the median, quartiles, and outliers for a continuous variable. They serve the same purpose as histograms, but visually convey the quartile distribution, median, and outlier information for continuous variables.
- Line Plot: Displays a set of data points to illustrate a trend over time or across continuous variables using straight line segments. Commonly utilised for analysing and visualizing the temporal changes of ECG waveforms and extracted features to determine/identify rhythm irregularities and dynamic cardiac performance.

The integration of statistical descriptors like mean, median, skewness and kurtosis for RR intervals & QRS widths with correlation matrices and QRS morphology overlays in the EDA process helps to establish both temporal patterns and structural relationships present in the ECG data. This results in increased-quality features for the proposed cardiac disease prediction model.

III.2.3 Feature Extraction- Pan-Tompkins Algorithm

Pan-Tompkin's algorithm is used to aid in identifying the R point in an ECG signal. This algorithm consists of five main steps, which include bandpass filtering, differentiation, squaring, sliding window integration, and thresholding.

Band-Pass Filtering for Noise Reduction

In the initial stage, the Pan-Tompkins algorithm utilizes a band-pass filter to reduce noise in the ECG signal. This is accomplished by using a series of low-pass and high-pass filters to create a band-pass filter with cascaded low-pass and high-pass filters. The result of this processing step is a more noticeable QRS complex and less baseline drift and high-frequency noise.

Differentiation

Differentiation to obtain the slope characteristics of the QRS complex, at a high-frequency resolution moving away from low frequencies. It provides an increased opportunity for successful detection of QRS complexes.

Squaring and Sliding Window Integration

Squaring and sliding window integration to provide further emphasis of the differences made by the previous operation as well as ensure that all values are positive. To accomplish this, the squared output is further integrated using a sliding window approach that is create a smoothed output with both QRS waveform energy and the least amount of residual noise possible.

Extracting Temporal Features in Relation to the R-Wave

Once the R-wave locations are identified, ECG signals are divided into segments based on their temporal proximity to the R-peaks. From these relationships, average-mean feature vectors to be obtained for the following items within each of the ECG signals:

- P Interval
- PR Interval
- QRS Duration
- QT Interval
- ST segment
- T-wave
- RR Interval

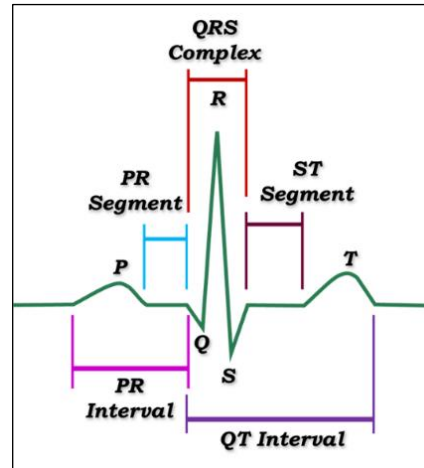


Figure 3: Basic structure of ECG signals.

Source: Authors, (2026).

The repeated wave sequences of P, QRS and T waves that are associated with each heartbeat characterized the ECG signal as illustrates in Figure 3.

RR interval Deviation

The deviation of RR intervals is then calculated to assess rhythm regularity:

$$RR_{deviation} = \frac{1}{RR_{number}} \sum |RR_i - RR_{avg}| \quad (4)$$

Here, RR_{number} stands total number of RR intervals.

QRS Complex Duration

A model based on physiological limitations utilized to determine the proper morphological characteristics of the QRS Complex to predict cardiovascular disease. A total QRS duration of 0.11s is used as the upper limit for a normal range of QRS Durations, and the Q-interval is less than 25% of the total duration of the QRS. The R peaks from the Pan-Tompkins Algorithm serve as temporal references for estimating the QRS boundaries. The QRS half-width is defined as half of the maximum QRS duration (0.055 s).

The Q interval is one-fourth of the QRS duration (0.0275 s), which meets clinical requirements, and the remaining QRS duration is assigned to the R and S waves.

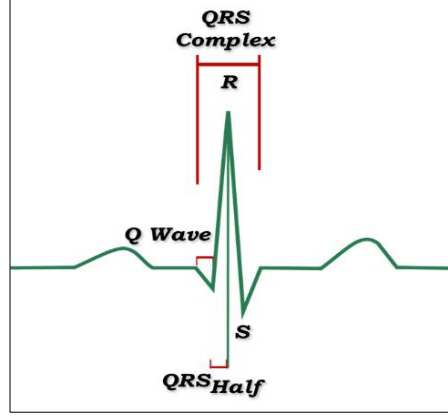


Figure 4: Intervals calculation of QRS complex.
Source: Authors, (2026).

Figure 4 illustrates intervals used for evaluation of QRS complex. This method of QRS modelling allows for the careful extraction of QRS Duration characteristics, which are important indicators of abnormalities in ventricular conduction, and therefore are an important component of automated systems for predicting CVD.

III.3 WT-MOE-SSM-ECHOFORMER

WT-MoE-SSM-EchoFormer classifier combines wavelet-domain tokenization for multiresolution feature extraction, Swin-1D hierarchical attention for localized temporal context modeling, mixture-of-experts routing for adaptive feature specialization, and state-space echo former dynamics for long-term temporal memory. This unified design allows for strong encoding of both transient and continuous cardiac rhythm variations.

Wavelet tokenization

Let the pre-processed 1D cardiac signal be denoted as,

$$x \in \mathbb{R}^{T \times B}, \quad (5)$$

Here, T stands temporal length and B stands number of signal channel. A 1D wavelet tokenization is applied independently to each channel:

$$LL^{(b)}, LH^{(b)} = WT(x^{(b)}), b = 1, \dots, B \quad (6)$$

Here, $LL^{(b)}$ and $LH^{(b)}$ represents approximation and detail coefficients. The same type of coefficient is stacked to form:

$$LL \in \mathbb{R}^{T/2 \times B}, LH \in \mathbb{R}^{T/2 \times B} \quad (7)$$

Wavelet coefficients are mapped into Low Frequency (LF) and High Frequency (HF) feature streams using learnable temporal mappings:

$$F_L = \phi_L(LL), F_H = \phi_H(LH) \quad (8)$$

Here, F_L denotes global rhythm and baseline trends and F_H denotes emphasize rapid cardiac abnormalities. The wavelet-based feature representation is obtained by;

$$F_{wave} = \text{Concat}(F_L, F_H) \quad (9)$$

Token Alignment

To match the Swin-based temporal tokens, feature alignment is applied:

$$\hat{F}_{wave} = \text{Align}(F_{wave}) \quad (10)$$

Here, $\text{Align}(\cdot)$ represents consistency in token length and embedding dimension. Let $F_{swin} \in \mathbb{R}^{T \times D_s}$ denote Swin features. Query, key and value projections are defined as;

$$Q_s, K_s, V_s = \phi_s(F_{swin}), Q_w, K_w, V_w = \phi_w(\hat{F}_{wave}) \quad (11)$$

$$A_{w \rightarrow s} = \text{Softmax}\left(\frac{Q_s K_w^T}{\sqrt{d_k}}\right) V_w, A_{s \rightarrow w} = \text{Softmax}\left(\frac{Q_w K_s^T}{\sqrt{d_k}}\right) V_s \quad (12)$$

The attention enhanced features are fused back using;

$$F'_{swin} = F_{swin} + \alpha \pi_s(A_{w \rightarrow s}), \hat{F}'_{wave} = \hat{F}_{wave} + \beta \pi_w(A_{s \rightarrow w}) \quad (13)$$

Here, α and β stands learnable scalar weights that dynamically balance temporal and frequency domain contributions. The fused representation is obtained as:

$$F_{fused} = Conv_{fuse} \left(Concat(F'_{swin}, \hat{F}'_{wave}) \right) \quad (14)$$

Mixture of Experts State Space Model

A SSM layer captures the long-range temporal dynamics of ECG sequences with linear complexity, while a MoE layer replaces the feed forward Multi-Layer Perceptron (MLP) to allow adaptive learning of heterogeneous cardiac patterns. The single MoE-SSM block is defined as;

$$x_{l+1} = x_l + MoE \left(LN \left(x_l + SSM(LN(x_l)) \right) \right) \quad (15)$$

Here, x_l denotes unput ECG feature representation at layer l . SSM capture rhythm level temporal evolution, while MoE selectively activates experts' subnetworks specializing in various ECG structure.

EchoFormer

EchoFormer architecture utilizes the Echo State Network (ESN) dynamic modelling capabilities and the transformer model learning global dependencies, to effectively capture these properties of ECG Signals. An ESN consists of an input layer, a randomly initialized reservoir and a trainable output layer. The reservoir serves to project ECG signal inputs into a high-dimensional latent space allowing for the extraction of many different and highly temporal features. The evolution of states in the ESN reservoir is described as follows:

$$S_{t+1} = \lambda S_t + (1 - \lambda) \psi_a(s_t, x_t, y_t) \quad (16)$$

Here, $s_t \in \mathbb{R}^{N_s}$ denotes reservoir state, $x_t \in \mathbb{R}^{N_x}$ denotes ECG input time t , and $\psi_a(\cdot)$ stands nonlinear activation function. The activation function defined as,

$$\psi_a = (s_t, x_t, y_t) = \tanh(W_{res}s_t + W_{in}x_t + W_{back}y_t) \quad (17)$$

Here, W_{res} , W_{in} and W_{back} represents fixed randomly generated weight matrices. After a short free-running period, the ESN output is computed as:

$$\hat{y}_t = W_{out}[x_t, s_t, \hat{y}_{t-1}] + b_y \quad (18)$$

Here, W_{out} stands trainable parameter set. The output weights are optimized via;

$$\min_{W_{out}} \sum_{i=1}^N \frac{1}{2} \|y_i - \hat{y}_i\|_2^2 + \frac{\beta}{2} \|W_{out}\|_F^2 \quad (19)$$

Here, $\hat{y}_i$ denotes predicted output of EchoFormer, N stands number of training samples, β stands regularization parameter controlling model complexity. The reservoir stage passed through a transformer encoder with multi-head self-attention heads that model global temporal relationships between ECG cycle. The EchoFormer combined structure provides the ability to learn from both local cardiac rhythms and distant pathological relationships, and thus achieve more accurate classifications of CVD.

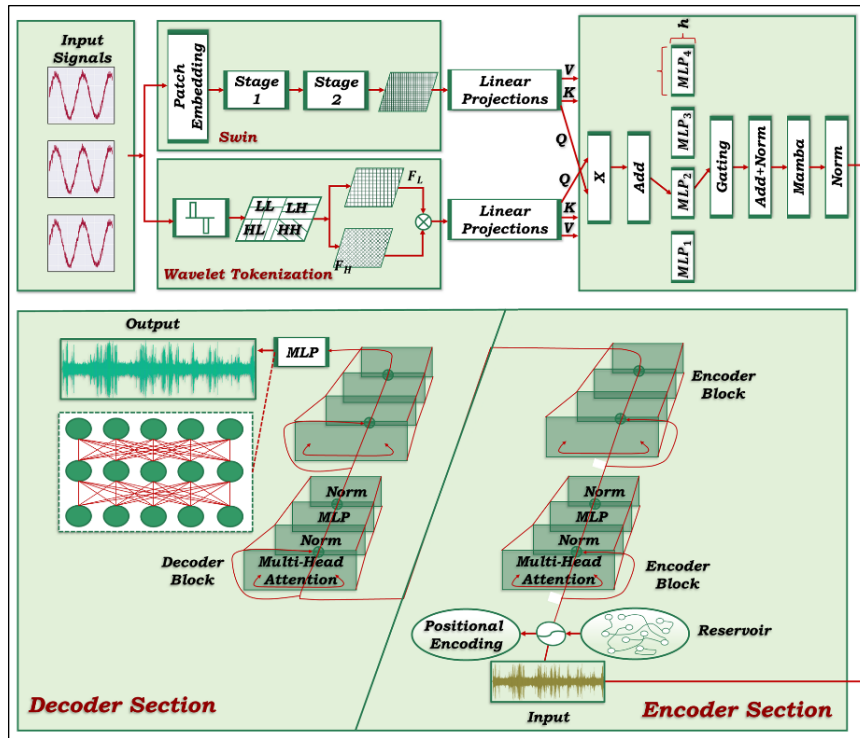


Figure 5: Architecture of WT-MoE-SSM-EchoFormer.

Source: Authors, (2026).

The WT-MoE-SSM-EchoFormer architecture as shown in Figure 5 combines wavelet decomposition with the principle of collaborative attention to represent temporal structures, through expert-guided selective learning and echo-state spaces as a means of generating a discriminative feature embedding for cardiac classification tasks. The final classifier efficiently distinguishing between normal and pathological cardiac activity without compromising the computational efficiency or stability of the predictor system.

III.4 GOLDEN TORTOISE BEETLE OPTIMIZER

GTBO is used to optimally tune the hyperparameters of the proposed WT-MoE-SSM-EchoFormer for accurate prediction of CVD as illustrates in Figure 6. It is a population-based metaheuristic based on the color-switching and survival processes of golden tortoise beetles. It provides a good level of exploration and exploitation for the complex optimization tasks involved in DL.

Solution Representation and Fitness Evaluation

Each beetle is a potential configuration of hyperparameters for the WT-MoE-SSM-EchoFormer, which include learning rate, number of attention heads, expert weightings for MoE, state-space parameters and regularisation weights. The population is represented by:

$$P_{beetle} = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,n} \end{pmatrix} \quad (20)$$

Here, P_{beetle} stands population matrix signifying all candidate GTBO solutions. m and n denotes population size and number of tunable WT-MoE-SSM-EchoFormer parameters. The fitness of each beetle is evaluated as,

$$F_{beetle} = \begin{pmatrix} f([b_{1,1}, b_{1,2}, \dots, b_{1,n}]) \\ f([b_{2,1}, b_{2,2}, \dots, b_{2,n}]) \\ \vdots \\ f([b_{m,1}, b_{m,2}, \dots, b_{m,n}]) \end{pmatrix} \quad (21)$$

Here, F_{beetle} represents fitness vector of all beetles.

Initialization

The initial value of k^{th} hyperparameter of i^{th} beetle at $G=1$ is given as,

$$x_{i,1}^k = x_{min}^k + rand(0,1)(x_{max}^k - x_{min}^k) \quad (22)$$

Here, $x_{i,1}^k$ denotes initial value of k^{th} hyperparameter for i^{th} beetle. x_{max}^k and x_{min}^k denotes upper and lower bounds. This initialization assures extensive coverage of WT-MoE-SSM-EchoFormer hyperparameter search space.

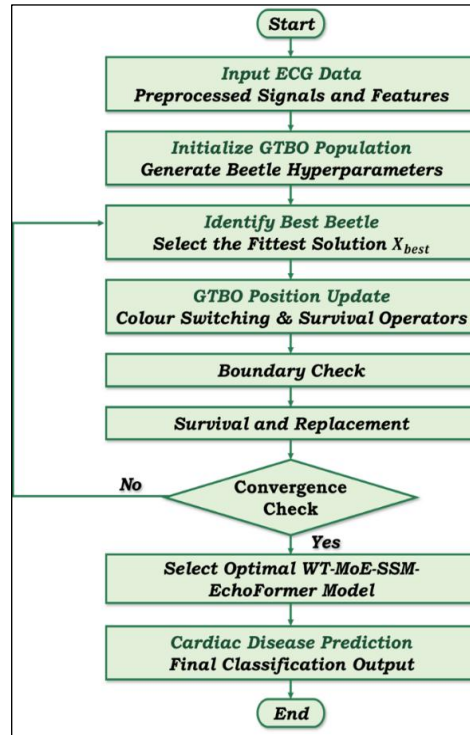


Figure 6: Flowchart for GTBO with WT-MoE-SSM-EchoFormer.
Source: Authors, (2026).

Color Switching Operator

The color switching operator updates the beetle position as,

$$V_t^G = X_t^G + S_{color}(X_{r1}^G - X_{best}^G) \quad (23)$$

Here, X_t^G stands current hyperparameter vector, X_{r1}^G stands randomly selected beetle, X_{best}^G denotes best performing configuration in generation G . S_{color} denotes color switching coefficient controlling exploration strength. These operators support the exploration of DL parameter space. The color switching coefficient is defined as,

$$S_{color} = (d_x n_x \cos \theta_x + d_y n_y \cos \theta_y) + h\lambda \quad (24)$$

With the dominant wavelength (λ) expressed as,

$$\lambda = \frac{2\alpha\bar{\phi}^2 - \sin^2(\theta_z)}{k} \quad (25)$$

The stochastic variables are generated as,

$$\begin{cases} d_x, n_x = \text{Randn}(), \\ d_y, n_y = \text{Randn}() \cdot \beta, \\ \bar{\phi} = \text{Cauchy}(\sigma, \mu), \\ \theta_x, \theta_y = \beta, \\ \alpha, k, h = \text{rand}(), \\ \theta_z = 2\pi \cdot \text{rand}() \end{cases} \quad (26)$$

Here, d_x and d_y stands directional displacement coefficient, n_x, n_y denotes normalized directional factor, $\theta_x, \theta_y, \theta_z$ represents random angular parameter. α and β stands random control parameter balancing exploration and exploitation. $\bar{\phi}$ stands random variable drawn from Cauchy distributions. σ and μ stands mean and scale parameters of Cauchy distributions. These randomized parameters prevent premature convergence and enhance training stability of WT-MoE-SSM-EchoFormer.

Survival Operator

The survival operator creates refined hyperparameter solution using,

$$\text{Beetle}_1 = \alpha x_{r1} + (1 - \alpha)(x_{r2} - \sigma_1) \quad (27)$$

$$\text{Beetle}_2 = \alpha x_{r2} + (1 - \alpha)(x_{r1} - \sigma_2) \quad (28)$$

Where,

$$\begin{cases} \sigma_1 = (1 - \delta)(x_{best} - x_{r1}), \\ \sigma_2 = (1 - \delta)(x_{best} - x_{r2}) \\ \delta = \alpha\eta|p|^\beta \end{cases} \quad (29)$$

Here, Beetle_1 and Beetle_2 stands offspring solutions generated through survival operators. σ_1 and σ_2 represents adaptive correction, δ stands survival rate, η and p denotes learning coefficient and random probability variables.

Table 1: Optimization details of GTBO.

Parameter	Value
Population Size	30
Maximum Iteration	50
Acceleration Coefficient	1.5
Random Factor	0.7
Spiral Probability	0.5
Convergence Coefficient	0.9
Exploration rate	0.6
Step size Factor	0.05
Inertia Weight	0.8
Escape radius/ threshold	0.3
Batch Size	32
Epoch	10

Source: Authors, (2026).

Table 1 signifying optimization parameter specifications. By combining GTBO and WT-MoE-SSM-EchoFormer attaining optimal hyperparameter tuning, faster convergence and increased generalization resulting in accurate and precise ECG based CVD prediction.

IV. RESULT AND DISCUSSION

This section presents experimental results and performance evaluation of the proposed WT-MoE-SSM-EchoFormer model using the GTBO. The experiments are performed in Python using MIT-BIH Arrhythmia Database in order to evaluate classification accuracy, robustness, and discriminative ability across multiple classes of ECGs.

Dataset details

Dataset Name: MIT-BIH Arrhythmia Database
 Dataset Type: ECG Time-Series Signal
 Sampling Frequency: 360 Hz
 Recording Duration: Approximately 30 minutes per subject
 Number of Subjects: 47 patients (48 ECG recordings)
 Class Names: N – Normal beat, V – Ventricular ectopic beat, F – Fusion beat, Q – Unknown beat
 Train & Test Split: 80:20

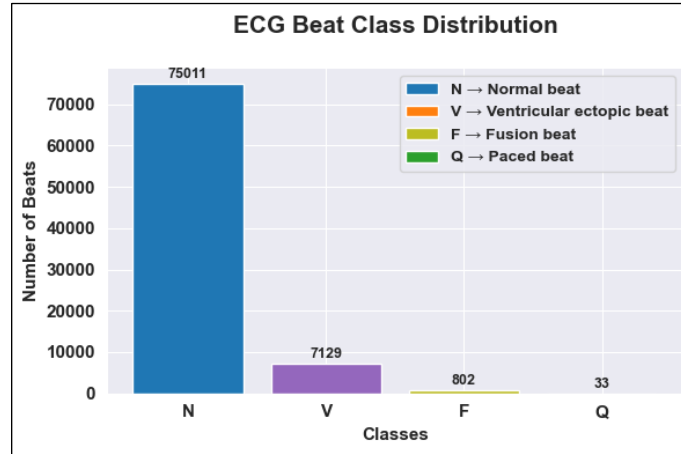


Figure 7: ECG beat class distribution in the MIT-BIH Arrhythmia dataset.
Source: Authors, (2026).

Figure 7 represents class wise distribution of ECG beats in MIT-BIH Arrhythmia dataset. The dataset is highly imbalanced with Normal (N) beats forming the majority of 77,011 beats followed by Ventricular ectopic (V) with 7,129 beats. Fusion (F) class contains fewer samples of 802 beats and Unknown beat (Q) is severely underrepresented with 33 beats. This balance emphasizes the difficulty of developing effective classifiers for minority arrhythmia classes and supports the use of improved feature representation and adaptive learning processes in CVD prediction.

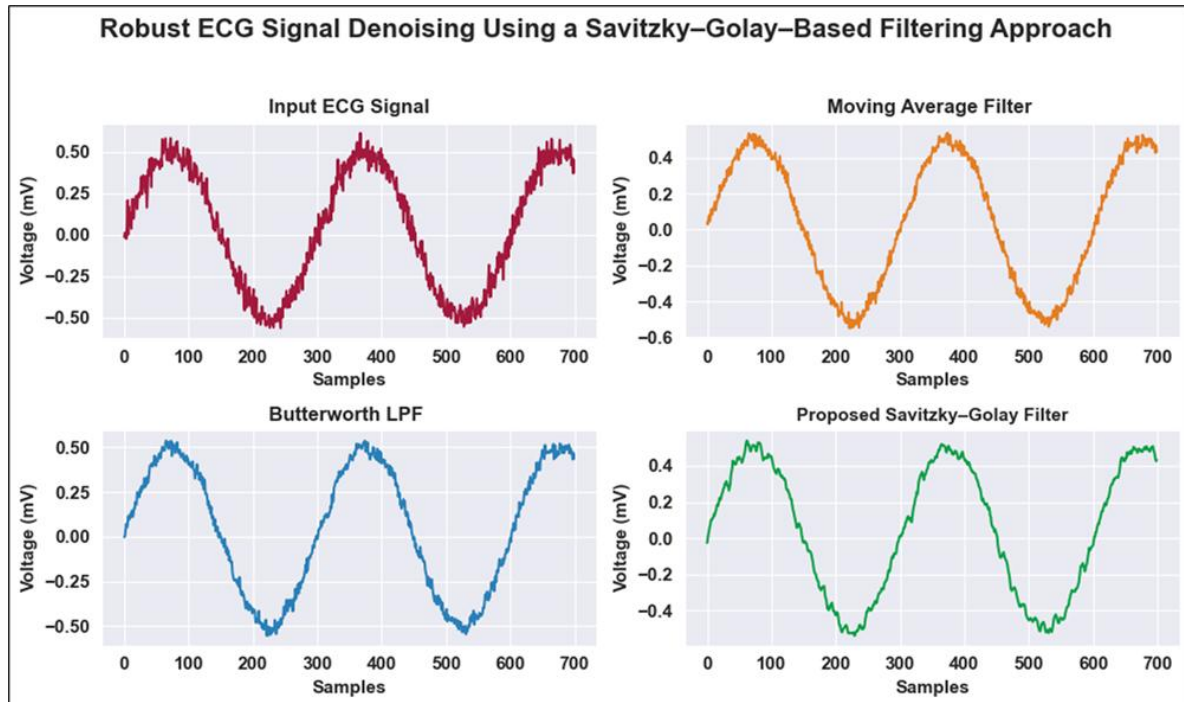


Figure 8: Comparative evaluation of ECG signal denoising using different filtering techniques.
Source: Authors, (2026).

Figure 8 compares different technique for denoising ECG signals. The raw ECG signal is HF noise and baseline wandering, making it difficult to understand the main features of the ECG. Applying a moving average filter to the signal help to reduce some of the noise, but the shape of the signal is greatly distorted, and reduced the important parts of the ECG components such as QRS peaks. The Butterworth Low Pass Filter (LPF) filter supresses the HF noise very well, but it slightly blurs the QRS complex edges, due to phase smoothing, which affect the accuracy of feature extraction.

However, the proposed method, a SG filter denoises the signal and keeps the main ECG features intact such as the sharpness of the peaks and the smoothness of the waveform. Thus, the proposed method is capable of accurately identifying QRS complexes and reliably measuring temporal intervals, thereby making it very effective for cardiac analysis and disease prediction purposes.

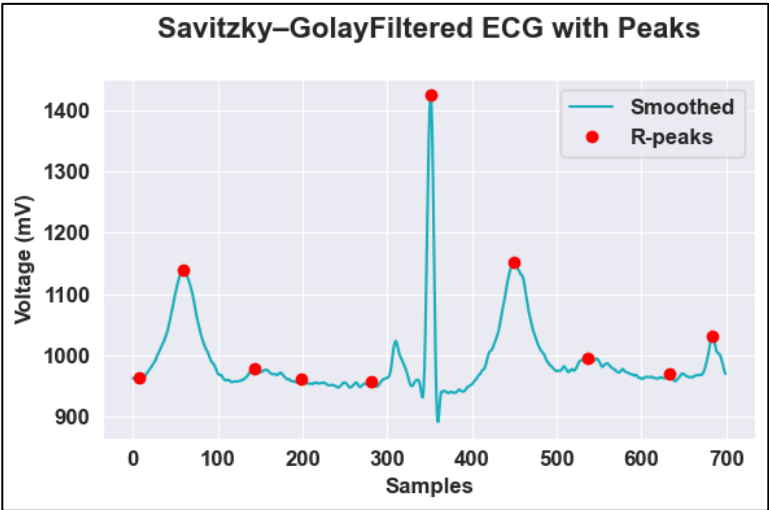


Figure 9: R-Peak detection on SG filtered ECG signal.
Source: Authors, (2026).

Figure 9 illustrates the detection of R-peaks on an ECG signal after denoising using SG filter. The smoothed ECG waveform preserves essential morphological characteristics while effectively decreasing HF noise. The observed R-peaks denoted by red dots, correspond to QRS complex maxima and serve as crucial reference points for calculating RR intervals, Heart Rate Variability (HRV) and temporal characteristics. Accurate R-peak localization is required for subsequent ECG feature extraction such as QRS duration and interval-based analysis which directly contributes to accurate cardiac disease categorization.

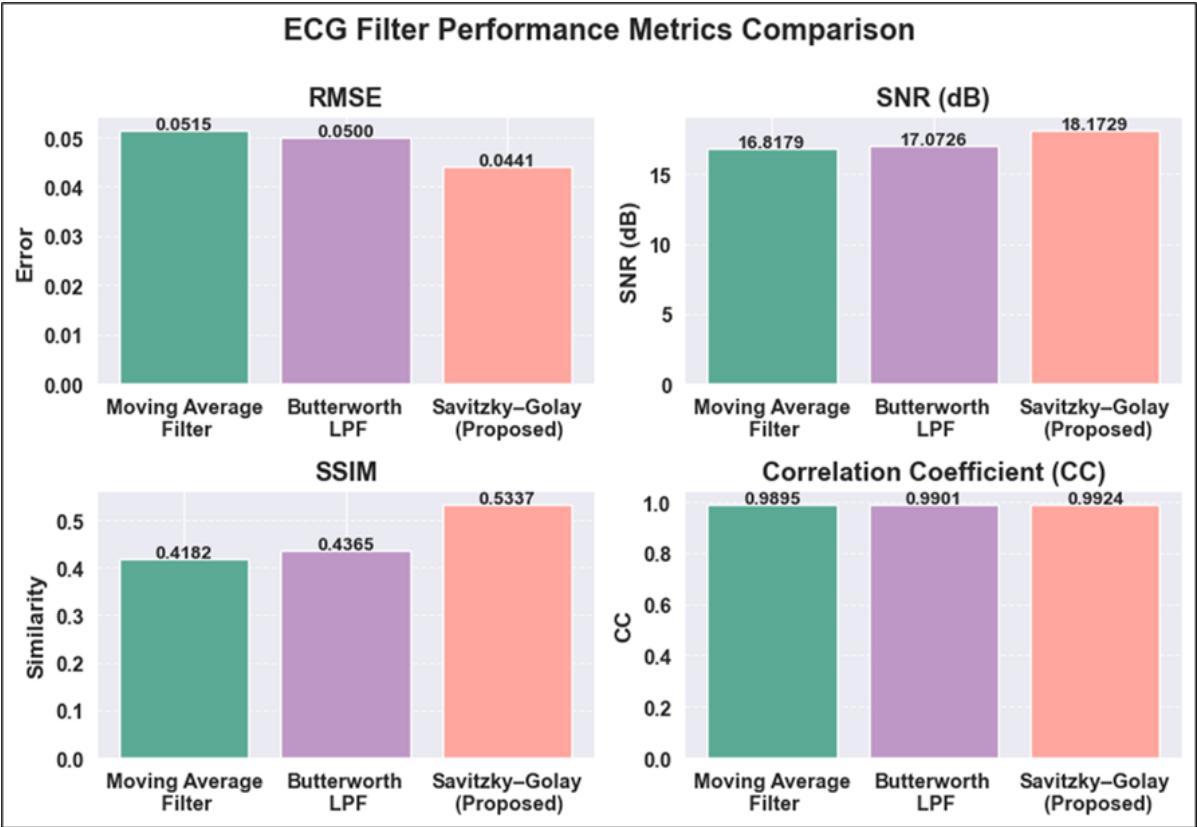


Figure 10: Performance comparison of ECG denoising filter.
Source: Authors, (2026).

Performance comparison shows Figure 10, that the proposed SG filter consistently beats traditional approaches on all evaluation metrics. It has the lowest RMSE of 0.0441 and greatest SNR of 18.1729 dB, suggesting excellent noise suppression and signal reconstruction accuracy. Furthermore, the proposed model has an SSIM of 0.5337 and correlation coefficient of 0.9924, signifying improved structure preservation and high similarity to the original ECG signal. These findings clearly show the proposed approach as the most effective ECG denoising method for accurate cardiac signal interpretation.

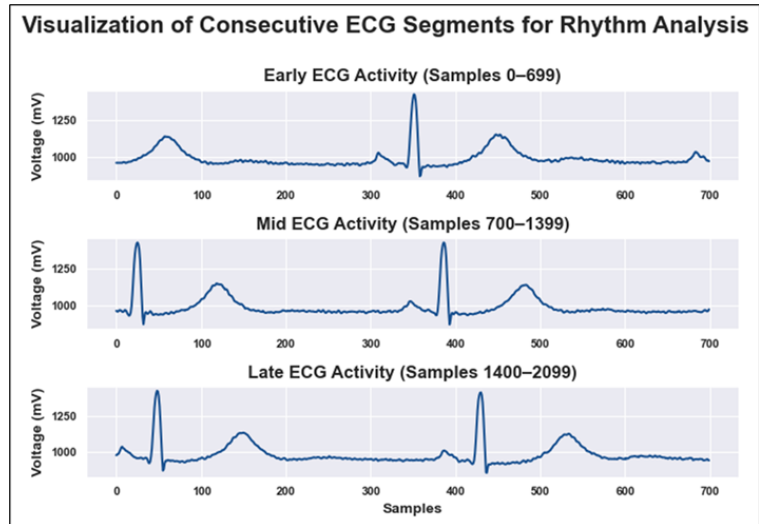


Figure 11: Visualization of consecutive ECG segments for Rhythm Analysis.
Source: Authors, (2026).

Figure 11 illustrates three consecutive ECG segments showing early, mid and late cardiac activity obtained from continuous recordings. Each section shows distinct P-QRS-T wave patterns allowing for visual analysis of rhythm consistency and beat morphology across time. The segmentation implies for effective temporal analysis of cardiac dynamics and consistent feature extraction for ECG classification and arrhythmia diagnosis.

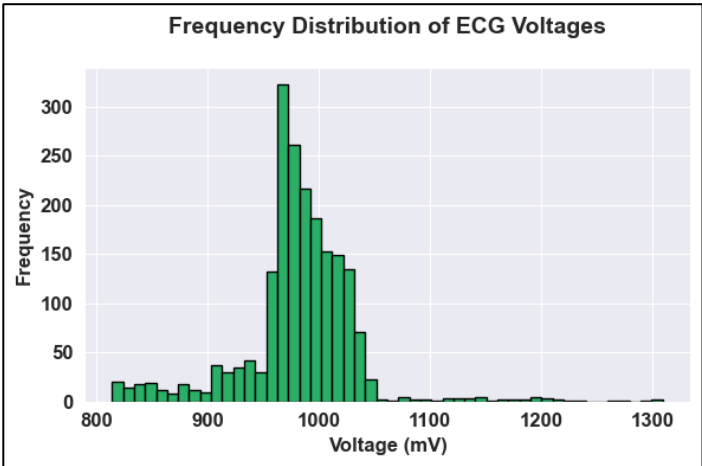


Figure 12: Frequency distribution of ECG voltages.
Source: Authors, (2026).

Figure 12 illustrates histogram of ECG voltage amplitudes, showing the distributions of signal values across the recordings. A high level within the mean voltage implies consistent heart activity, while the spread indicates physiological variability and the existence of notable waveform component such as QRS complex.

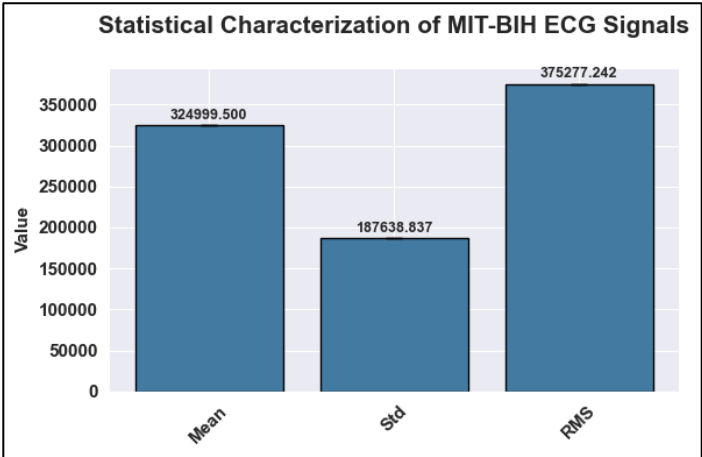


Figure 13: Statistical characteristics of MIT-BIH ECG signals.
Source: Authors, (2026).

Figure 13 summarizes key statistical measures of the ECG signals, including mean, standard deviation and Root Mean Square (RMS). The ECG signals exhibit a mean value of 324,999.5, a standard deviation of 187638.837 and RMS value of 375277.242, reflecting moderate variability and high signal energy suitable for accurate cardiac analysis.

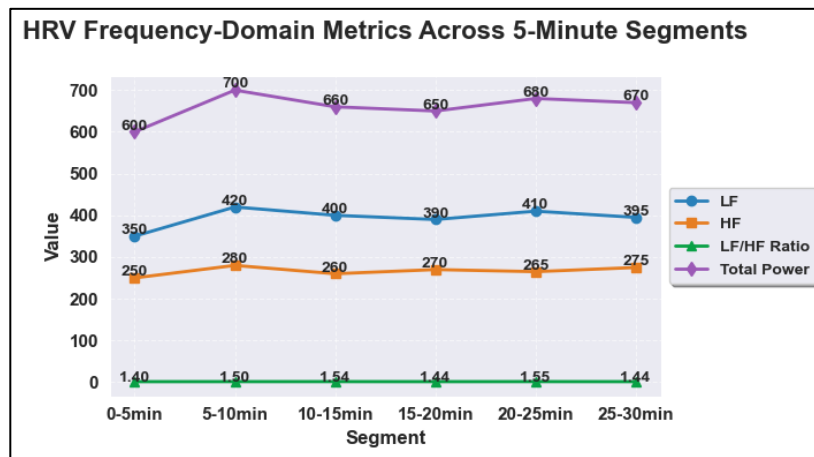


Figure 14: HRV frequency-domain metrics across 5-minute segments.

Source: Authors, (2026).

Figure 14 represents variation of HRV frequency domain parameters such as LF, HF, LF/HF ratio, and total power across successive 5-minute segments. Across consecutive 5-minutes intervals, LF varies between 350 and 420, HF ranges from 250 to 280, the LF/HF ratio remains stable at 1.40-1.55 and total power fluctuates between 600 and 700, signifying consistent automatic regulation over time.

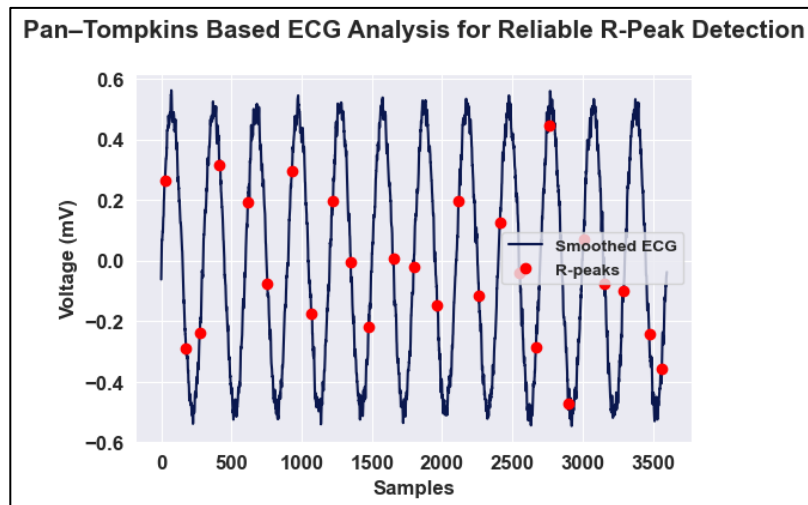


Figure 15: Pan-Tompkins based ECG analysis.

Source: Authors, (2026).

Table 2: Extracted ECG features using Pan-Tompkin's algorithm.

R-Peak Index	RR-Interval (s)	Heart Rate (bpm)	QRS Duration (s)
177	0.417	144.0	0.067
275	0.272	220.4	0.069
415	0.389	154.3	0.128
617	0.561	106.9	0.108
756	0.386	155.4	0.089
931	0.486	123.4	0.131
1067	0.378	158.8	0.094
1223	0.433	138.5	0.097
1348	0.347	172.8	0.092
1476	0.356	168.8	0.083
1653	0.492	122.0	0.106
1796	0.397	151.0	0.100
1966	0.472	127.1	0.089

2117	0.419	143.0	0.108
2261	0.400	150.0	0.094
2412	0.419	143.0	0.144
2551	0.386	155.4	0.083
2666	0.319	187.8	0.086
2764	0.272	220.4	0.164
2901	0.381	157.7	0.067
3006	0.292	205.7	0.103
3154	0.411	145.9	0.094
3291	0.381	157.7	0.086
3473	0.506	118.7	0.089
3563	0.250	240.0	0.089

Source: Authors, (2026).

Figure 15 represents the smoothed ECG waveform processed using Pan-Tompkins algorithm. The detected R-peaks are highlighted on the signal, signifying accurate localization of QRS complexes across successive cardiac cycles, which is crucial for strong heartbeat detection and rhythm analysis. Table 2 represents numerical features extracted from the detected R-peaks including R-Peaks index, RR intervals, heart rate and QRS duration. These features provide quantitative information of cardiac activity and utilized for further analysis and classification of cardiac conditions.

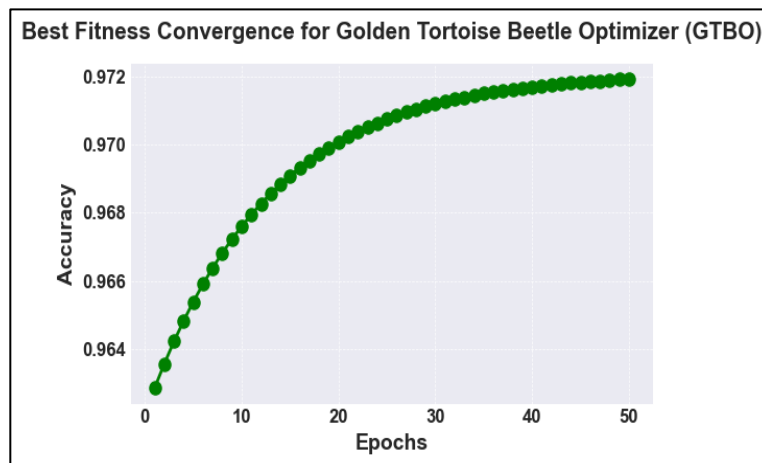


Figure 16: Best fitness convergence curve of GTBO.

Source: Authors, (2026).

Figure 16 shows convergence process of the GTBO with the number of epochs. The smooth saturation of the accuracy towards the optimal value shows the good exploration-exploitation trade-off and the strong ability of the optimizer to reach high-quality solutions.

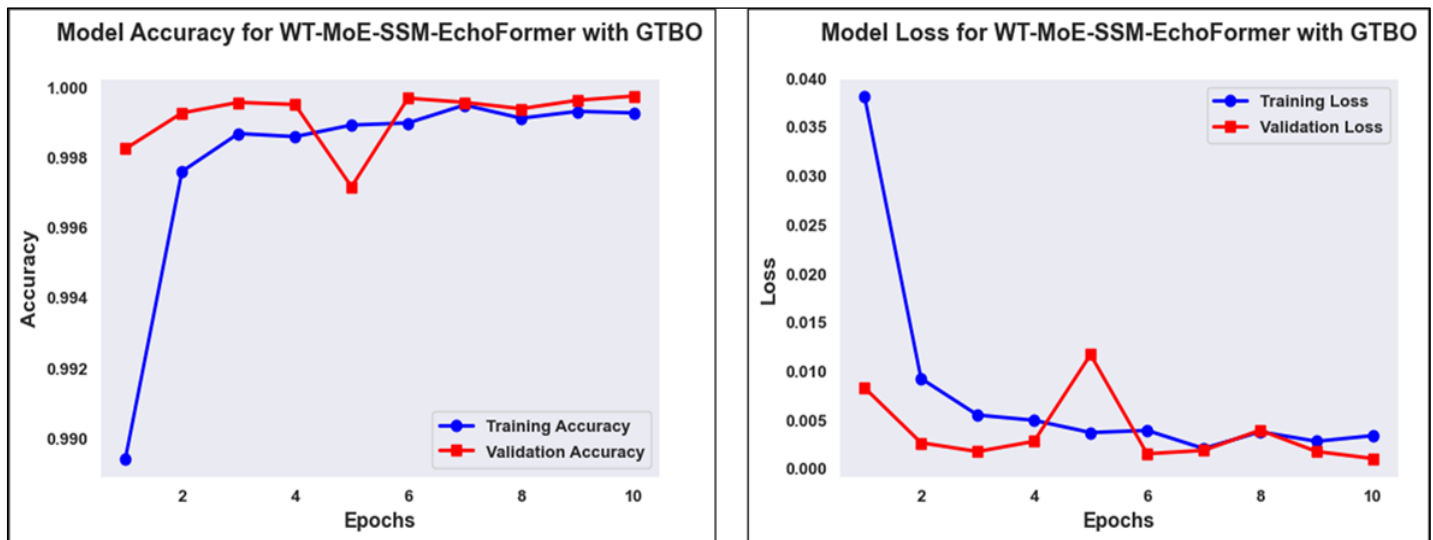


Figure 17: Model accuracy and loss curves for WT-MoE-SSM-EchoFormer with GTBO.

Source: Authors, (2026).

Figure 17 shows training and validation performance of the WT-MoE-SSM-EchoFormer model optimized by GTBO. It shows that the training and validation accuracy of model increases very quickly and converges to values above 99% and displays loss of the model, which converges to very low values.

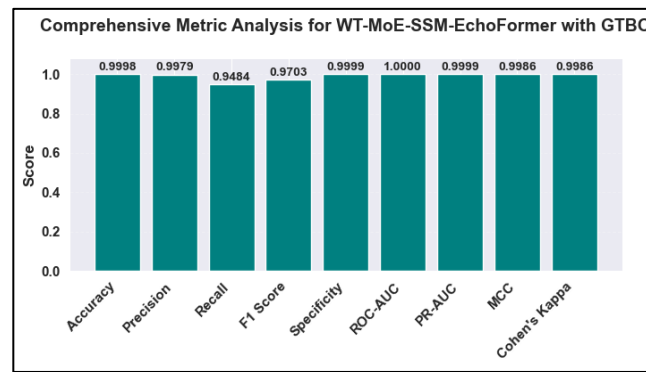


Figure 18: Performance metrics for proposed model.

Source: Authors, (2026).

Figure 18 illustrates a comprehensive comparison of various performance metrics obtained by the proposed WT-MoE-SSM-EchoFormer model optimized GTBO. The WT-MoE-SSM-EchoFormer model optimized by GTBO obtains an accuracy of 0.9998 and precision of 0.9979, signifying highly positive class predictions. Recall of 0.9484 and F1-score of 0.9703 also confirm a well-balanced sensitivity and classification stability. The model obtains specificity of 0.9999, ROC-AUC of 1.0000, and PR-AUC of 0.9999, signifying a highly dependable discriminative power. MCC of 0.9986 and Cohen's Kappa of 0.9986, signifying the excellence of the proposed method.

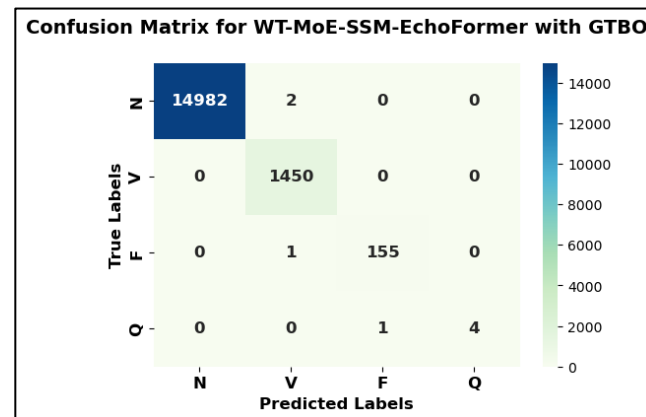


Figure 19: Proposed model confusion matrix structure.

Source: Authors, (2026).

Figure 19 shows confusion matrix that indicates classification performance of the WT-MoE-SSM-EchoFormer model optimized by GTBO on ECG beat classes (N, V, F, Q). The model has been able to correctly classify most of the beats, obtaining 14,982 correct classifications for Normal (N) beats, 1,450 for Ventricular ectopic (V) beats, 155 for Fusion (F) beats, and 4 for Unknown (Q) beats, with only a negligible number of misclassifications. The high diagonal dominance of confusion matrix indicates high accuracy and robustness of the proposed model.

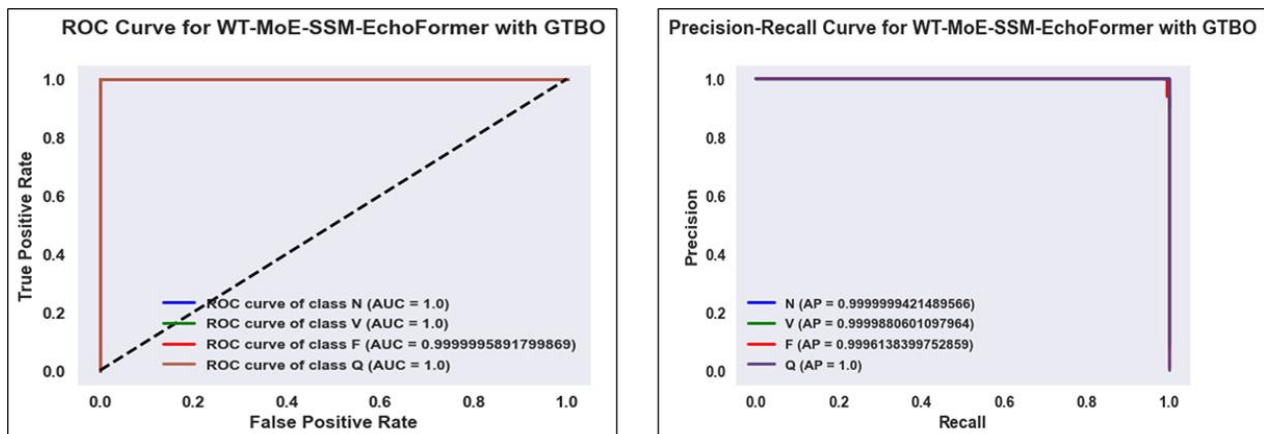


Figure 20: ROC and Precision-Recall Curves for proposed model.

Source: Authors, (2026).

Figure 20 shows performance of the classification task of the proposed WT-MoE-SSM-EchoFormer model optimized with GTBO on the ROC and Precision-Recall curves for all classes of ECG signals (N, V, F, and Q). The ROC curves clearly show the near-perfect discrimination power of the model with an AUC value of 1.0 for all classes, which is a clear indication of the excellent trade-off and Precision-Recall curves show an excellent average precision value of 1.0 for all classes.

Existing Approaches:

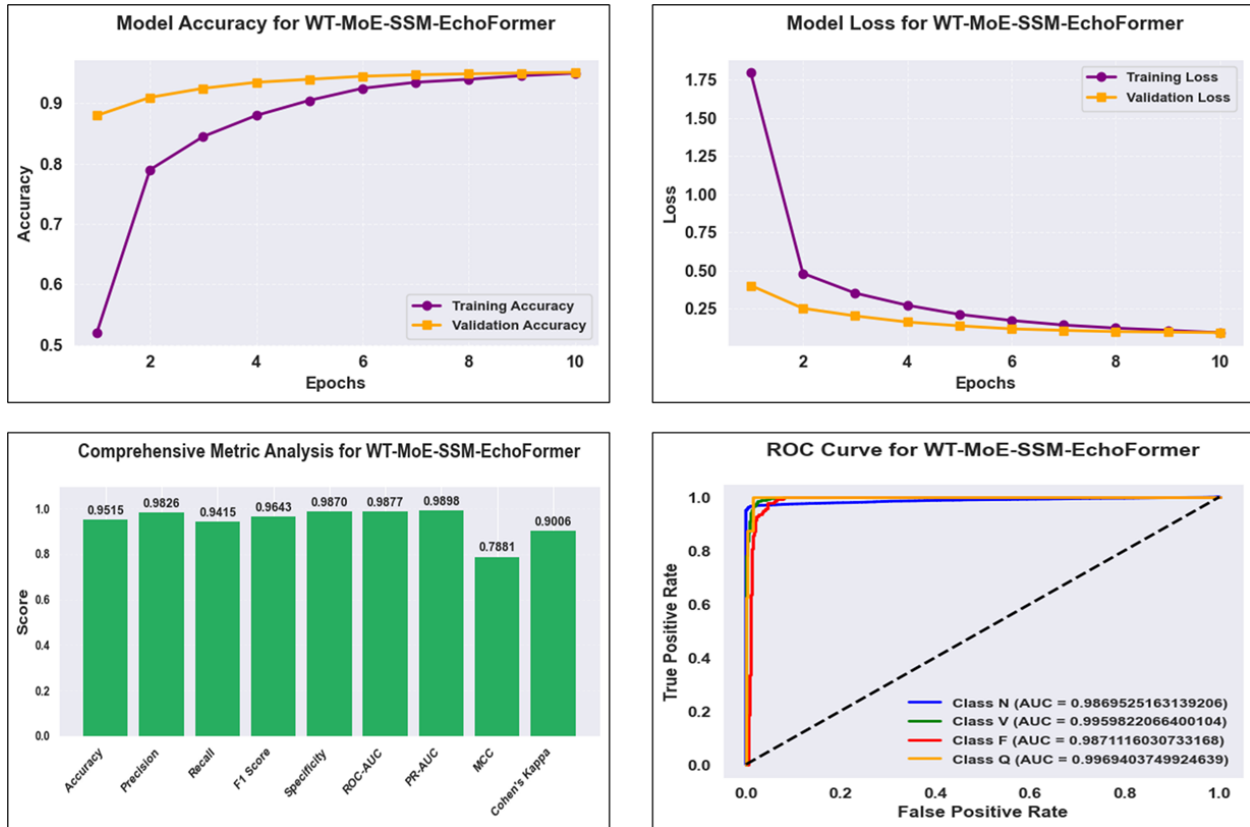


Figure 21: Performance analysis of existing WT-MoE-SSM-EchoFormer model.

Source: Authors, (2026).

Figure 21 illustrates performance analysis of existing WT-MoE-SSM-EchoFormer model. Training and validation accuracy and loss of existing WT-MoE-SSM-EchoFormer model, along with its overall classification metrics. The accuracy plots indicate a steady improvement in the model with convergence at 95%, whereas the loss plots indicate a gradual decrease, signifying a good learning process. The overall classification metrics obtains an accuracy of 0.9515, precision of 0.9826, recall of 0.9415, F1-score of 0.9643, specificity of 0.9870, ROC-AUC of 0.9877, PR-AUC of 0.9898, MCC of 0.7881 and Cohen's Kappa of 0.9006, clearly indicate that the existing model performs well on the classification task. However, when compared to the proposed GTBO-optimized WT-MoE-SSM-EchoFormer model, the proposed model achieves better convergence, accuracy, and generalization. The ROC analysis provides a measure of the performance of the classification task, obtaining AUC values of 0.98695 for class N, 0.99598 for class V, 0.98711 for class F, and 0.99694 for class Q. While these values clearly indicate a high predictive capability, they are relatively lower compared to the ones obtained by the proposed GTBO-optimized WT-MoE-SSM-EchoFormer model.

Confusion Matrix for WT-MoE-SSM-EchoFormer				
True Labels	N	V	F	Q
N	14302	168	439	99
V	3	1350	78	8
F	1	7	131	1
Q	0	0	1	7
Predicted Labels				
	N	V	F	Q

Figure 22: Confusion matrix for WT-MoE-SSM-EchoFormer.

Source: Authors, (2026).

Figure 22 shows the performance of the existing WT-MoE-SSM-EchoFormer model based on confusion matrix. The model is able to correctly classify 14,302 samples of class N, 1,350 samples of class V, 131 samples of class F, and 7 samples of class Q. However, there are some samples that are misclassified, such as 439 samples of class N to F, 168 samples of N to V, and 78 samples of V to F.

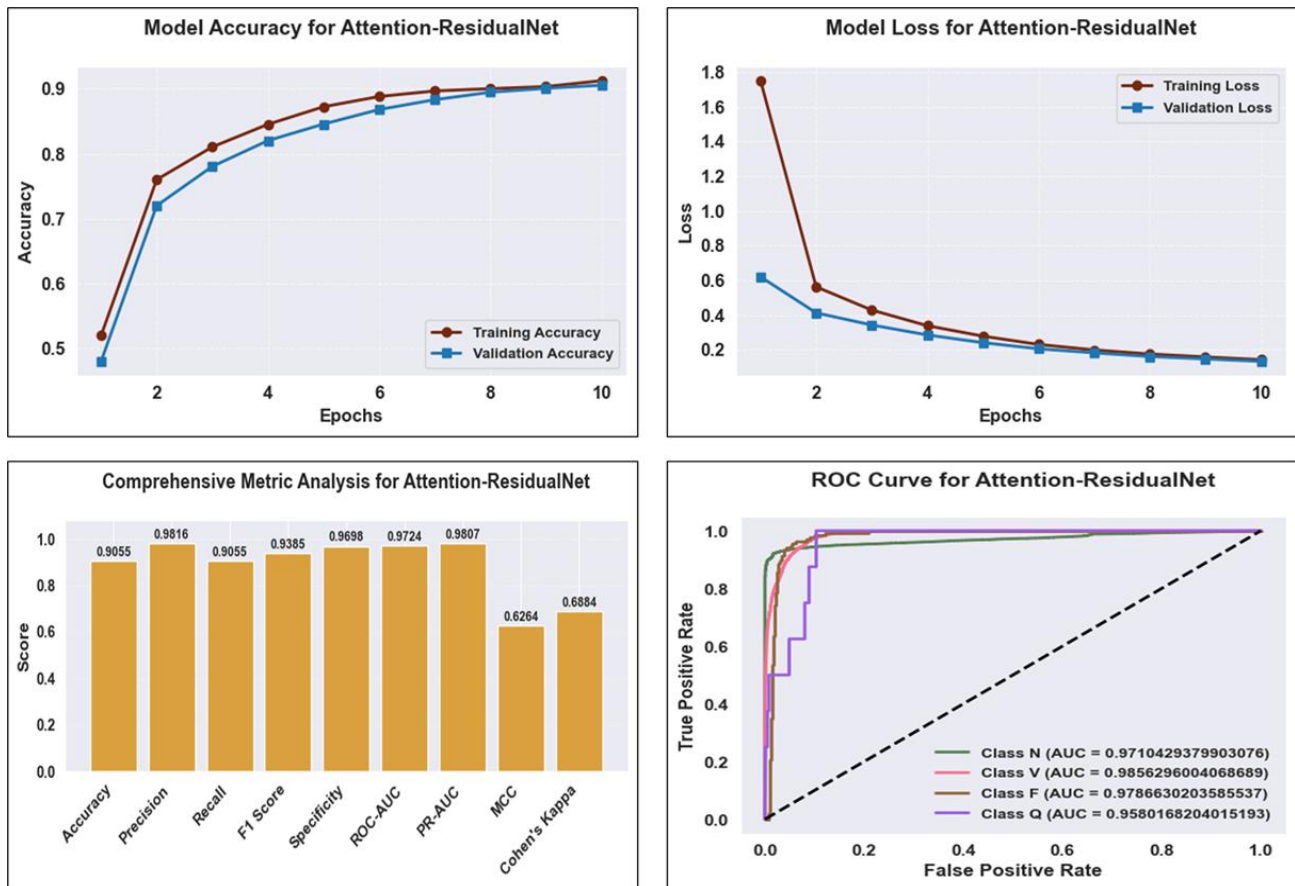


Figure 23: Performance evaluation of Attention ResidualNet Model.

Source: Authors, (2026).

Figure 23 represents training and validation performance of Attention-ResidualNet in terms of accuracy, loss, overall metrics, and ROC analysis. The model has an overall accuracy of 0.9055, precision of 0.9816, recall of 0.9055, and F1-score of 0.9385, which is well balanced. The ROC analysis also shows high discriminative power with AUC values of 0.9710 (Class N), 0.9856 (Class V), 0.9787 (Class F), and 0.9580 (Class Q). However, these values are lower compared to the proposed GTBO-optimized model, which proves that the proposed model is better in terms of accuracy, robustness, and class separability.

Confusion Matrix for Attention-ResidualNet				
True Labels	N	V	F	Q
	13887	94	554	473
	42	1003	207	187
	2	6	132	0
	0	0	4	4
Predicted Labels				

Figure 24: Confusion matrix of Attention ResidualNet.

Source: Authors, (2026).

Confusion matrix of the Attention-ResidualNet on four classes (N, V, F, and Q) is shown in Figure 24. The Attention-ResidualNet correctly classifies 13,887 samples of class N, 1,003 samples of class V, 132 samples of class F, and 4 samples of class Q, which shows fair classification of most classes.

However, there are some noticeable misclassifications, such as class N vs. F (554 samples) and Q (473 samples), and class V vs. F (207 samples) and Q (187 samples). These misclassifications show the limitations of class discrimination, particularly for minority classes, while the proposed GTBO-optimized model performs better in classification accuracy and misclassification.

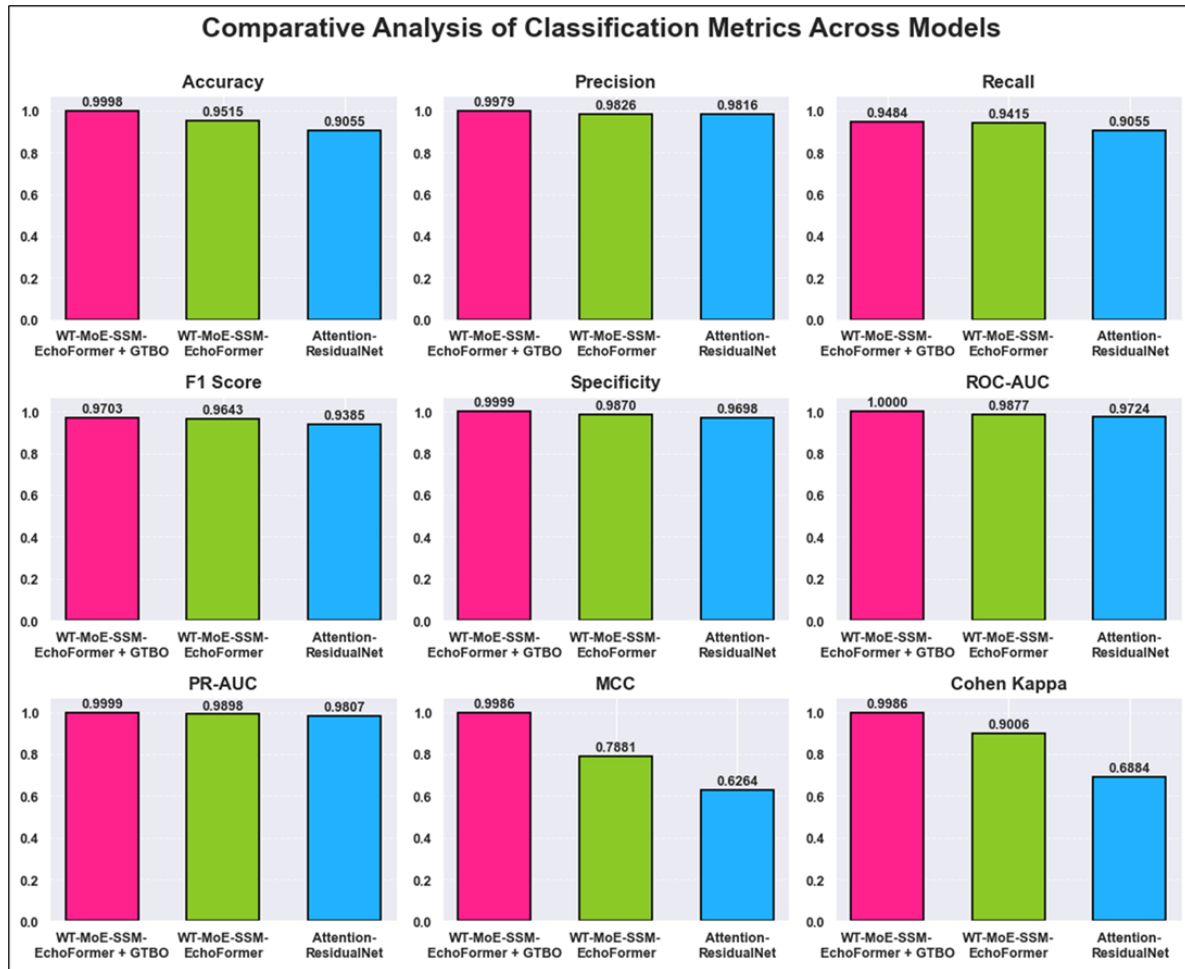


Figure 25: Comparative analysis of classification metrics across models.
Source: Authors, (2026).

Figure 25 compares the performance of three models, namely WT-MoE-SSM-EchoFormer with GTBO, WT-MoE-SSM-EchoFormer, and Attention-ResidualNet, on various evaluation metrics. The proposed WT-MoE-SSM-EchoFormer with GTBO outperforms all other models on all evaluation metrics, including accuracy of 0.9998, precision of 0.9979, recall of 0.9484, F1-Score of 0.9703, specificity of 0.9999, ROC-AUC of 1.0000, PR-AUC of 0.9999, MCC of 0.9986, and Cohen's Kappa of 0.9986. In contrast, the baseline WT-MoE-SSM-EchoFormer and Attention-ResidualNet perform relatively poorly on most evaluation metrics.

Table 3: Evaluation of accuracy across different previous literature.

REF	Methods	Dataset used	Accuracy
[11]	XGBoost	PTB-XL dataset	88.44%
[12]	Extra-Trees	Open-source dataset for heart disease prediction	95.48%
[13]	ANN	Heart disease dataset	92.50%
[14]	Rhythm-CNN	MIT-BIH + BIDMC dataset	98.52%
[15]	VGG16	ECG Images dataset of Cardiac and COVID-19 Patients	95%
[16]	Google ViT Resnet50	ECG Images of Cardiac and COVID-19 patients	98%
[17]	HybDeepNet	MIT-BIH	98.6%
[18]	Swin	ECG images dataset	95.5%
[19]	TransConvNet	MIT-BIH arrhythmia	98.6%
[20]	Self-Attention Transformer	Cleveland dataset	96.51%
Ours	PROPOSED	MIT-BIH Arrhythmia Database	99.98%

Source: Authors, (2026).

Table 3 indicates comparative accuracy analysis of existing ECG based cardiac disease prediction methods reported in the literature. The GTBO with WT-MoE-SSM-EchoFormer model, which is tested on the MIT-BIH Arrhythmia Database, outperforms all the other methods in terms of classification accuracy and attains a value of 99.98%.

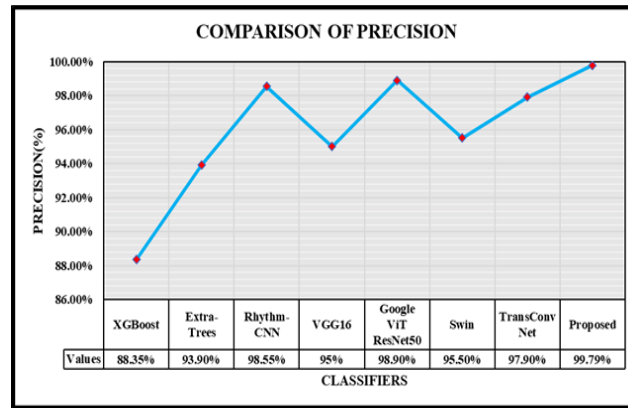


Figure 26: Evaluation of precision.
Source: Authors, (2026).

Figure 26 evaluate precision of various classifiers. XGBoost has 88.35%, and Extra-Trees increase to 93.90%. Rhythm-CNN and Google ViT-ResNet50 have a precision of 98.55% and 98.90%, respectively, whereas VGG16 and Swin have a precision of 95.0% and 95.50%, respectively. TransConvNet has a precision of 97.90%, and the proposed model has the maximum precision of 99.79%.

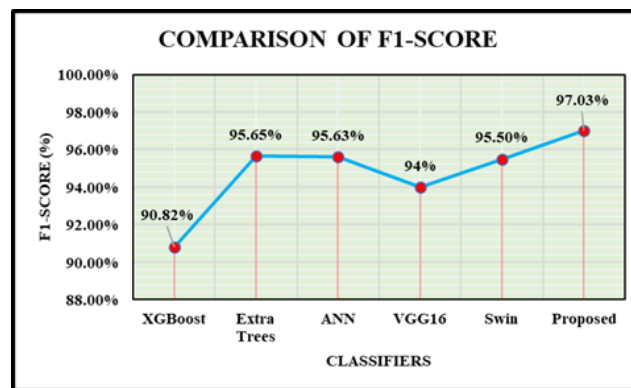


Figure 27: Evaluation of F1-Score.
Source: Authors, (2026).

Figure 27 represents F1-score of each classifier, where XGBoost attained 90.82%, followed by Extra-Trees of 95.65% and ANN of 95.63%. The F1-score for VGG16 is slightly lower at 94.0%, Swin attain to 95.50% and the proposed model achieves highest F1-score of 97.03%.

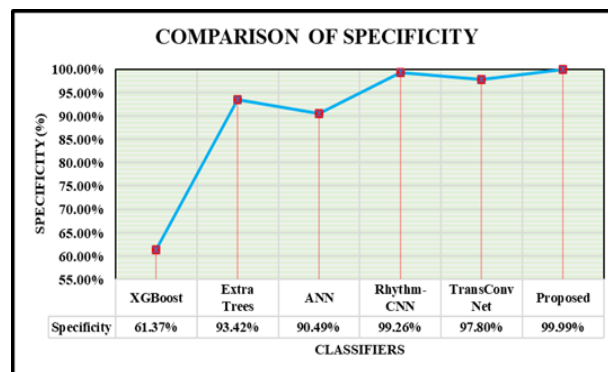


Figure 28: Evaluation of specificity.
Source: Authors, (2026).

Figure 28 compares specificity of each classifier in the existing literature. The specificity of XGBoost is low at 61.37%, while that of Extra-Trees and ANN is 93.42% and 90.49%, respectively. The specificity of Rhythm-CNN and TransConvNet is attained 99.26% and 97.80%, respectively. The specificity of the proposed model is maximum value of 99.99%, which proves its effectiveness in identifying negative cases correctly.

V. CONCLUSION

In this work, a new framework of ECG-based cardiac disease strong and intelligent approach that combines GTBO with WT-MoE-SSM-EchoFormer Architecture. With the aid of SG filtering technique to denoise ECG signals without losing essential morphological aspects such as QRS complexes and RR intervals crucial in the clinical analysis of ECGs. Using EDA including statistical, temporal, and morphological pattern achieved effective ECG signals analysing which leading to a highly informative signal representation. Clinical features of heart rhythm dynamics such as RR interval, QRS duration, HRV extracted accurately using the Pan-Tompkin's algorithm.

The hybrid WT-MoE-SSM-EchoFormer integrates the ability to identify and extract features from multiple resolutions, adaptive learning capabilities of multiple experts, and model Temporal information over long-term periods, resulting in accurate classifications of ECG-based cardiac diseases. Hyperparameter tuning performed with the aid of the GTBO achieved highly stable training process, faster training epochs and improved generalization. Results from extensive testing by Python software with MIT-BIH Arrhythmia Database yielded a classification accuracy of 99.98%, precision of 99.79%, recall of 94.84%, F1 score of 97.03%, specificity of 99.99%, and a ROC-AUC of 1.0000. Thus, the proposed framework demonstrates a very high level of discrimination across all ECG type classes.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: S. Nagakumararaj, P. Archana Devi, D. Justin Jose, S. Sree Vidhya.

Methodology: S. Nagakumararaj, P. Archana Devi.

Investigation: S. Nagakumararaj, P. Archana Devi.

Discussion of results: S. Nagakumararaj, P. Archana Devi, D. Justin Jose, S. Sree Vidhya.

Writing – Original Draft: S. Nagakumararaj.

Writing – Review and Editing: S. Nagakumararaj, P. Archana Devi.

Resources: D. Justin Jose, S. Sree Vidhya.

Supervision: D. Justin Jose, S. Sree Vidhya.

Approval of the final text: S. Nagakumararaj, P. Archana Devi, D. Justin Jose, S. Sree Vidhya.

VII. REFERENCES

- [1] Kumar, K. Rathor, S. Vaddi, D. Patel, P. Vanjarapu, M. Maddi, "ECG based early heart attack prediction using neural networks", In 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 1080-1083. IEEE, 2022.
- [2] G. N. Ahmad, H. Fatima, S. Ullah, A. S. Saidi, "Efficient medical diagnosis of human heart diseases using machine learning techniques with and without GridSearchCV", IEEE Access, vol. 10, pp. 80151-80173, 2022.
- [3] S. Karimulla, D. Patra, "An optimal methodology for early prediction of sudden cardiac death using advanced heart rate variability features of ECG signal", Arabian Journal for Science and Engineering, vol. 49, no. 5, pp. 6725-6741, 2024.
- [4] V. Thamilarasi, R. Roselin, P. Pushpa, M. Kannan, B. S. Vignesh, "Intelligent Deep Learning Algorithms for Autism Spectrum Disorder Diagnosis", In Computational Intelligence Algorithms for the Diagnosis of Neurological Disorders, pp. 109-123. CRC Press.
- [5] C. H. Lee, S. H. Kim, "ECG measurement system for vehicle implementation and heart disease classification using machine learning", IEEE Access, vol. 11, pp. 17968-17982, 2023.
- [6] A. L. Golande, T. Pavankumar, "Optical electrocardiogram based heart disease prediction using hybrid deep learning", Journal of Big Data 10, no. 1, pp. 139, 2023.
- [7] R. Anand, A. Rath, P. K. Sahoo, P. Jain, G. Panda, X. Wang, H. Liu, "Winograd Transform Based Fast Detection of Heart Disease using ECG Signals and Chest X-Ray Images", IEEE Access, 2025.
- [8] D. Karthikeyan, R. Nivasini, K. T. Raja, D. Madhivadham, B. Sivakumar, A. K. Priya, , "Hybrid Deep Learning Algorithm For MRI Brain Alzheimer's Disease Prediction", In 2023 International Conference on Energy, Materials and Communication Engineering (ICEMCE), pp. 1-6. IEEE, 2023.
- [9] P. Pachiyannan, M. Alsulami, D. Alsadie, A. K. J. Saudagar, M. AlKhatami, R. C. Poonia, "A novel machine learning-based prediction method for early detection and diagnosis of congenital heart disease using ECG signal processing", Technologies, vol. 12, no. 1, pp. 4, 2024.
- [10] U. Sumalatha, K. K. Prakasha, S. Prabhu, V. C. Nayak, "Deep learning applications in ecg analysis and disease detection: An investigation study of recent advances", IEEE Access, 2024.
- [11] S. Karthik, M. Santhosh, M. S. Kavitha, A. C. Paul, "Automated Deep Learning Based Cardiovascular Disease Diagnosis Using ECG Signals", Computer Systems Science & Engineering, vol. 42, no. 1, 2022.
- [12] M. S. I. Sumon, M. S. B. Islam, M. S. Rahman, M. S. A. Hossain, A. Khandakar, A. Hasan, M. Murugappan, M. E. Chowdhury, "CardioTabNet: a novel hybrid transformer model for heart disease prediction using tabular medical data", Health Information Science and Systems, vol. 13, no. 1, pp. 44, 2025.
- [13] A. T. Maadapoosi, V. Balamurugan, V. Vedanarayanan, S. A. Nisha, R. Narmadha, "Grey Wolf Optimization-Based Artificial Neural Network in the Development of an Automated Heart Disease Prediction Model", International Journal of Fuzzy Logic and Intelligent Systems, vol. 24, no. 3, pp. 231-241, 2024.
- [14] A. Eleyan, E. AlBoghbaish, A. AlShatti, A. AlSultan, D. AlDarbi, "RHYTHMI: A deep learning-based mobile ECG device for heart disease prediction", Applied System Innovation, vol. 7, no. 5, pp. 77, 2024.
- [15] L. Mhamdi, O. Dammak, F. Cottin, I. B. Dhaou, "Artificial intelligence for cardiac diseases diagnosis and prediction using ECG images on embedded systems", Biomedicines, vol. 10, no. 8, pp. 2013, 2022.
- [16] R. Patra, S. Dutta, I. K. Roy, P. Basak, A. Ghosh, "Heart Disease Detection using Vision-Based Transformer Ensemble Models", Procedia Computer Science, no. 258, pp. 3554-3569, 2025.
- [17] R. S. Ram, J. Akilandeswari, M. V. Kumar, "HybDeepNet: a hybrid deep learning model for detecting cardiac arrhythmia from ECG signals", Information Technology and Control, vol. 52, no. 2, pp. 433-444, 2023.
- [18] Z. H. Kilimci, M. Yalcin, A. Kucukmanisa, A. K. Mishra, "Heart disease detection using vision-based transformer models from ECG images", arXiv preprint arXiv, no. 2310, pp. 12630, 2023.
- [19] H. A. Shah, F. Saeed, M. Diyan, N. A. Almujaally, J. M. Kang, "ECG-TransCovNet: a hybrid transformer model for accurate arrhythmia detection using electrocardiogram signals", CAAI transactions on intelligence technology, 2024.
- [20] A. U. Rahman, Y. Alsenani, A. Zafar, K. Ullah, K. Rabie, T. Shongwe, "Enhancing heart disease prediction using a self-attention-based transformer model", Scientific Reports, vol. 14, no. 1, pp. 514, 2024.