Indonesian Journal of Electrical Engineering and Computer Science
Vol. 39, No. 2, August 2025, pp. 1164~1175
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v39.i2.pp1164-1175

Advanced deep attention neural inference network for enhanced arrhythmia detection and accurate classification

H. Sumitha, M. Devanathan
Department of Electrical and Computer Engineering, Reva University, Bangalore, India

Article Info

Article history:

Received Aug 17, 2024
Revised Apr 17, 2025
Accepted Jul 2, 2025

Keywords:

Arrhythmia detection
Attention network
Deep learning
ECG signal analysis
Feature extraction

ABSTRACT

Arrhythmias are irregular heartbeats that can lead to severe health risks, including sudden cardiac death, necessitating accurate and timely detection for effective treatment. Traditional diagnostic methods such as stress tests, resting electrocardiograms (ECGs), and 24-hour Holter monitors are limited by their monitoring capacity and often result in delayed diagnoses, compromising patient safety. To address these challenges, this paper introduces the deep attention neural inference network (DANIN) methodology. DANIN integrates one-dimensional ECG signals with two-dimensional spectral images using multi-modal feature fusion, capturing comprehensive cardiac information in both temporal and frequency domains. The methodology employs advanced deep attention network-based models for superior feature extraction, recognizing intricate patterns and long-range dependencies within the data. Additionally, the inclusion of an inference model system enhances interpretability and usability, making the model highly suitable. Further, DANIN is evaluated considering the MIT-BIH dataset, and extensive comparative analysis with state-of-the-art techniques demonstrates that DANIN significantly improves accuracy, precision, recall, and F1-score, highlighting its potential to revolutionize arrhythmia detection and improve patient outcomes.

Corresponding Author:

H. Sumitha
Department of Electrical and Computer Engineering, Reva University
Bangalore, India
Email: sumithah_12@rediffmail.com

1. INTRODUCTION

Arrhythmias are a significant class of cardiac diseases that require proper diagnosis. This is because arrhythmias raise the possibility of high-risk incidents, such as unexpected cardiac death. An accurate diagnosis must be made quickly to treat the changes in the electrocardiogram (ECG) which have proven to be difficult to correctly detect using traditional automated approaches for arrhythmia detection. Standard diagnostic techniques including stress tests, resting ECGs, and 24-hour Holter monitors are frequently used to identify arrhythmias [1]. The limited cardiac recording capacity, which allows about 2,000 heartbeats to be recorded during a stress test, still restricts the monitoring capabilities of the tests. The patients must endure lengthy wait times for examination results, doctor appointments, and hospital stays, it might take up to four months to get test results. The conventional methods that are being proposed involve making appointments with medical doctors in person and performing a series of diagnostic tests. The patient's health may be affected by these circumstances since there is a decreased likelihood of early detection. The most common method used to diagnose arrhythmias is an ECG recording. To record and analyze the electrical activity of the heart, electrodes can be placed on the skin and left there for a predefined period. The electrical activity of the heart may be assessed from a variety of angles and orientations using ECGs. As such, they can act as indicators for a range of diseases. The primary responsibility of the ECG leads is to detect and assess anomalies in waveforms and rhythms. An ECG is a comprehensive record of the electrical characteristics of a heartbeat. Each of the several waveforms that comprise an ECG signal represents a single pulse. Over the past several years, there has been a sizable surge in the usage of ECG monitoring technology in the medical industry. The observed pattern can be attributed to advancements in enabling technologies, which have greatly increased the capabilities of these systems. For these systems to function correctly, several technologies must be integrated, including edge computing, mobile computing, and the internet of things. These technologies have several other uses in addition to being employed in the diagnosis and treatment of health problems. The device satisfies specific mode-related needs, tracks activities, and enhances sports performance [2]. Numerous studies have been conducted on various techniques for using deep learning (DL) and machine learning (ML) algorithms to classify cardiac arrhythmias also various techniques have been developed by ECG analysis researchers to automatically identify heart arrhythmias. These techniques employ traditional ML techniques. These methods typically consist of three primary stages: preprocessing the signal, extracting features, and recognizing and categorizing patterns. The process of feature extraction has a major impact on how well heartbeat categorization works. Among the often-used feature extraction techniques are principal component analysis (PCA), wavelet transform (WT), discrete wavelet transforms (DWT), independent component analysis (ICA), and other manually developed features. They employed a PCA to identify features and decrease the dimensionality of the ECG data. To construct the atrial fibrillation (AF) detector, it combines the DWT with morphology to extract features. When combined with this method, the implementation of an artificial neural network (ANN) classifier produced positive results. It has been demonstrated that WT is effective in interpreting ECG data due to the signals' inherent non-stationarity.

By extracting certain characteristics from ECG arrhythmia signals, classification models that can distinguish between different types of arrhythmias are developed. support vector machines (SVMs) and ANNs are the two techniques for handling categorization problems which have presented a method for the automatic classification of ECG data using multiple SVMs. Elhaj *et al.* [3] employed a hybrid approach utilizing two SVMs to detect AF. It is widely accepted that the multilayer perceptron (MLP) is the most often used ANN design for categorizing arrhythmias.

The subject of recognizing arrhythmias has been studied extensively and used with DL with efficient results. Because DL performs well in so many different applications-such as photo identification, speech recognition, and machine vision-it is incredibly potent. Twelve distinct rhythm patterns were classified using a deep neural network (DNN) developed. The DNN accuracy was arbitrated sufficient for this task, to classify arrhythmias, which employ a DL methodology. The researchers employed an long short-term memory network (LSTM)-a specific type of neural network-to achieve this. Badr *et al.* [4] described an automated technique for categorizing cardiac arrhythmias in a research study. The model employed a 1D convolutional neural network (CNN) technique. The authors of the study developed a 2D-CNN model to classify arrhythmias. The system uses the whole feature maps of heartbeats that are generated via empirical modal decomposition. Previous research employed recurrent neural networks (RNNs) to discriminate between abnormal and normal heartbeats.

It is not obvious how feature extraction and module classification vary in DL systems. Instead, these two tasks are integrated into a single, smooth process. The DL algorithms employ large volumes of ECG data to automatically identify the crucial elements needed for categorization. Even in cases when operator interaction is not required, interpretability remains a significant challenge when using DL approaches. One benefit of DL systems is their ability to automatically extract characteristics from unprocessed, raw data. Due to its potential to inflict pain on medical staff, this issue is of utmost importance in the field of medical applications [5]. Scientists have used a variety of advanced deep-learning algorithms to classify arrhythmias. However, there has been little improvement in categorization performance.

The motivation for this research on arrhythmia detection and classification is driven by the need to enhance the accuracy and timeliness of diagnosing these potentially fatal cardiac conditions. Arrhythmias, which can lead to sudden cardiac death, pose a significant health risk that necessitates prompt and precise identification. Traditional diagnostic methods, including stress tests, resting ECGs, and 24-hour Holter monitors, are often limited in their monitoring capacity and can result in delayed diagnoses, thereby compromising patient safety. With the advent of advanced ECG monitoring technologies and the integration of cutting-edge DL and ML algorithms, there is an unprecedented opportunity to revolutionize arrhythmia detection. is research aims to leverage these technological advancements to develop more accurate, efficient, and reliable diagnostic tools, ultimately improving early detection and patient outcomes, and reducing the burden on healthcare systems.

- Feature fusion: the deep attention neural inference network (DANIN) methodology introduces a novel approach by integrating one-dimensional ECG data with two-dimensional spectral images through multi-modal feature fusion. This comprehensive analysis in both the temporal and frequency domains enhances the accuracy and robustness of arrhythmia detection.

- Advanced deep attention network-based feature extraction: leveraging the power of deep attention network models, DANIN excels in recognizing and extracting complex patterns and long-range dependencies in ECG data. This advanced feature extraction capability significantly improves the model's ability to accurately classify various types of arrhythmias.
- Enhanced interpretability and usability: by incorporating an inference model system, the DANIN methodology not only boost diagnostic accuracy but also improves the interpretability of the results. This ensures that the models are practical and reliable for clinical use, facilitating better decision-making and patient care.

The paper is organized into 4 sections; the first section gives a brief introduction to arrhythmia the second section gives a thorough literature survey. The third section defines the proposed methodology. The fourth section determines the performance evaluation where the results are given in the form of graphs and tables.

2. RELATED WORK

ECGs have been classified using DNNs in the recent past. DNNs may directly derive a feature extraction function from the raw input data by utilizing the dataset's probability distribution, this is how these approaches are different from traditional ones. Features derived from a DNN model can be more comprehensive than features produced manually when a large enough quantity of training data is available. Ventricular arrhythmias were identified by training an appropriate feature mapping with a stacked denoising autoencoder (SDAE). Then, by adding a SoftMax regression layer to the hidden representation layer, DNNs are used [6]. Automatic identification of cardiac arrhythmias has been made possible by the parallel use of CNNs, by using CNNs to identify AF. The application of a multiscale fusion of deep convolutional neural networks (MS-CNN) has been suggested as a solution to the AF problem [7]. By using filters of various sizes, the method makes use of a two-stream convolutional network architecture to extract information at numerous scales. Dekimpe and Bol [8], a CNN using the residual network design was created to precisely categorize 12 rhythm classes. In the field of arrhythmia classification, CNNs are frequently utilized to classify arrhythmias at the beat level. In these kinds of situations, the model's input data is usually much shorter, frequently numbering just in the hundreds of samples [9]. A nine-layer CNN example was created to automatically recognize and categorize five distinct kinds of heartbeats. Two more network topologies that are often used in the field of ECG classification are the restricted Boltzmann machine (RBM) and the autoencoder. A unique approach based on DL is used as a solution to the previously described problem. The previous method integrates a SVM with an autoencoder network that uses LSTM architecture to classify arrhythmias [10]. According to research, CNN and LSTM may be integrated to automatically classify arrhythmias. The classification performance utilizing various recording times for ECG data was also investigated in this work. Using an ensemble network model based on DL improved the performance of a single network. Three different networks are incorporated into the model's architecture to recognize and gather data. The previously outlined procedure yields an extremely efficient method for data identification and gathering.

According to preliminary research, several algorithms have demonstrated potential in the automated categorization of arrhythmias using ECG data. Before these algorithms are successfully applied in real-world circumstances, several challenges need to be resolved. It is noteworthy that the features of obtained individual ECGs might yield valuable clinical data for automated cardiac arrhythmia identification. However, it's crucial to remember that ECG signals from people with various medical disorders frequently have unique temporal and morphological features [11]. Individual differences might cause each person's ECG signal to fluctuate differently over time. Furthermore, even people with the same medical condition could have different ECG morphologies. It is essential to remember that different heart diseases might present with identical ECG features. One major obstacle is analyzing and extracting characteristics to detect cardiac diseases. The distinct characteristics of each patient's rhythm that may vary from the training set are not taken into account by the arrhythmia classification algorithms currently in use. These algorithms make use of relatively tiny training datasets. As a result, it's possible that the existing techniques won't work as well in practical situations. The gathering of a patient's long-term ECG records is made easier by long-term ECG monitoring equipment, which helps to address this problem. With these gadgets, automated categorization methods are applied [12].

Kiranyaz's real-time patient-specific ECG classification technique is the only way to classify long-term ECG recordings of a patient; Zhang proposed a patient-specific ECG classification method using RNN to classify ECG beats with different heart rates and capture temporal correlation from ECG signal samples. 1D CNNs serve as the foundation for this method. A technique for categorizing five typical kinds of arrhythmia signals using a one-dimensional CNN (1D-CNN) [13]. A deep two-dimensional CNN was used by Jun as a useful method for identifying ECG arrhythmias. In the area of pattern recognition, the aforementioned neural network has demonstrated exceptional performance. Through the use of transfer learning from 2D deep CNN features, this method was able to classify ECG disorders. The methodology outlined was employed to identify and classify four discrete ECG patterns. In their study, [14] introduced a novel technique for carrying out feature extraction across many domains. The proposed approach combines WT with kernel-ICA. The data was compressed using PCA. The ICA approach was applied to the ECG data to detect features and minimize dimensionality [15]. Selecting specific traits is sometimes based on the evaluation of experience. Among the characteristics are higher-order spectra, linear and nonlinear features, sparse features, entropy-based features, and statistical features.

3. PROPOSED METHOD

The proposed DANIN methodology consists of a detailed account of the procedures and techniques employed to collect data and take ECGs. Initially, the process of converting ECG impulses into spectrograms is analyzed, and the proposed approach integrates one-dimensional ECG data with the corresponding two-dimensional spectral images using multimodal feature fusion. Because of this integration, the data is collected in both the temporal and frequency domains, which allows a comprehensive understanding of the cardiac signals. Additionally, a deep attention network-based feature extraction encoder was specifically developed to extract features from one-dimensional ECG data signals and spectrum images. Deep attention networks are often employed in several domains, including natural language processing. However, their application in ECG signal analysis is novel, particularly when integrated with inference model logic.

Deep attention networks are an economical means of representing complex patterns and long-range connections in ECG data. Furthermore, a more accurate categorization and forecasting of ECG data is carried out by examining the combined multi-modal attributes through the utilization of an inference model. The integration of the inference model system overcomes a major barrier in applying DL to medical data and enhances the accuracy of the model's predictions. The inference model module enhances the diagnostic procedure's robustness by providing a means to handle the inherent uncertainties and variabilities in ECG readings. Figure 1 presents the comprehensive schematic of the model's frame.

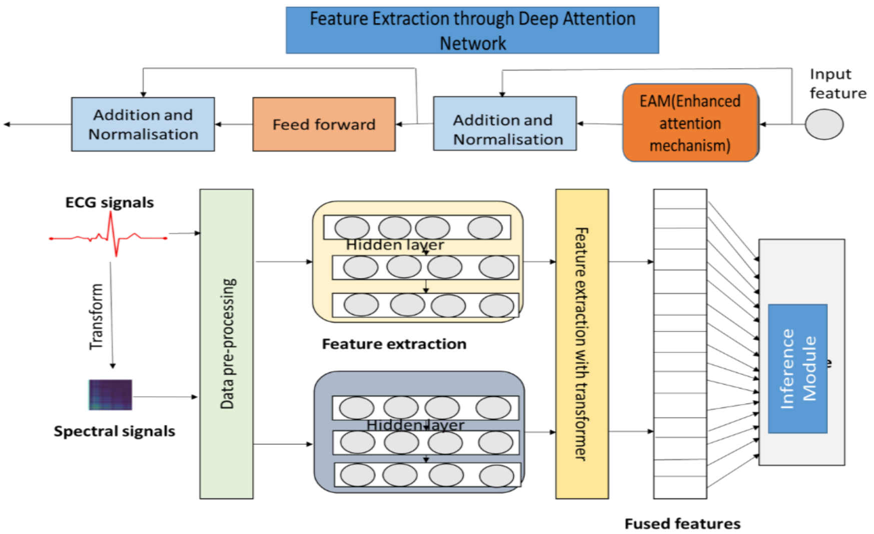

Figure 1. Proposed DANIN architecture

3.1. Efficient signal transformation

The fundamental module of the windowed Fourier transform (WFT) is to decompose the signal into an unlimited number of smaller segments and subsequently perform window processing to each of these segments. The windowed small segment signal may be analyzed using the Fourier transform subsequently it is considered as a stationary signal. Assume the signal is $z(v)$, the transformation of the signal is then applied as given in (1).

$$trans_z(time, hz) = \int z(\alpha) j(\alpha - v) e^{-l2\pi h\alpha} d\alpha \quad (1)$$

Where in hz is frequency and $j(v)$ is the window function, for further processing of the Fourier transform on the signal for the local time-frequency plane, a limited window function $j(v)$ is added before the Fourier transforms so this function segments the non-stationary signal into unlimited no of segments, within the action of the window function $j(v)$, the Fourier transformation is performed on each segment of the signal hence obtaining the local spectrum difference at various times. The frequency associated with each segment of the signal determines the action of the window function $j(v)$, the smaller it is higher its resolution is, the time and bandwidth of the window are contradictory and the limit size of the window is arbitrarily very small, once after windowing the signal $j(v)$ the Fourier transform is further connected to time variable. The window function is evaluated by the (2).

$$j^{(r,s)}(v) = e^{lrv} j(v-s), j(v) \in N^2(T) \tag{2}$$

Here r denotes the frequency shift parameter for the window whereas s denotes the time shift parameter window, by assuming operations like $r = or_q$ and $s = ps_q$, this signal transformation is represented by (3). Where $E(z)$ is the coefficient of the WFT transform, from the above equation wherein $ps_q, or_q(o, q\epsilon\mathrm{B})$ transforms the signal as a segmented series of rectangular equal blocks at equal intervals within the variable j_{or_q}, ps_q . The equal resolution of the signal mapped by $N^2(T)$- $N^2(T^2)$ for the time-frequency plane is an essential feature of the WFT transform.

$$E_{o,p}(z) = \{j_{o,p}, z\} = \{j_{or_q}, ps_q, z\} \tag{3}$$

3.2. Deep attention network-based feature extraction

In this section, the analysis of the structure of the feature extraction module is carried out. Deep attention network designs are the basis of this technique for recognizing and extracting crucial information from spectral images and ECG signals. This bimodal approach utilizes the deep attention network's ability to distinguish complex patterns and distant connections from the input, interpreting it as particularly valuable for decoding the complex structures observed in spectral images and ECG data. The deep attention network model has achieved significant use due to its effectiveness in handling sequential data since its initial development for natural language processing applications. Deep attention network-based models surpass computer vision tasks such as classification, detection, and segmentation, as well as addressing challenges in natural language processing.

Figure 1 illustrates the structure of the typical deep attention network, which consists of a group of encoders and a group of decoders. The self-attention mechanism is the most essential component in both the encoder and decoder. The deep attention network model utilizes the self-attention approach to accurately detect global relationships across different regions. Deep attention network surpasses CNN and RNN in addressing sequence problems due to the same underlying cause. The self-attention mechanism, which is the main mechanism of the model, allows it to prioritize different parts of the input data. This enables the model to reduce the impact of less significant characteristics and concentrate on the significant ones. The deep attention network architecture is adapted for suitable analysis, enabling us to analyze both two-dimensional spectral images and one-dimensional ECG data. The model consists of two parallel branches, each specialized for processing a distinct data modality. The next, step in this process is calculating data for both ECG signals and the corresponding spectral images in the feature extraction model based on deep attention networks. This process concludes with the integration of characteristics. The objective of this process is to extract and merge the most essential data from both modalities for further analysis using a series of computer phases.

For the ECG signal analysis, the deep attention network processes the input signal U consisting of P data points resulting in a sequence of feature vectors $z = [z_1, z_2, \dots, z_P]$. Each z_k is a feature denoted by a time step, the computation is denoted at the deep attention network layer for the ECG branch is denoted by the self-attention mechanism defined as shown in (4). The self-attention mechanism evaluates the attention scores at various positions of the input sequence to focus on the interdependencies irrespective of the position as follows in (4). Here S, M and X is the query, key, and value matrices derived through the input, and f_m is the key dimension evaluated as given in (5).

$$A(S, M, X) = softmax\left(\frac{SM^V}{\sqrt{f_m}}\right)X \tag{4}$$

$$\begin{aligned} S &= Y_s * r \\ M &= Y_m * r \end{aligned} \tag{5}$$

$X = Y_x * r$

Y_s, Y_m, Y_x are weight matrices randomly initialized at the initial of the network training, their value is adjusted as the backpropagation network and r is given as the input to the self-attention layer. Enhanced attention mechanism (EAM) is the extension of the attention mechanism that uses multiple parallel attention to enhance the model's ability that extract the data features. Each head captures the varied attention focus at a similar time aspect the EAM has certain similarities that capture various features irrespective of way. The specific implementation here is a similar mechanism for several sets of weight matrices. The evaluation is as shown in (6). After this, the output for each position goes through a feed-forward network applied at the individual position, through the position-wise feed-forward network as given in (7).

$$h_{k=}A(SY_s^k, MY_m^k, XY_x^k)$$
$$\xi(S, M, X)=concat(head_1, \dots \dots .., head_j)Y^q \quad (6)$$

$$\vartheta_{(Z)}=max(0, zY_1+d_1)Y_2 + d_2 \quad (7)$$

This applies to each element of the sequence that ensures that the model captures the local context within each point of the signal. However, for the spectral image analysis the input spectral image K is denoted by the sequence of vectors as $a = [a_1, a_2, \dots, a_O]$ wherein each a_l correlates to a processed region of the image. The vectors are processed by the deep attention networks relevant to ECG signal analysis, through self-attention and feed-forward networks as spatial and frequency-domain features. Upon receiving the feature representation H_{ECG} and H_{spec}, the next phase is to integrate these features within an integrated representation that captures the insights through both time-domain and frequency-domain data. This integration is shown below in (8). Algorithm 1 shows the DANIN algorithm.

$$H_{hused} = \omega(H_{ECG} + H_1) \quad (8)$$

Algorithm 1. Proposed DANIN algorithm

Input Input: $ECG\ signal_{dataset} F = \{(z_k, a_k)\}_{k=1}^{P}$, here z_k is the $k-th$ ECG signal along with the relevant label;
Step 1 For each signal z_k in the dataset F do
Step 2 Normalize ECG signals z_k , through the min-max normalization
Step 3 Use WFT to get spectral images;
Step 4 End for
Step 5 For each normalized $z_k, norm$ in the dataset F do
Step 6 Use the Deep Attention Network-based model for feature extraction through $z_k, norm$
Step 7 Use the Deep Attention network-based model for feature extraction from spectral images;
Step 8 End for
Step 9 Features extracted are H_{ECG}, H_1;
Step 10 For the set of features extracted the H_{ECG} and H_1 do
Step 11 Integrated features through integration accordingly as to eq (11)
Step 12 Get integrated features as H_{hused}
Step 13 Use adaptive input encoding (AIE) to get H_{hused} according to eq 13
Step 14 Use output extraction (OE) to get a_k according to eq 13
Step 15 End for
Step 16 For each step output a_k do
Step 17 Predict the class of the ECG signal
Step 18 End for
Step 19 Return predicted variables as $\{a_k\}_{k=1}^{P}$
output Predicted variables a_k for each ECG signal

4. PERFORMANCE EVALUATION

The performance evaluation is carried out with the existing state-of-art techniques and the proposed model using metrics such as accuracy (ACC), precision (PRE), recall (RE), and F1-score. The baseline methods for comparison include random forest (RF), logistic regression (LR), K-means clustering, Gaussian Naive Bayes, K-nearest neighbors (KNNs), SVM, decision trees (DT), CNN, RNN, CNN+RNN, ES, and DANIN. The evaluation aims to compare the performance of these models to determine the most effective methods in terms of these metrics. The results are shown in the form of graphs and tables.

4.1. Results

Figure 2 depicts the ACC of various ML and DL models. The models include RF, LR, K-means clustering, Gaussian Naive Bayes, KNNs, SVM, DT, CNN, RNN, CNN+RNN, ES, and DANIN. The chart reveals that the DANIN and ES models exhibit the highest accuracy, closely followed by CNN+RNN, RNN, and CNN models, all of which surpass the 95% mark. DT and SVM show moderately high accuracy, above 80%. In contrast, KNNs, Gaussian Naive Bayes, and K-means clustering perform below 80%, with LR and RF having the lowest accuracies, around 70%. This analysis indicates a clear advantage of DL models, particularly those that combine CNN and RNN, over traditional ML approaches in terms of accuracy.

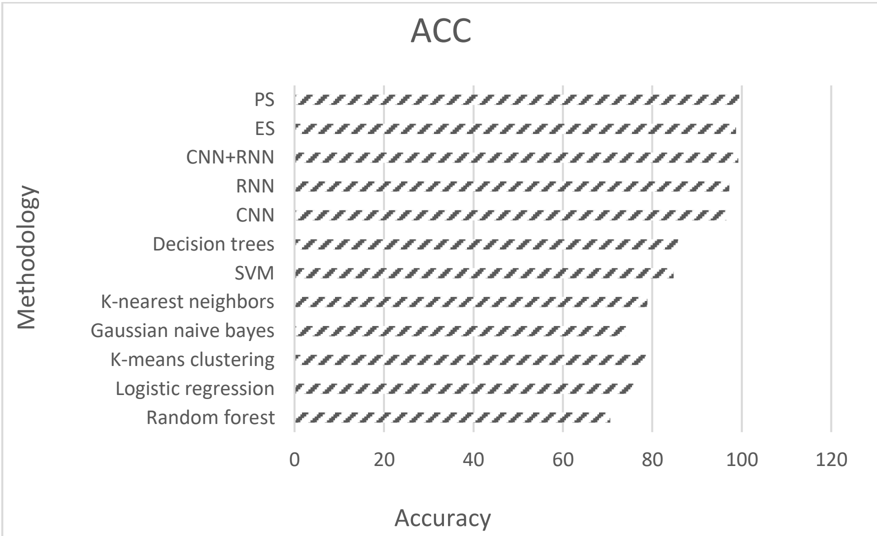

Figure 2. Accuracy comparison of existing state-of-art techniques with DANIN

Figure 3 illustrates the precision (PRE) of various ML and DL models, including DANIN, ES, CNN+RNN, RNN, CNN, DT, SVM, KNNs, Gaussian Naive Bayes, K-means clustering, and LR. The DANIN and ES models again top the chart with the highest precision values, indicating their superior performance in correctly identifying positive instances. The CNN+RNN, RNN, and CNN models also demonstrate high precision, reflecting their effectiveness in minimizing false positives. DT and SVM exhibit moderately high precision, falling just below the top-tier models. KNNs, Gaussian Naive Bayes, and K-means clustering show lower precision, indicating a higher rate of false positives compared to the top-performing models. LR, while slightly better than K-means clustering, still lags behind the other methods. This analysis highlights the superior precision of DL models, particularly the combined CNN+RNN approach, compared to traditional ML models.

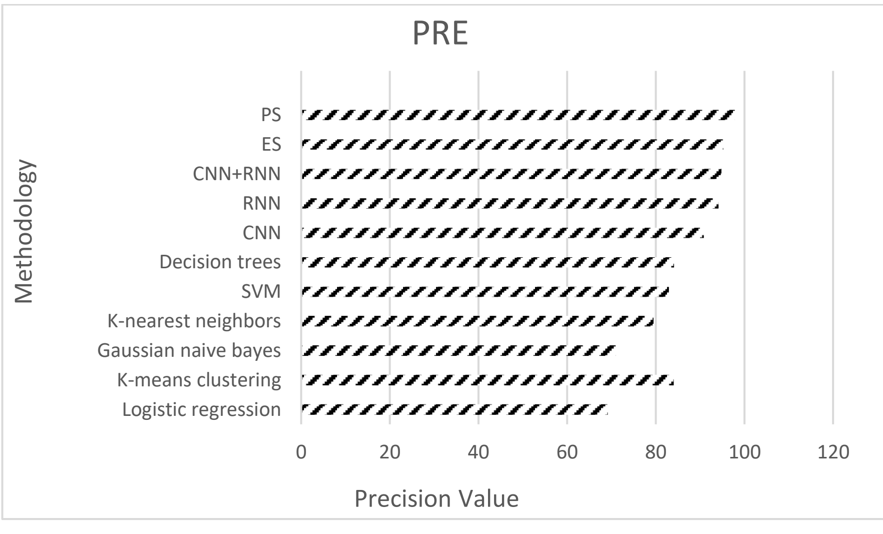

Figure 3. Precision comparison of existing state-of-art techniques with DANIN

Figure 4 presents the (RE) of various ML and DL models, including DANIN, ES, CNN+RNN, RNN, CNN, DT, SVM, KNNs, Gaussian Naive Bayes, K-means clustering, LR, and RF. The DANIN and ES models exhibit the highest recall values, both at 100, indicating their exceptional performance in identifying all relevant instances (true positives). The CNN+RNN model follows closely with a recall value of 93.76, demonstrating its effectiveness in minimizing false negatives. RNN and CNN models also show high recall values, at 87.25 and 88.12 respectively, showcasing their ability to identify a large portion of relevant instances. DT and SVM show moderately high recall values, at 87.55 and 86.28 respectively, indicating a good balance between true positives and false negatives. In contrast, KNNs (78.22), Gaussian Naive Bayes (79.44), and K-means clustering (70.33) exhibit lower recall values, reflecting a higher rate of false negatives compared to the top-performing models. LR and RF perform similarly, with recall values of 91.08 and 84.85 respectively, slightly below those of the better-performing models. This analysis underscores the superior recall of DL models, particularly the combined CNN+RNN approach, compared to traditional ML models, indicating their robustness in identifying relevant instances.

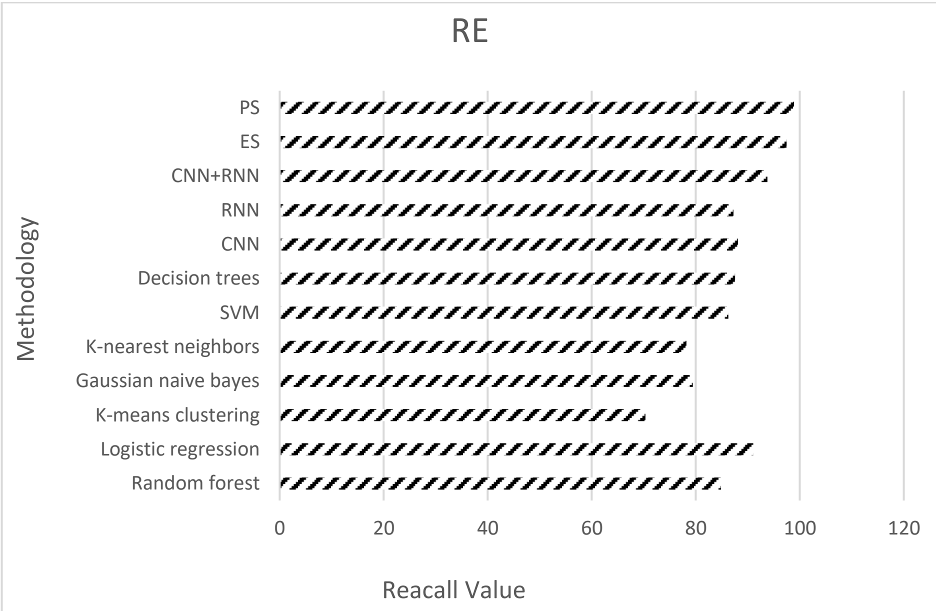

Figure 4. Recall comparison of existing state-of-art techniques with DANIN

Figure 5 shows the F1-score of various ML and DL models, including DANIN, ES, CNN+RNN, RNN, CNN, DT, SVM, KNNs, Gaussian Naive Bayes, K-means clustering, LR, and RF. The DANIN and ES models achieve the highest F1-scores, both at 100, indicating their balanced performance in terms of precision and recall. The CNN+RNN model follows closely with an F1-score of 94.25, demonstrating its effective handling of both true positives and minimizing false positives and negatives. RNN and CNN models also perform well with F1-scores of 90.56 and 89.46, respectively, showcasing their robust performance. DT and SVM exhibit moderately high F1-scores, at 85.81, and recall. In contrast, KNNs (78.86), Gaussian Naive Bayes (74.95), and K-means clustering (76.57) show lower F1-scores, indicating less balanced performance compared to the top-performing models. LR and RF have the lowest F1-scores, at 78.65 and 74.72, respectively, indicating their lower overall performance. This analysis highlights the superior F1-scores of DL models, particularly the combined CNN+RNN approach, compared to traditional ML models.

Table 2 presents a tabular comparison of various models across multiple metrics, including recall (RE) and precision (PRE) for different classes (N, S, V, F). Figure 6 illustrates the performance of various models across multiple metrics, including recall (RE) and precision (PRE) for different classes (N, S, V, F). The models compared are MRFO-SVM, fused transformer, DNN-ensemble, light transformer, WaveINet-Db6, WaveINet-Sym4, MAHA, ISFnet-14, ECGTransForm, ES, and DANIN. DANIN and ES models consistently achieve high values across all metrics, indicating their robust performance with near-perfect recall and precision values. ECGTransForm and MAHA also demonstrate strong performance, particularly in recall and precision for most classes, positioning them among the top models. In contrast, ISFnet-14 exhibits a significant drop in some metrics, particularly in recall (N) and precision (N), suggesting a higher rate of false negatives and less accurate positive predictions for this class. Light transformer and fused transformer maintain balanced performance across most metrics, though not as high as the top-performing DANIN and ES models. WaveINet-Db6 and WaveINet-Sym4 show varying results, with notable drops in certain metrics, reflecting less consistency. DNN ensemble displays lower performance, especially in recall (V) and precision (V), indicating challenges in correctly identifying and classifying relevant instances in these classes. Overall, the chart highlights the superior and consistent performance of DANIN and ES models across all metrics, while other models show varying degrees of effectiveness, with some excelling in specific areas but not across the board.

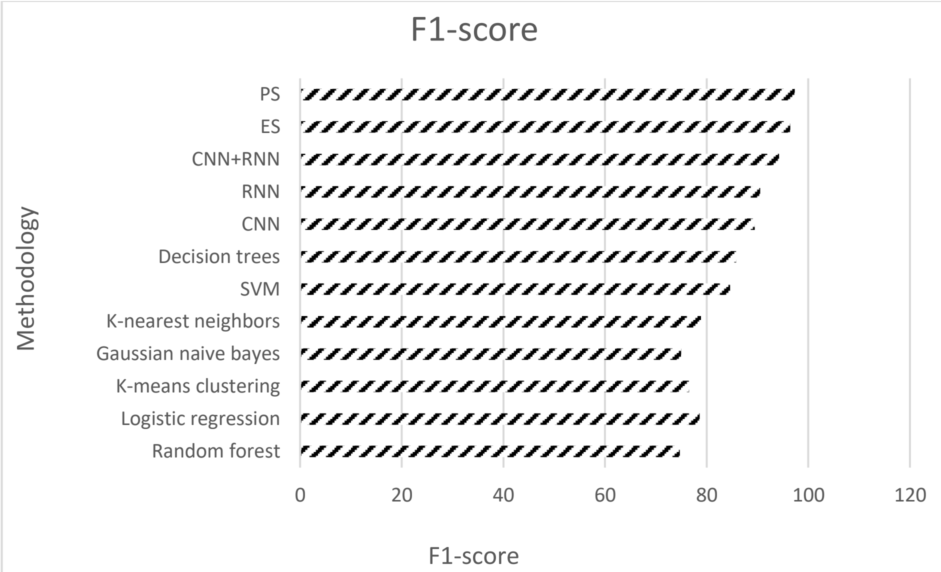

Figure 5. F1-score comparison of existing state-of-art techniques with DANIN

Table 2. Comparison table

Methods	RE (N)	PRE (N)	RE (S)	PRE (S)	RE (V)	PRE (V)	RE (F)	PRE (F)	Average RE	Average PRE	Average F1-score	
MRFO-SVM [16]	98.8	98.7	98	99.3	98.2	96.4	96	97.8	97.8	97.6	97.7	97.7
Fused transformer [17]	99.2	99.2	94.8	91.4	69.3	87.4	86.5	87.7	92.6	90.1	90.1	90.1
DNN-ensemble [18]	93.1	98.1	83.1	90.9	49.4	80.2	49.4	76.4	68.7	86.4	78	78
Light transformer [19]	99.8	99.7	91.5	94.4	83	93.8	91.4	91.4	91.4	93.8	94.8	93.6
WaveINet-Db6 [20]	92	95.3	74.6	64.8	64.8	64.8	64.8	64.8	74.1	64.8	61.3	61.3
WaveINet-Sym4 [21]	91.4	97.7	91.4	65.7	49.3	25.6	64.8	77.3	63.8	66.6	69.4	69.4
MAHA [22]	99.5	91.9	99.4	85	84.9	90.1	91.8	91.8	93.9	89.7	89.7	93.2
ISFnet-14 [23]	63.5	63.8	96.1	91.8	99.4	98.6	85.4	83.1	86.1	84.3	84.2	84.2
ECGTransForm [24]	99.7	97.8	99.7	89.2	93.8	86.5	96.6	93.6	97.6	91.8	93.3	95
ES [25]	99.8	99.7	97.5	97.7	91.7	95.7	100	96.8	97.5	95.2	96.4	96.4
DANIN (PS)	99.9	99.8	98.4	98.87	93.43	96.76	100	97.86	98.76	96.76	97.86	97.87

Figure 7 presents an average comparison of various models based on three metrics: average recall (RE), average precision (PRE), and average F1-score. The models compared include MRFO-SVM, fused transformer, DNN-ensemble, light transformer, WaveINet-Db6, WaveINet-Sym4, MAHA, ISFnet-14, ECGTransForm, ES, and DANIN. From the chart, it is evident that the DANIN and ES models exhibit superior performance across all three metrics, with values close to or reaching 100, indicating their excellent balance between recall, precision, and overall effectiveness. ECGTransForm and MAHA also perform well, showing high average values across all metrics, placing them among the top performers.

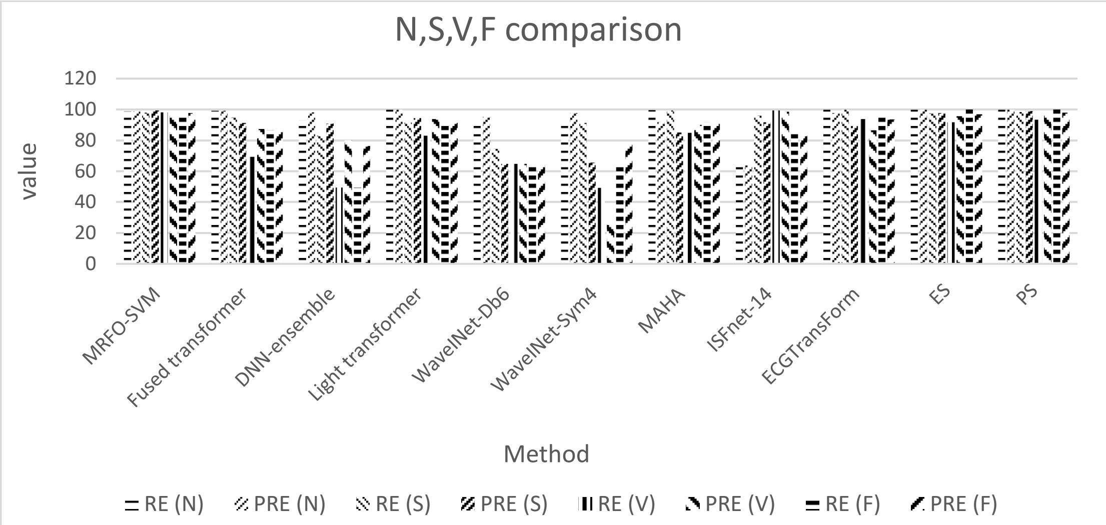

Figure 6. N, S, V, F comparison of existing state-of-the-art techniques with DANIN

Light transformer and fused transformer demonstrate balanced performance, with average values consistently high, although slightly lower than the top models DANIN and ES. WaveINet-Db6 and WaveINet-Sym4 show more variability, with significant drops in some metrics, reflecting less consistent performance. DNN-ensemble and ISFnet-14 have noticeably lower average values, particularly in the recall, indicating challenges in identifying all relevant instances. This lower performance impacts their overall F1-score, suggesting a need for improvement in these models. Overall, the chart highlights the exceptional and consistent performance of DANIN and ES models across all metrics. Other models like ECGTransForm and MAHA also show strong performance, while some models like DNN-ensemble and ISFnet-14 require further optimization to improve their average recall and precision.

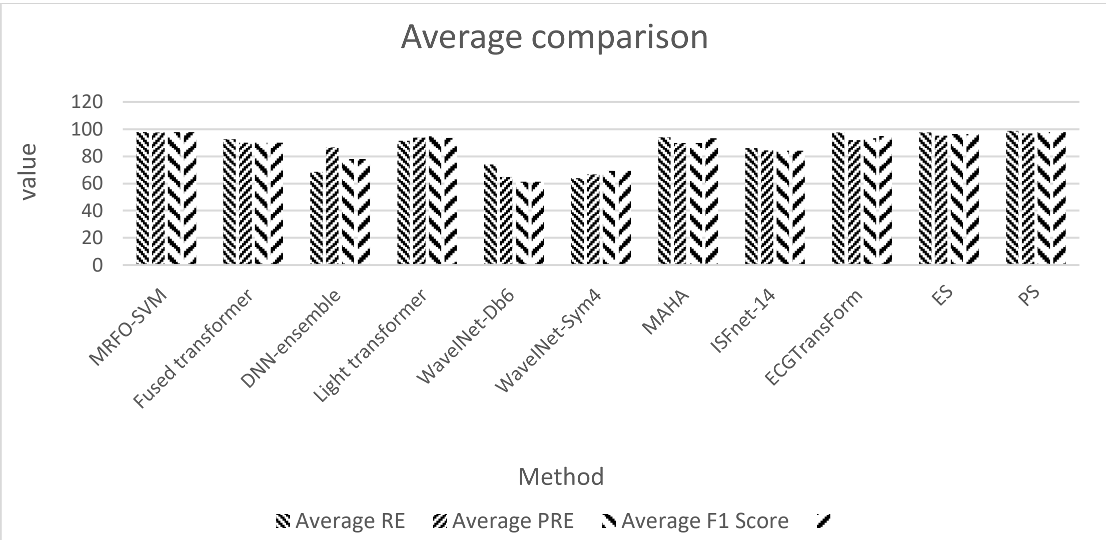

Figure 7. Average comparison of existing state-of-art techniques with DANIN

5. CONCLUSION

In conclusion, the DANIN methodology significantly enhances the accuracy and timeliness of arrhythmia detection by integrating one-dimensional ECG signals with two-dimensional spectral images. This innovative approach leverages deep attention network-based models for superior feature extraction and incorporates an inference model system for improved interpretability and clinical usability. Comparative analysis with traditional and state-of-the-art methods demonstrates that DANIN achieves higher performance metrics, including accuracy, precision, recall, and F1-score. This methodology effectively addresses the limitations of existing diagnostic tools, offering a reliable solution for early and accurate arrhythmia detection, ultimately improving patient outcomes and reducing healthcare problems.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to all those who have supported and contributed to this research project. Primarily, we extend our heartfelt thanks to our guide for his unwavering guidance, invaluable insights, and encouragement throughout the research process.

FUNDING INFORMATION

No funding is raised for this research.

AUTHOR CONTRIBUTION

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
H. Sumitha	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	
M. Devanathan								✓	✓	✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST

Author declares no conflict of interest.

DATA AVAILABILTY

No dataset is utilized for this research.

REFERENCES

[1] X. Dong and W. Si, "Heartbeat dynamics: a novel efficient interpretable feature for arrhythmias classification," *IEEE Access*, vol. 11, pp. 87071–87086, 2023, doi: 10.1109/ACCESS.2023.3305473.

[2] H. Li, D. Yuan, Y. Wang, D. Cui, and L. Cao, "Arrhythmia classification based on multi-domain feature extraction for an ECG recognition system," *Sensors (Switzerland)*, vol. 16, no. 10, p. 1744, Oct. 2016, doi: 10.3390/s16101744.

[3] F. A. Elhaj, N. Salim, A. R. Harris, T. T. Swee, and T. Ahmed, "Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals," *Computer Methods and Programs in Biomedicine*, vol. 127, pp. 52–63, 2016, [Online]. Available: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L2000581011%0Ahttp://dx.doi.org/10.1016/j.cmpb.2015.12.024.

[4] M. Badr, S. Al-Otaibi, N. Alturki, and T. Abir, "Detection of heart arrhythmia on electrocardiogram using artificial neural networks," *Computational Intelligence and Neuroscience*, vol. 2022, 2022, doi: 10.1155/2022/1094830.

[5] M. Hammad, A. M. Iliyasu, A. Subasi, E. S. L. Ho, and A. A. A. El-Latif, "A multitier deep learning model for arrhythmia detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, 2021, doi: 10.1109/TIM.2020.3033072.

[6] R. He *et al.*, "Automatic cardiac arrhythmia classification using combination of deep residual network and bidirectional LSTM," *IEEE Access*, vol. 7, pp. 102119–102135, 2019, doi: 10.1109/ACCESS.2019.2931500.

[7] H. K. Kim and M. H. Sunwoo, "An automated cardiac arrhythmia classification network for 45 arrhythmia classes using 12-lead electrocardiogram," *IEEE Access*, vol. 12, pp. 44527–44538, 2024, doi: 10.1109/ACCESS.2024.3380892.

[8] R. Dekimpe and D. Bol, "ECG arrhythmia classification on an ultra-low-power microcontroller," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 16, no. 3, pp. 456–466, 2022, doi: 10.1109/TBCAS.2022.3182159.

[9] C. H. Lin, C. D. Kan, J. N. Wang, W. L. Chen, and P. Y. Chen, "Cardiac arrhythmias automated screening using discrete fractional-order integration process and meta learning based intelligent classifier," *IEEE Access*, vol. 6, pp. 52652–52667, 2018, doi: 10.1109/ACCESS.2018.2870689.

[10] P. Pujar, A. Kumar, and V. Kumar, "Efficient plant leaf detection through machine learning approach based on corn leaf image classification," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 1, p. 1139, 2024, doi: 10.11591/ijai.v13.i1.pp1139-1148.

[11] S. H. Sreedhara, V. Kumar, and S. Salma, "Efficient big data clustering using adhoc fuzzy c means and auto-encoder CNN," in *Lecture Notes in Networks and Systems*, vol. 563, 2023, pp. 353–368.
[12] S. Sahoo, M. Dash, S. Behera, and S. Sabut, "Machine learning approach to detect cardiac arrhythmias in ECG signals: a survey," *IRBM*, vol. 41, no. 4, pp. 185–194, Aug. 2020, doi: 10.1016/j.irbm.2019.12.001.
[13] H. Fujita and D. Cimr, "Decision support system for arrhythmia prediction using convolutional neural network structure without preprocessing," *Applied Intelligence*, vol. 49, no. 9, pp. 3383–3391, Sep. 2019, doi: 10.1007/s10489-019-01461-0.
[14] C. R. Dell'Aquila, G. E. Canadas, and E. Laciar, "Embedded system for the simultaneous study of SAHS and cardiac arrhythmia," *IEEE Embedded Systems Letters*, vol. 15, no. 2, pp. 53–56, 2023, doi: 10.1109/LES.2022.3208469.
[15] Y. K. Kim, M. Lee, H. S. Song, and S. W. Lee, "Automatic cardiac arrhythmia classification using residual network combined with long short-term memory," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, 2022, doi: 10.1109/TIM.2022.3181276.
[16] J. Li, Y. Zhang, L. Gao, and X. Li, "Arrhythmia classification using biased dropout and morphology-rhythm feature with incremental broad learning," *IEEE Access*, vol. 9, pp. 66132–66140, 2021, doi: 10.1109/ACCESS.2021.3076683.
[17] E. H. Houssein, I. E. Ibrahim, N. Neggaz, M. Hassaballah, and Y. M. Wazery, "An efficient ECG arrhythmia classification method based on Manta ray foraging optimization," *Expert Systems with Applications*, vol. 181, p. 115131, Nov. 2021, doi: 10.1016/j.eswa.2021.115131.
[18] G. Yan, S. Liang, Y. Zhang, and F. Liu, "Fusing transformer model with temporal features for ECG heartbeat classification," in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Nov. 2019, pp. 898–905, doi: 10.1109/BIBM47256.2019.8983326.
[19] H. Wu *et al.*, "A deep neural network ensemble classifier with focal loss for automatic arrhythmia classification," *Journal of Healthcare Engineering*, vol. 2022, pp. 1–11, Sep. 2022, doi: 10.1155/2022/9370517.
[20] L. Meng, W. Tan, J. Ma, R. Wang, X. Yin, and Y. Zhang, "Enhancing dynamic ECG heartbeat classification with lightweight transformer model," *Artificial Intelligence in Medicine*, vol. 124, p. 102236, Feb. 2022, doi: 10.1016/j.artmed.2022.102236.
[21] N. Kim, W. Seo, J. Kim, S. Y. Choi, and S.-M. Park, "WavelNet: A novel convolutional neural network architecture for arrhythmia classification from electrocardiograms," *Computer Methods and Programs in Biomedicine*, vol. 231, p. 107375, Apr. 2023, doi: 10.1016/j.cmpb.2023.107375.
[22] E. Kıymaç and Y. Kaya, "A novel automated CNN arrhythmia classifier with memory-enhanced artificial hummingbird algorithm," *Expert Systems with Applications*, vol. 213, p. 119162, Mar. 2023, doi: 10.1016/j.eswa.2022.119162.
[23] H. Tung, C. Zheng, X. Mao, and D. Qian, "Multi-lead ECG classification via an information-based attention convolutional neural network," *Journal of Shanghai Jiaotong University (Science)*, vol. 27, no. 1, pp. 55–69, Feb. 2022, doi: 10.1007/s12204-021-2371-8.
[24] H. El-Ghaish and E. Eldele, "ECGTransForm: empowering adaptive ECG arrhythmia classification framework with bidirectional transformer," *Biomedical Signal Processing and Control*, vol. 89, p. 105714, Mar. 2024, doi: 10.1016/j.bspc.2023.105714.
[25] Y. Tao, B. Xu, and Y. Zhang, "Refined self-attention transformer model for ECG-based arrhythmia detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–14, 2024, doi: 10.1109/TIM.2024.3400302.

BIOGRAPHIES OF AUTHORS

H. Sumitha earned her Bachelors of Engineering B.E. degree in ECE from IFET College of Engineering, Anna University in 2009. She has obtained her master's degree in M.Tech. (ECE) from Sri Manakula Vinayagar Engineering College in 2011. And currently she is a research scholar at REVA University (Deemed to be University) doing her Ph.D. in electronics and communication engineering and also working as assistant professor at Srinivasa Ramanujan Institute of Technology (Autonomous), Anantapur. She has attended many workshops and induction programs conducted by various universities. Her areas of interest are AI and DL. She can be contacted at email: sumithah_12@rediffmail.com.

Dr. M. Devanathan is an associate professor in the Electronics and Communication Engineering Department at REVA University, Bangalore with an experience of 12.5 years in Teaching. He is qualified in Bachelor and Master Degrees in electronics and communication engineering, and Ph.D. in electronics and communication engineering in the area of Satellite Technology and Space Exploration. His areas of interest are satellite technology, machine learning, artificial intelligence, internet of things, and image processing. He can be contacted at email: devanathan.m@reva.edu.in.