Enhanced Electrocardiogram Arrhythmia Diagnosis with Deep Learning and Selective Attention Mechanism

Hasanain Shakir Mansour[1,2], Morteza Valizadeh[2], Alaa Hussein Abdulaal[3*], Mehdi Chehl Amirani[2]

[1]Department of Computer Technical Engineering, College of Information Technology, Imam Ja'afar Al-Sadiq University, Al-Muthanna, Iraq

[2]Department of Electrical Engineering, Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran

[3]Department of Electrical Engineering, College of Engineering, Al-Iraqia University, Baghdad, Iraq

Received 21 July 2024; received in revised form 05 December 2024; accepted 06 December 2024

DOI: https://doi.org/10.46604/aiti.2024.14034

Abstract

The study aims to improve the diagnosis of arrhythmia in cardiovascular disease management. A novel approach using a deep convolutional network combined with a selective attention mechanism is proposed for electrocardiogram signal classification. The deep convolutional network extracts relevant features directly from raw electrocardiogram signals, while the selective attention mechanism focuses on the most critical regions of the signals and suppresses irrelevant or noisy components. This method achieves an accuracy of 99.70% in multi-class arrhythmia classification and 99.85% in binary classification, significantly outperforming traditional classification algorithms. Furthermore, the selective attention mechanism improves the localization of critical electrocardiogram segments, offering valuable insights for clinicians and aiding in the diagnosis process. This enhanced approach increases diagnostic accuracy and provides a clearer understanding of the electrocardiogram signals, which is crucial for effective patient management in cardiovascular diseases.

Keywords: arrhythmia diagnosis, electrocardiogram (ECG), deep convolutional network (DCNN), selective attention mechanism (SAM)

1. Introduction

Cardiovascular diseases constitute the leading cause of death globally [1], and the World Health Organization (WHO) has listed them as the cause of the maximum number of deaths worldwide, accounting for about 31% of all deaths every year [2]. Cardiac arrhythmia is a condition in which the heartbeat becomes irregular due to some fault in the electrical system of the heart. An electrocardiogram (ECG) is an established diagnostic tool for cardiac arrhythmias since it captures the heart's physiological activity over time [3]. Global annual ECG recordings exceed 300 million and are projected to increase. The popularity of ECG stems from its simplicity, affordability, non-invasiveness, and ability to provide valuable information about the heart's electrical activity, heart rate, and the presence of conditions like arrhythmias or heart attacks [4]. ECG tests are painless, easy, quick to perform, and can be repeated to monitor the progress of certain conditions. Fig. 1 shows the ECG heart cycle [5].

The main part of an ECG contains a P wave, a QRS complex, and a T wave. The P wave indicates atrial depolarization. The QRS complex consists of a Q wave, R wave, and S wave, representing ventricular depolarization. The T wave comes after the QRS complex and indicates ventricular repolarization. Structural, electrical, and circulatory are the three types of cardiovascular systems [6].

* Corresponding author. E-mail address: Engineeralaahussein@gmail.com

English language proofreader: Yu-Rui Huang

Accurate diagnosis of the type of arrhythmia is necessary to select appropriate treatment. However, considerable morphological differences exist in the correct manual identification of ECG components. Moreover, visual identification, the current standard, might bring subjective biases between observers. As illustrated in Fig. 2, ECG signal classes exhibit distinct heartbeat characteristics and patterns, such as fusion, F; normal, N; supra-ventricular ectopic, S; and ventricular ectopic, V, beats [7]. In this regard, researchers have explored alternative methods—deep learning being one—to remove visual and manual interpretations.

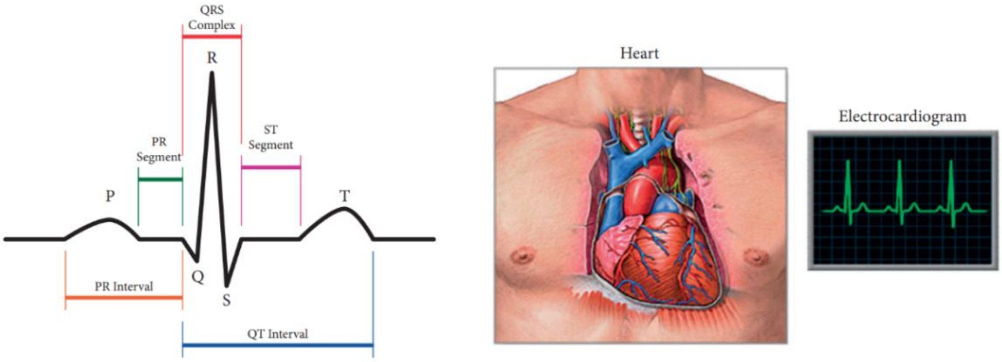


Fig. 1 ECG heart cycles [5]

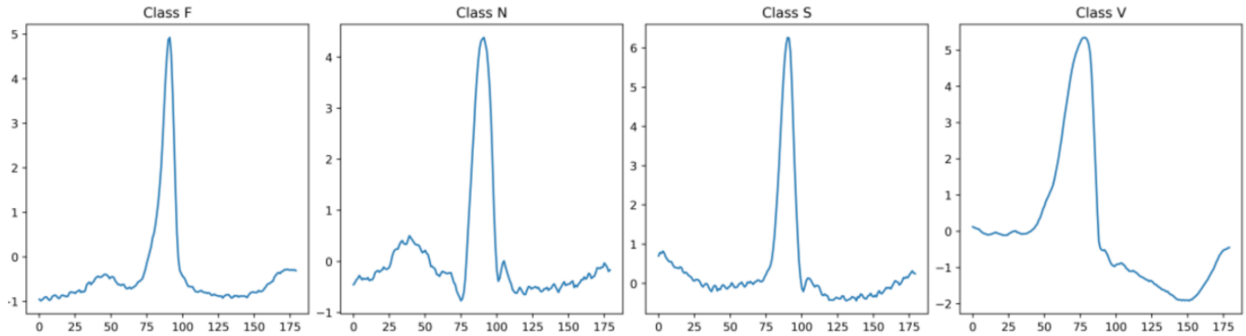


Fig. 2 Heartbeat ECG patterns [7]

Traditionally, clinicians diagnose arrhythmias through manual analysis of ECG signals, which can be time-consuming and susceptible to human error. These methods often lack the precision and efficiency needed for large-scale screenings. Deep learning (DL) offers a promising alternative to cardiac arrhythmia classification, as it can automatically learn significant features and class distinction [8]. DL has advantages in handling massive and noisy datasets, automatically reducing features, and finding applications in different domains. The automatic detection of cardiac arrhythmia has been the topic of many researchers over the past decades. Most of them use the MIT-BIH Arrhythmia Dataset [9] and the Physikalisch Technische Bundesanstalt(PTB) Diagnostic ECG Database [10], which are publicly available and represent the most frequently used database in arrhythmia research. To achieve this, deep genetic hybrid classifiers were used by [11-12] in diagnosing the arrhythmias from the long-term ECG data in the MIT-BIH database with an accuracy of 94.6%.

The proposed approach addresses the deficiencies by including a deep convolutional neural network along with a selective attention mechanism. This allows the model to capture more relevant features from the ECG signals, thereby improving classification accuracy and speed. It leverages state-of-the-art feature extraction with pre-trained models like ResNet-50 and Visual Geometry Group (VGG19), reducing the need for manual intervention and improving model generalization. This method significantly outperforms traditional algorithms, offering a more reliable and scalable solution for arrhythmia diagnosis.

2. Literature Review

Since ECG signals are nonlinear, variations in real-life signals can be detected using higher-order statistical methods like the nonlinear dynamic method [13]. Principal component analysis reduces the dimensionality of the derived bispectrum features. From these reduced features, a least squares-support vector machine and a four-layer feed-forward neural network are used for automatic pattern recognition. This approach achieved the highest average accuracy among the researchers at 93.48%. While the accuracy of deep convolutional network (DCNN) training decreases with increasing network depth, [14] has reported an accuracy of 97.5% using only a simple convolutional neural network (CNN) model with five layers. This technology worked on the principle of wavelet transform based on quadratic waves for identifying individual ECG waveforms and generating a fiduciary marker array. Data classification was performed using a probabilistic neural network with an accuracy rate of 92.7% [15].

Dewangan and Shukla [16] classified heartbeats into five different types by employing artificial neural networks and discrete wavelet transformation. The authors have reported that using wavelet coefficients and morphological features improved the performance of ANN, increasing the accuracy to 87%. The accuracy further increased with the increase in the number of neurons in the hidden layer. Jha and Kolekar [17] contributed an efficient approach to classify seven types of ECG beats based on the 12 approximation coefficients derived through the tunable Q-wavelet transform of ECG beats from a dissimilar record of the MIT-BIH database. Extracted features were used as the input to support the vector machine classifier, which yielded an average classification accuracy of 99.27% for an additional class of eight ECGs.

In [18], researchers proposed an efficient classification methodology in which five types of heartbeats were classified using a one-dimensional CNN with 12 layers. Before classification, the noise was removed from the ECG beats using the threshold denoising technique, which was accurate at 97.41%. In [19], four types of heartbeats from various datasets were categorized through five machine-learning algorithms, including Random Forest. Wavelet decomposition and frequency content-based sub-band coefficients were used to reduce the dataset's dimensionality, improving the performance of the classification technique. The results showed that the Random Forest algorithm achieved a classification accuracy of up to 97%.

Murat et al. [20] provided a survey for background information, and further research into deep learning models became one of the standard approaches by which ECG data had to be classified. They worked on ECG data from 5 classes with 100,022 beats from the MIT-BIH rhythmic database and focused on testing the most commonly used DL strategies available in the literature. Acharya et al. [21] identified general and predictive classes using 13 deep layers of a fully CNN. Like the artificial neural network (ANN), the final CNN model performance judgment depends on network structure weights and previous layer preferences. The pooling process reduces the output neurons' dimension co-evolutionary layer to reduce calculation amplitude and avoid overfitting. The suggested method's accuracy, specificity, and sensitivity were 88.67%, 90.00%, and 95.00%, respectively.

DCNN can use attention mechanisms to diagnose cardiac arrhythmias by focusing only on salient parts of the ECG signals. This provides a new way to isolate and detect relevant features in accurately classifying arrhythmias[22]. The attention mechanism allows for the identification of the ECG signals that drive the classification decision. This implies the localization of the relevant segments of the signals, promoting an understanding of the rationale behind diagnosis and giving valuable insights that enable informed decisions about patient management. Moreover, most of the problems associated with inter-observer variability and subjective biases of visual identification will be decreased due to the attention mechanism. Most arrhythmia diagnosis methods using ECG signals have low accuracy or suffer from similar-looking patterns of arrhythmias. Most conventional methods involve huge feature engineering, which is normally time-consuming and inefficient in most cases. Additionally, many models lack interpretability, making them difficult for clinicians to adopt practically.

The present study exploits the deep learning potential based on CNNs to automatically extract relevant features from ECG signals in classifying various arrhythmias. Another milestone in this area is the integration of a selective attention mechanism within DCNN to diagnose arrhythmias. This provides deep learning with its power while maintaining the interpretability and transparency of the model; further bridging the gap between automated analysis and human understanding. Thus, deep learning embedded with selective attention holds high promise to raise the accuracy, efficiency, and reliability in the diagnosis of arrhythmias for improved patient outcomes in the management of cardiovascular diseases.

3. Materials and Methods

An arrhythmia is an irregular heartbeat resulting from abnormal electrical activity in the heart, which can lead to ineffective blood pumping. It is defined as a deviation from normal heart rate. Tachycardia, bradycardia, and irregular heartbeat are terms used to describe several arrhythmia problems. Bradycardia is characterized by a slow resting heart rate, fewer than 60 beats per minute, whereas tachycardia is characterized by a high resting heart rate, often exceeding 100 beats per minute. The heart's abnormal electrical activity can be fatal. People with coronary artery disease, diabetes, and high blood pressure are more likely to experience arrhythmias.

3.1. Dataset overview

This paper uses the PhysioNet MIT-BIH arrhythmia dataset [11] and The PTB Diagnostic ECG Database [12] as data sources of labeled ECG records. This demonstrates how the knowledge from previous databases can be successfully transferred to train inference models. The ECG lead II resampled at 125Hz is used as an input. The MIT-BIH database contains the ECG recordings of 47 different subjects. The sampling rate is 360Hz. Each beat is annotated with at least two independent cardiologists' estimates. In this paper, annotations from the dataset are used to separate the five beat types under the EC57 standard of the Association for the Advancement of Medical Instrumentation. The PTB Diagnostics dataset consists of ECG recordings of 290 subjects: 148 with MI diagnosis, 52 healthy controls, and the rest diagnosed with 7 diseases. Each record includes ECG signals from the 12 leads sampled at 1000 Hz. This paper will only work on ECG lead II and two categories: MI and health controls.

3.2. Preprocessing

Since ECG beats are used as inputs for this method, Kachuee et al. [23] introduced an efficient approach to preprocessing the ECG signal and extracting its beats. Figs. 3 and 4 illustrate the steps involved in extracting beats from the ECG signal. The continuous ECG signal is divided into 10-second windows, with a specific window selected for further analysis. To enhance signal quality, a combination of noise reduction techniques is employed, including a high-pass filter with a cutoff frequency of 0.5 Hz for baseline wander removal, a notch filter at 50/60 Hz to eliminate power line interference, and a low-pass filter with a cutoff frequency of 40 Hz to reduce high-frequency noise. Additionally, min-max normalization rescales the ECG signal amplitudes to ensure they fall within the range of zero to one.

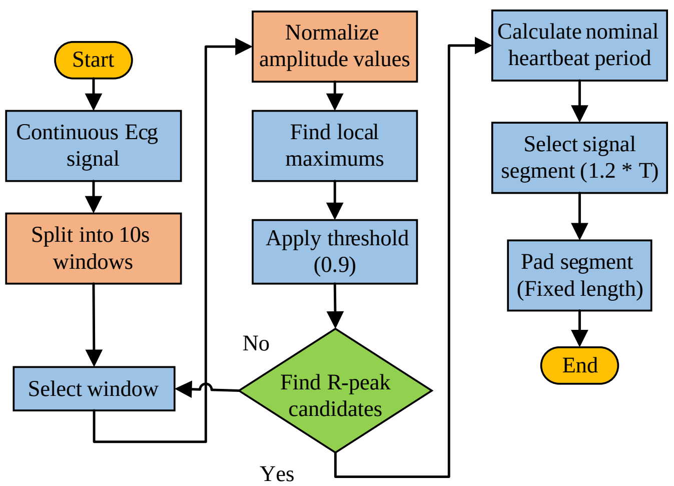


Fig. 3 Flow chart for bit extraction

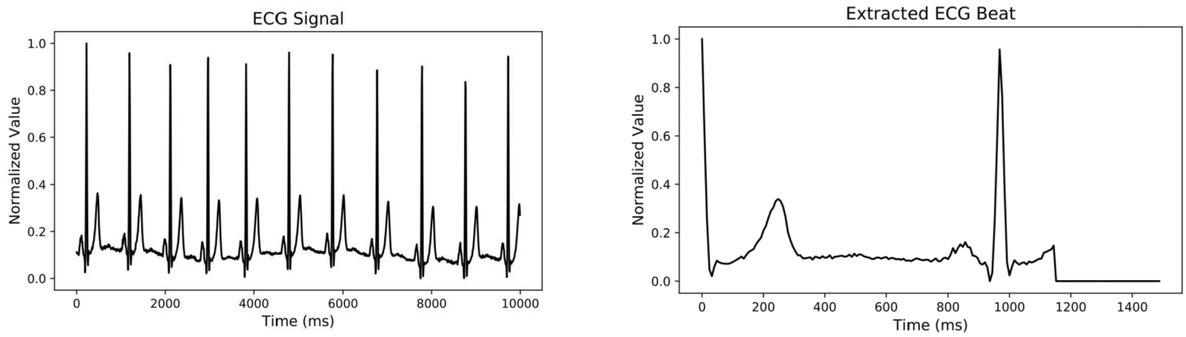


Fig. 4 An extracted beat from 10s ECG window [23]

Zero-crossings of the first derivative are utilized to identify all local maximum points in the signal. A threshold of 0.9 is then applied to the normalized values of these local maximums to identify potential R-peak candidates corresponding to the peaks of heartbeats. The median value of the R-R intervals, representing the time between consecutive R-peaks, is calculated to determine the nominal heartbeat period for that specific window (T). For each R-peak candidate, a signal segment with a length equal to 1.2 times the nominal heartbeat period (1.2T) is selected. To ensure a fixed length for further analysis, each selected segment is padded with zeros, if necessary, to meet a predefined length requirement.

The beat extraction methodology is effective in capturing R-R intervals from ECG signals that exhibit diverse morphological characteristics. Applying this technique, all extracted beats are standardized to an equal length, enhancing the reliability of subsequent analyses. This uniformity is crucial for accurate interpretation, the assessment of heart rate variability, and the evaluation of overall cardiac health.

3.3. Pre-traind CNN

ResNet-50 and VGG19 are popular deep-learning architectures used in computer vision tasks, including image classification, object detection, and feature extraction [24]. While they have different architectural designs, both models have achieved state-of-the-art performance on benchmark datasets.

3.3.1. ResNet-50

ResNet-50 stands for Residual Network with 50 layers. It is a deep residual network proposed by Microsoft Research in 2015. One of the most critical innovations introduced by ResNet-50 is residual or skip connections. These connections enable the network to learn residual mappings, which are differences between a layer's input and output, instead of learning it directly. Residual connections overcome the problem of vanishing gradients during backpropagation and allow for the training of deep networks. In the ResNet-50 architecture, several building blocks of residual blocks are used. Each residual block comprises multiple convolutional layers integrated with batch normalization, ReLU activation, and dropout. ResNet-50 implements skip connections to facilitate gradient flow. Such is the architecture that enables the network to learn more efficiently and effectively. The architecture combines global average pooling and a fully connected layer for classification.

3.3.2. VGG19

VGG19 is a DCNN developed by Visual Geometry Group, VGG, at the University of Oxford. Introduced in 2014, it became widely adopted due to its simplicity and notable performance. VGG19 is an extension of VGG16, with 19 layers composing the architecture [24]. Another critical element of VGG19 is using 3x3 convolution filters throughout the network. This size of filter allows deep networks without excessively increasing the parameter count. The architecture consists of a stack of convolutional layers and a max-pooling layer for down-sampling. The last few layers in VGG19 are fully connected to facilitate the classification task. The architecture of VGG19 is straightforward and uniform; this simplicity makes the model easy to understand and implement. It provides a balanced model complexity and performance and is thus broadly employed as a baseline model for several tasks in computer vision [24]. Due to its deeper architecture, VGG19 contains significantly more parameters than other models.

ResNet-50 and VGG19 have been among computer vision's most influential deep models. ResNet-50 introduced residual connections that enable the training of deep networks, while VGG19 is straightforward yet effective; thus, it is perfectly suitable for use as a baseline model. These models have been successfully applied to many diverse areas with pre-trained representations and powerful features that enable an accurate classification.

4. Proposed Model

In this first stage, the goal is to train ResNet-50 and VGG19 using a public ECG bit dataset containing multi-class classification labels for various cardiac beats. The beat dataset contains labeled ECG signals, and every signal corresponds to a specific cardiac condition or beat type. Continuous ECG signals are segmented into small time windows before training.

Preprocessing of the ECG signals is the first step. Preprocessing techniques include amplitude normalization to a range between zero and one, noise filtering, and resampling the signals to the desired frequency [23]. Later, the previously pre-trained models of ResNet-50 and VGG19, which were trained on large-scale datasets of images, are loaded. These models are used for initialization for training with the ECG bit dataset. Fine-tuning of the pre-trained models on the ECG bit dataset is performed using backpropagation and gradient descent methods. In the process of fine-tuning, the weights are optimized for the multi-class classification task. During training, the models take ECG signal windows as input, and the predicted beat category is compared with the ground truth labels. The iterative optimization process enables the models to learn the patterns and features indicative of each beat category. Evaluation metrics such as accuracy, precision, recall, and the F1 score, can be used to evaluate the performance of the models during training [25-26]. These parameters indicate the classification performance of the models' different types of cardiac beats.

Training ResNet-50 and VGG19 on the ECG bit dataset is for obtaining a model to accurately classify multiple categories of ECG signals: normal beats, supraventricular premature beats, premature ventricular contractions, fusion beats, unclassifiable beats, and myocardial infarction. Subsequently, an ECG signal is divided into equal-sized window fragments to capture meaningful cardiac cycles. Each fragment is then converted into a spectrogram using a Short-Time Fourier Transform (STFT) (1). This involves computing the Fourier transform on short and overlapping time windows to produce a time-frequency representation. This creates a visual image of how the frequency content of the ECG signal changes over time. The spectrograms resulting from this are well-suited for use with CNNs, as these networks can exploit this rich frequency and temporal information for the classification task. This effectively highlights patterns that are less discernible in the raw waveform and thus facilitates better analysis and interpretation of the ECG data.

$$X(t, f) = \int_{-\infty}^{\infty} x(\tau) \cdot w(\tau - t) \cdot e^{-j2\pi f \tau} d\tau \tag{1}$$

where x (τ) is the ECG signal in the time domain, and $w\ (\tau - t)$ is A window function, typically a smooth, finite-duration function, centered at time t.

4.1. Selective attention mechanism (SAM)

SAM allows attention to be focused on specific regions or features of the input. A SAM in the context of ECG signal analysis could highlight the salient patterns or segments responsible for the accurate classification of the different arrhythmia types. The selective attention mechanism involves segmenting the ECG signals into smaller fragments at the front end (2) and subsequently assigning attention scores to them (3). The attention scores are computed by assigning weights to different segments of the ECG signal. These weights are determined based on the relevance of each segment to the classification task, allowing the model to prioritize key features that contribute to identifying specific arrhythmias. These attention scores indicate the diagnostic significance of the segment toward arrhythmia diagnosis. The segments with higher attention scores are considered more informative and are thus given greater weights during classification (4), (5). The one with the highest attention score is selected for further analysis. This can be realized in ECG signal classification by applying a hand-crafted layer in a model's architecture: "SelectiveAttentionLayer (6)." This mechanism ensures the model's focus on essential signal components, enhancing its diagnostic performance.

The ECG segment X is divided into n segments

$$X = \{x_1, x_2, x_3, x_4, \ldots, x_n\} \tag{2}$$

The attention score a_i for segment X_i can be represented as

$$a_i = \{a_1, a_2, a_3, \ldots.., a_n\} \tag{3}$$

The highest attention score X_{Max} is

$$X_{Max} = \arg Max_i \{a_i\} \tag{4}$$

$$X_{weighted} = \sum_{i=1}^{n} a_i X_i \tag{5}$$

The SelectiveAttentionLayer function $SAM(X)$,taking X as input, is

$$SAM(X) = X_{weighted} \tag{6}$$

4.2. Deep convolutional neural network with a selective attention mechanism s

The integration of SAM with the DCNN presents a significant advancement in this study. SAM enhances the DCNN by dynamically weighting the features extracted from ECG signals; thus, the model can concentrate its attention on the most relevant components of the signal, which improves both feature extraction and accuracy in classification. In more detail, SAM accomplishes this by attaching the attention score to different features, assigning higher priority to the features that contribute the most to proper classification. This mechanism is seamlessly embedded within the DCNN framework, enhancing its ability to distinguish between complex arrhythmic patterns.

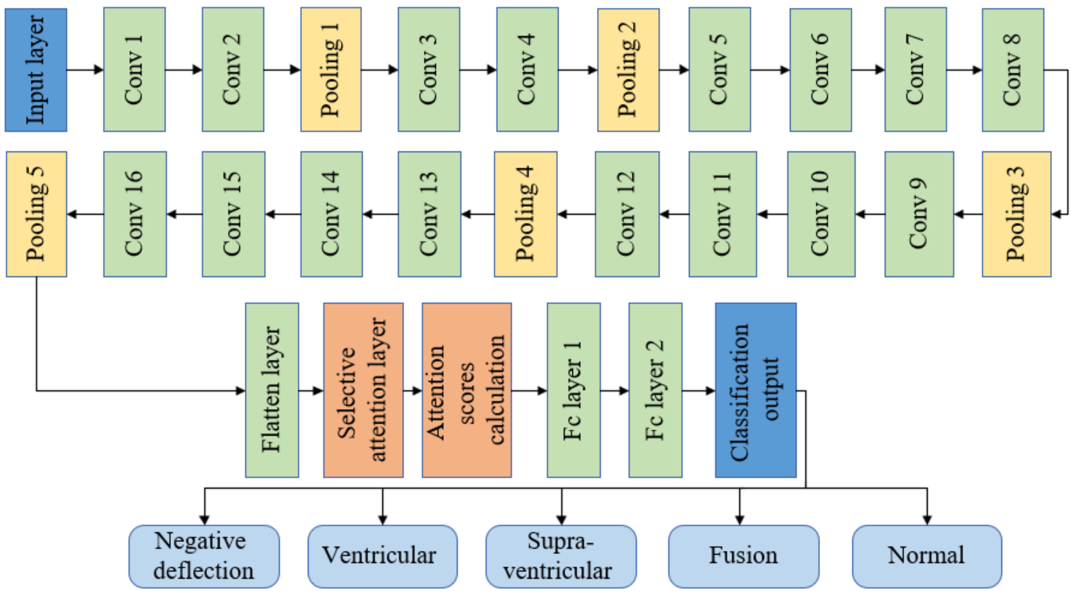


Fig. 5 VGG19-SAM architecture

The second stage centers around the integration of the pre-trained ResNet-50 and VGG19 models with the selective attention mechanism to classify ECG data. This will leverage the rich, learned representations from these pre-trained models while enhancing classification performance by focusing on salient temporal patterns. These pre-trained ResNet-50 and VGG19 models have learned discriminative features from the variably transformed ECG signals. An ECG signal is segmented into window fragments of equal size; and transformed into a spectrogram-like representation. These pre-trained models extract features from every segment to capture high-level representations in complete signal classification. These features are input into SAM, to select the most relevant segments based on their attention scores. It focuses on salient parts of the ECG signal at different steps, emphasizing patterns that contribute more to arrhythmia classification. Fig. 5 illustrates the integration of VGG19 with SAM.

The combination of the pre-trained models with SAM improves the accuracy and robustness of arrhythmia classification by considering only the most informative segments while harnessing the benefit of learned representations from the pre-trained models. SAM improves classification performance by highlighting salient patterns in the ECG signal, as illustrated in Fig. 6. It can effectively utilize informative segments to identify and classify arrhythmia correctly. The model may finally identify the relevant segments strongly and give insights into the existence and type of arrhythmia, thereby helping in making more accurate diagnostic decisions. Further work in this area aims to optimize the integration of a pre-trained model with SAM by investigating variations or adaptations specific to ECG signal analysis. These developments have been crucial in enhancing precision and speed in arrhythmia diagnosis and thereby advancing cardiovascular care.

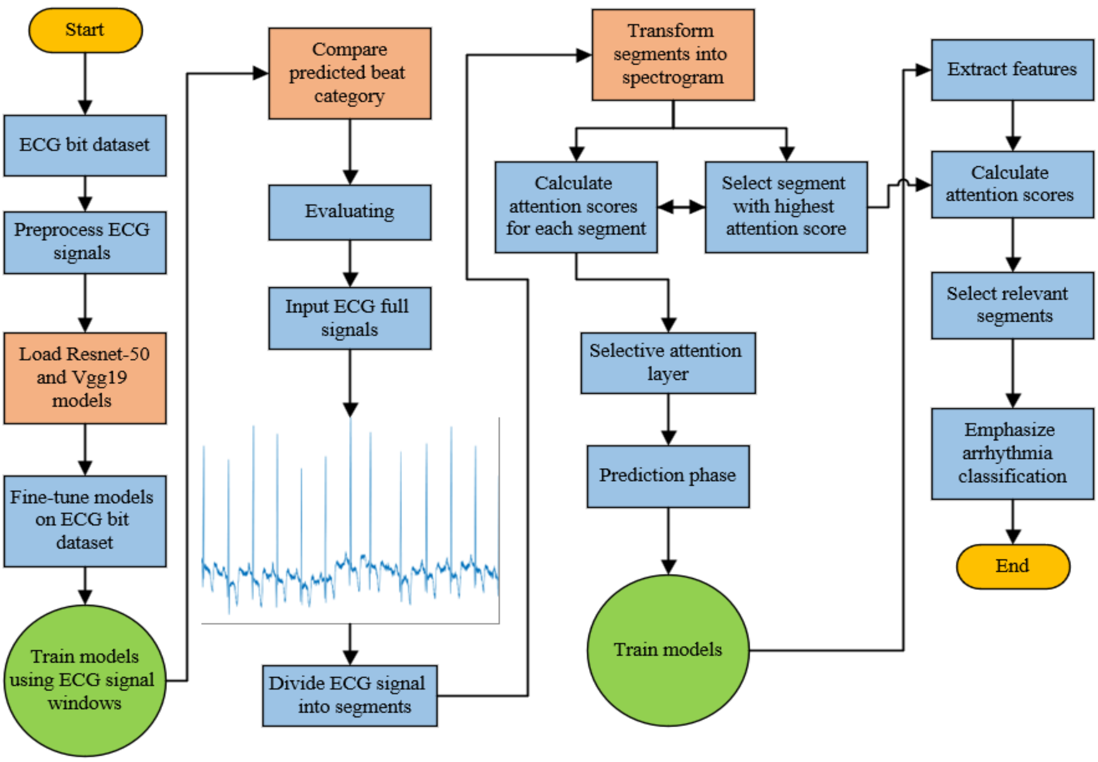


Fig. 6 Flow chart of the proposed model

In clinical settings, the computational costs and complexity of models like VGG19 and ResNet-50 must be considered. VGG19, with its 143 million parameters, requires substantial computational resources, whereas ResNet-50, with 25 million parameters, offers a more efficient option for real-time diagnostics. By integrating a selective attention mechanism (SAM), the interpretability of these models is significantly enhanced. SAM focuses on the most relevant parts of the ECG signal, providing attention maps that align with clinical reasoning. This approach not only improves diagnostic accuracy but also increases transparency, fostering trust in automated systems. By aligning machine predictions with human reasoning, SAM bridges the gap between model outputs and clinical insight, making automated diagnostics more reliable.

4.3. Evaluation metrics

The evaluation metrics used in this paper are as follows [26] :

Accuracy (7): The proportion of true results (both true positives and true negatives) among the total number of cases examined.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} X\ 100 \tag{7}$$

where TP is true positives, FP is false positives, FN is false negatives, and TN is true negatives.

Recall (8): The proportion of true positives among the total number of actual positives.

$$\mathrm{Re}call = \frac{TP}{TP + FN} X\ 100 \tag{8}$$

Precision (9): The proportion of true positives among the total number of positive predictions.

$$\mathrm{Pr}ecision = \frac{TP}{TP + FP} X\ 100 \tag{9}$$

Specificity (10): The proportion of true negatives among the total number of actual negatives.

$$Specificity = \frac{TN}{TN + FP} X\ 100 \tag{10}$$

F1 score (11): The harmonic mean of precision and recall

$$F1 - Score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} X\ 100 \tag{11}$$

The Receiver Operating Characteristic (ROC) curve plots the True Positive Rate (Recall) against the False Positive Rate to evaluate a binary classifier's performance across different threshold values.

5. Results

Experimental results demonstrate the superiority of the proposed approach over traditional methods for arrhythmia diagnosis. Utilizing a DCNN with a SAM for arrhythmia diagnosis has shown promising results in ECG signal classification. The combination of these techniques enhances accuracy by capturing important temporal patterns in the signals. Pretrained networks extract discriminative features, while the SAM highlights relevant segments. The evaluation of the model typically involves 5-fold cross-validation, which enhances the reliability and generalization ability of the performance assessment. This technique trains and evaluates a model on five subsets of a dataset to increase its ability estimation. It improves interpretability, robustness against noise, and generalization performance, demonstrating superior accuracy compared to traditional techniques, and making it valuable for real-time arrhythmia detection.

Table 1 presents the performance metrics for multi-class arrhythmia ECG signal classification. The models achieved high accuracy values: VGG19 with an accuracy of 91.51%, ResNet50 at 94.98%, VGG19+SAM at 97.12%, and ResNet50+SAM with the highest accuracy of 99.70%. These accuracy scores prove that the models are capable of accurately classifying the various classes of arrhythmia.

High precision was recorded for all models in this work, reflecting the reliability of the identification of arrhythmia cases among the predicted positives. VGG19 achieved a precision of 91.68%, and ResNet50 achieved 95.04%.VGG19+SAM achieved 97.14%, and ResNet50+SAM maintained a precision of 99.70%. These precision values indicate low false positive rates in model predictions.

The F1-Score, which considers both precision and recall, demonstrated outstanding performance across all models: VGG19 at an F1-Score of 91.45%, ResNet50 at 94.97%, VGG19+SAM at 97.11%, and ResNet50+SAM was the best at 99.70%. These scores indicate the ability models can balance precision and recall, yielding an accurate classification of Arrhythmia cases.

Specificity, which measures a model's ability to correctly classify subjects without Arrhythmia, showed remarkable results: the specificity was 97.88% for VGG19, 98.74% for ResNet50, and 99.28% for VGG19+SAM, with the highest being for ResNet50+SAM at a specificity of 99.93%. These specificity values highlight the ability of models to recognize subjects without arrhythmia.

Confidence intervals (CIs) estimate a range likely containing the true population parameter, often calculated at a 95% confidence level. In this study, the classification train accuracy ranged from 99.44%- 99.99%. P-values quantify the strength of evidence against the null hypothesis. In this case, ANOVA was used for statistical analysis, yielding a statistically significant p-value of 0.000006, well below the threshold of less than 0.05.

Table 1 Evaluation metrics of multi-class classification

VGG19								
Metrics	N	S	V	F	Q	Average	CIs	P-Value
Accuracy	99.08%	90.29%	91.98%	82.61%	93.59%	91.51%	84.1%-98.92%	0.0051
Precision	88.50%	87.87%	93.68%	94.11%	94.22%	91.68%		
Error	0.92%	9.71%	8.02%	17.39%	6.41%	8.49%		
F1-Score	93.49%	89.06%	92.82%	87.99%	93.91%	91.45%		
Specificity	96.78%	96.88%	98.45%	98.71%	98.56%	97.88%		
ResNet 50								
Accuracy	99.23%	94.28%	94.75%	90.68%	95.96%	94.98%	94.19%-99.94%	0.0043
Precision	93.59%	92.48%	95.63%	97.05%	96.43%	95.04%		
Error	0.77%	5.72%	5.25%	9.32%	4.04%	5.02%		
F1-Score	96.33%	93.37%	95.18%	93.76%	96.19%	94.97%		
Specificity	98.30%	98.08%	98.92%	99.31%	99.11%	98.74%		
VGG19+SAM								
Accuracy	99.49%	97.05%	98.00%	93.17%	97.89	97.12%	94.16%-99.99%	0.000021
Precision	96.62%	95.62%	98.02%	98.16%	97.27%	97.14%		
Error	0.51%	2.95%	2.00%	6.83%	2.11%	2.88%		
F1-Score	98.04%	96.33%	98.01%	95.60%	97.57%	97.11%		
Specificity	99.13%	98.89%	99.51%	99.56%	99.31%	99.28%		
ResNet 50+SAM								
Accuracy	99.94%	99.65%	99.79%	99.38%	99.75%	99.702%	99.44%-99.99%	0.000006
Precision	99.76%	99.18%	99.97%	99.80%	99.81%	99.704%		
Error	0.06%	0.35%	0.21%	0.62%	0.25%	0.298%		
F1-Score	99.85%	99.42%	99.88%	99.59%	99.78%	99.704%		
Specificity	99.94%	99.79%	99.99%	99.95%	99.95%	99.93%		

The ResNet50+SAM model exhibits outstanding performance in binary classification between myocardial and normal cases, as shown in Table 2. It achieves high scores across metrics such as accuracy, precision, recall, F1-Score, and specificity, ranging from 99.82% to 99.88%. These results indicate that this model efficiently classifies cases correctly, reliably detects true positives, and is extremely low in false positives. Overall, ResNet50+SAM worked well in distinguishing myocardial cases from normal ones with exceptional accuracy and reliability.

Table 2 Evaluation metrics for binary classification

CNN	Accuracy	Precision	Recall	F1-Score	Specificity
VGG19	92.33%	92.08%	92.53%	92.31%	92.13%
ResNet50	93.81%	93.87%	93.76%	93.82%	93.87%
VGG19+SAM	97.62%	97.56%	97.68%	97.62%	97.56%
ResNet50+SAM	99.85%	99.88%	99.82%	99.85%	99.88%

Fig. 7 displays deep dream images for features in the `selective_attention` layer in a neural network trained for multiclass arrhythmias from ECG signals. It consists of a 4x4 grid of subplots, each illustrating a deep dream image that maximizes the activation of a specific feature. These images highlight the patterns learn by the network's features for arrhythmia classification. Each subplot is titled with its corresponding feature number. This visualization helps interpret the network's representations and its ability to distinguish arrhythmia types from ECG data.

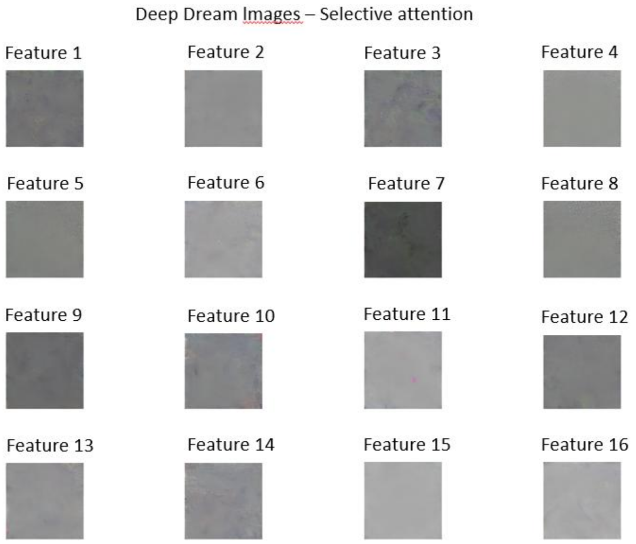


Fig. 7 Deep dream with SAM

Feature extraction, as shown in Fig. 8, visualizes the learned features of a neural network trained for multiclass classification of arrhythmias from ECG signals. Each subplot in this figure represents feature activations for the test image, described as bar plots where the x-axis denotes feature indices and the y-axis shows activation values. Higher bars indicate those features that are most important/relevant for classifying a particular image.

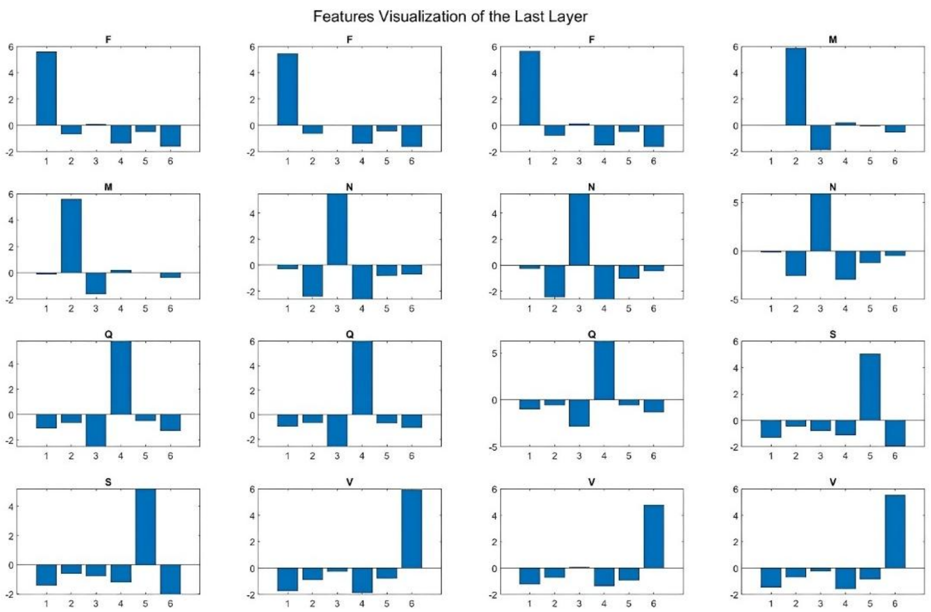
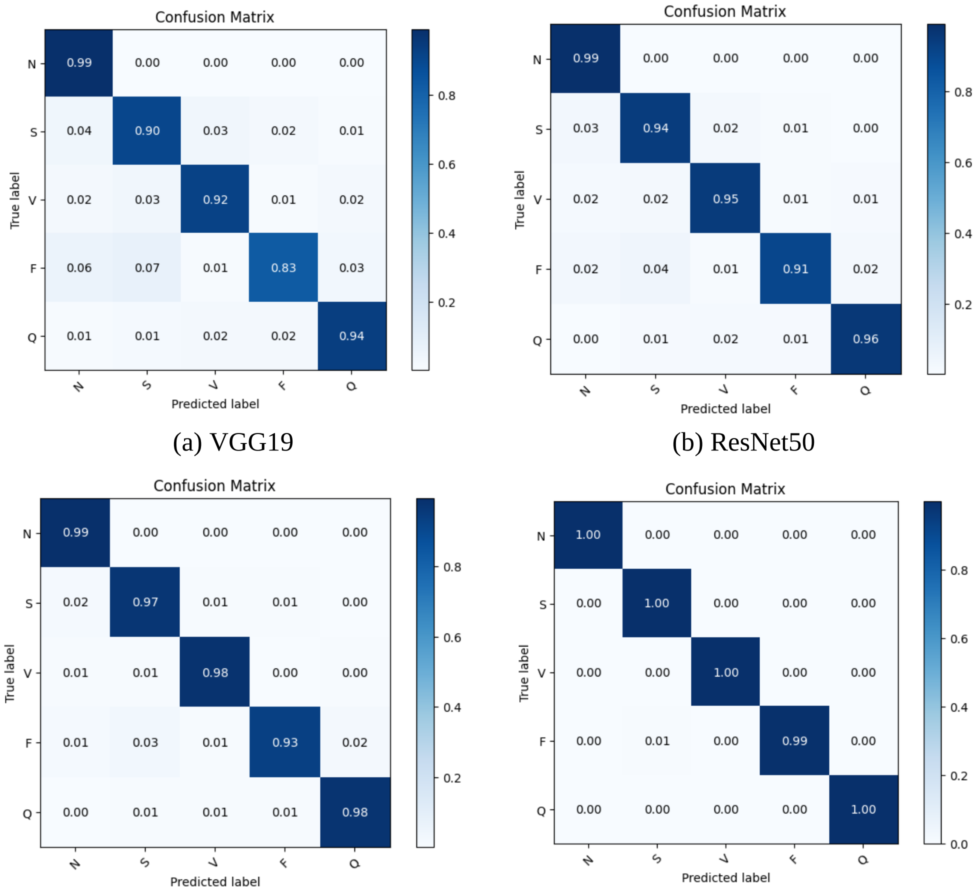


Fig. 8 Features visualization of the last layer

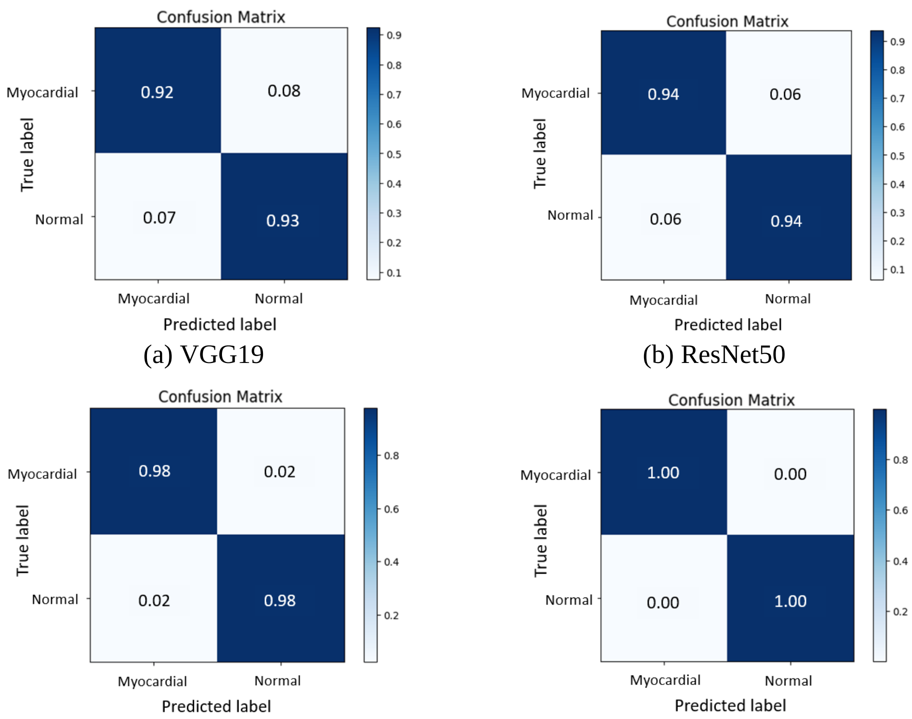
The titles of the subplots display the real class of arrhythmia for every signal; therefore, enabling a visual comparison across classes. This visualization helps with relevant features for separating the arrhythmia classes and provides hints into model decision-making, which may guide improvements in diagnostic tools and treatment strategies.



(a) VGG19 (b) ResNet50

(c) VGG19+SAM (d) ResNet50+SAM

Fig. 9 Confusion matrices (multi-class)



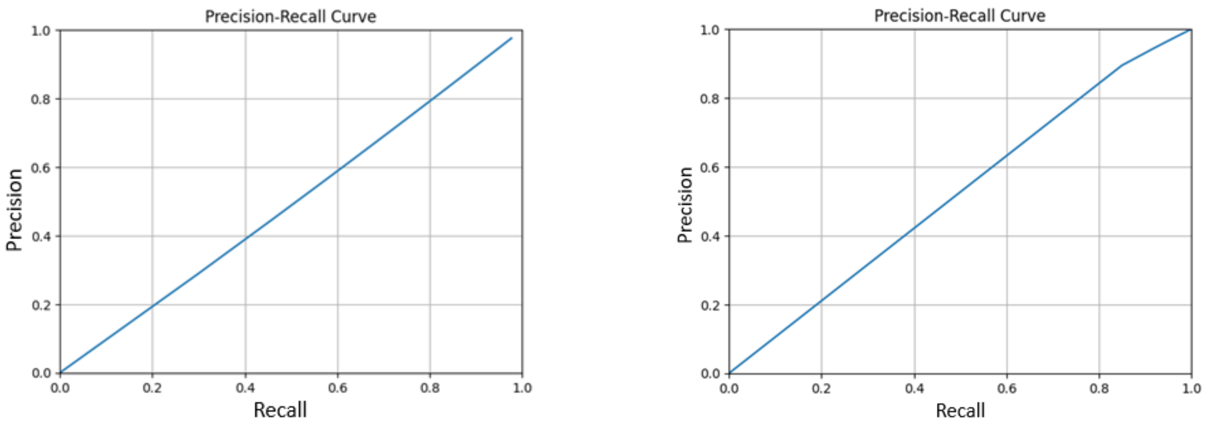
(a) VGG19 (b) ResNet50

(c) VGG19+SAM (d) ResNet50+SAM

Fig. 10 Confusion matrices (Binary)

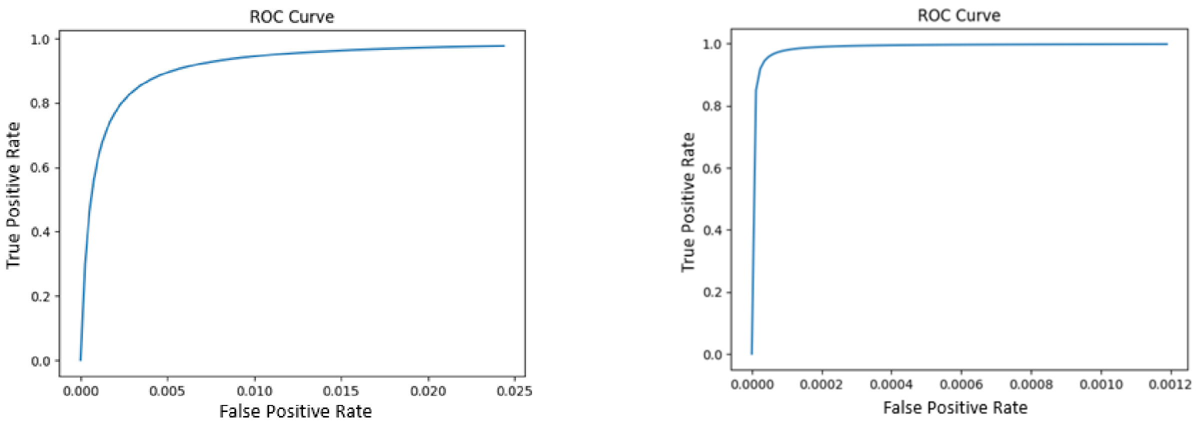
Confusion matrices corresponding to the proposed models for multiclass and binary classification are illustrated in Fig. 9 and 10. The confusion matrices comprehensively summarize the normalized process and provide a clear indication of how well each model fits the data.

The Precision-Recall and ROC curves are plots used to evaluate the performance of models in classifying the arrhythmias. The Precision-Recall curve highlighting the trade-off between the two metrics, precision and recall, is represented in Fig. 11. Fig. 12 presents the ROC curve, which plots the true positive rate against the false positive rate, assessing model performance across different classification thresholds. These curves provide insights into model effectiveness, thus making comparisons possible and full decisions based on such performance characteristics.



(a) VGG19+SAM (b) ResNet-50+SAM

Fig. 11 The Precision-recall curve



(a) VGG19+SAM (b) ResNet-50+SAM

Fig. 12 The ROC curve

Table 3 summarizes the most relevant research efforts reviewed. This collation helps provide input into the commonalities among the individual studies, which have tremendously advanced arrhythmia diagnosis using machine and deep learning techniques with ECG signal analysis.

Table 3 Comparison of proposed methodology with some existing method

Ref.	Model	Class	Accuracy	Specificity	Sensitivity	Precision	F1 score	Computational Complexity
[27]	NN	5	98.90%	98.90%	98.90%	-	-	Low
[28]	MLP	4	94.76%	96.50%	90.11%	-	-	Low
	SVM	4	98.20%	98.79%	96.45%	-	-	Medium
[29]	LSTM	5	99.37%	99.14%	94.89%	96.73%	95.77%	Medium
[30]	CNN	3	99.2%	99.6%	99.2%	-	-	High
2024	This Work CNN+SAM	5	99.702%	99.93%	-	99.704%	99.704%	High
		2	99.85%	99.88%	99.82%	99.88%	99.85%	

Implementing the deep convolutional network (DCNN) with a selective attention mechanism (SAM) in clinical practice offers significant potential to enhance arrhythmia diagnosis. Integrating the DCNN-SAM model into existing systems could streamline workflows, reduce diagnostic errors, and allow clinicians to concentrate on complex cases. Training programs will be necessary to help clinicians interpret the model's outputs appropriately. Challenges may include ensuring system compatibility and addressing data security concerns. These problems are key concerns for widespread adoption. Ultimately, by increasing the accuracy of diagnostics, this approach enhances diagnostic accuracy, provides clearer insights into ECG signals, and supports timely, individualized treatment plans, hence improving patients' outcomes in cardiovascular care.

6. Conclusion

This study proposed a model for arrhythmia diagnosis based on ECG signal classification using pre-trained ResNet-50 and VGG19 models combined with a selective attention mechanism to enhance accuracy and robustness by focusing on prominent signal patterns. The approach involved preprocessing ECG signals, fine-tuning the models for binary and multi-class classification, and utilizing attention scores to emphasize critical signal segments during feature extraction and classification. The proposed approach significantly outperformed traditional methods, achieving high accuracy rates of 99.70% and 99.85% in multi-class and binary arrhythmia classification between myocardial and normal cases, respectively. With its highly accurate ECG signal classification, this approach can improve diagnostic efficiency and accuracy in managing cardiovascular disease.

Conflicts of Interest

The authors declare no conflict of interest.

Statement of Ethical Approval

For this type of study, statement of human rights is not required.

References

[1] O. Gaidai, Y. Cao, and S. Loginov, “Global Cardiovascular Diseases Death Rate Prediction,” Current Problems in Cardiology, vol. 48, no. 5, article no. 101622, 2023.

[2] B. R. Smith and E. R. Edelman, “Nanomedicines for Cardiovascular Disease,” Nature Cardiovascular Research, vol. 2, pp. 351-367, 2023.

[3] V. A. Ardeti, V. R. Kolluru, G. T. Varghese, and R. K. Patjoshi, “An Overview on State-of-the-Art Electrocardiogram Signal Processing Methods: Traditional to AI-Based Approaches,” Expert Systems with Applications, vol. 217, article no. 119561, 2023.

[4] T. Anbalagan, M. K. Nath, D. Vijayalakshmi, and A. Anbalagan, “Analysis of Various Techniques for ECG Signal in Healthcare, Past, Present, and Future,” Biomedical Engineering Advances, vol. 6, article no. 100089, 2023.

[5] W. Ullah, I. Siddique, R. M. Zulqarnain, M. M. Alam, I. Ahmad, and U. A. Raza, “Classification of Arrhythmia in Heartbeat Detection Using Deep Learning,” Computational Intelligence and Neuroscience, vol. 2021, no. 1, article no. 2195922, 2021.

[6] O. Taylan, A. S. Alkabaa, H. S. Alqabbaa, E. Pamukçu, and V. Leiva, “Early Prediction in Classification of Cardiovascular Diseases with Machine Learning, Neuro-Fuzzy and Statistical Methods,” Biology, vol. 12, no. 1, article no. 117, 2023.

[7] A. A. Ahmed, W. Ali, T. A. A. Abdullah, and S. J. Malebary, “Classifying Cardiac Arrhythmia from ECG Signal Using 1D CNN Deep Learning Model,” Mathematics, vol. 11, no. 3, article no. 562, 2023.

[8] J. Cui, L. Wang, X. He, V. H. C. de Albuquerque, S. A. AlQahtani, and M. M. Hassan, “Deep Learning-Based Multidimensional Feature Fusion for Classification of ECG Arrhythmia,” Neural Computing and Applications, vol. 35, no. 18, pp. 16073-16087, 2023.

[9] G. B. Moody and R. G. Mark, “The Impact of the MIT-BIH Arrhythmia Database,” IEEE Engineering in Medicine and Biology Magazine, vol. 20, no. 3, pp. 45-50, 2001.

[10] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung Der EKG-Signaldatenbank CARDIODAT Der PTB Über Das Internet," Biomedizinische Technik/Biomedical Engineering, vol. 40, no. s1, pp. 317-318, 2009.

[11] Ö. Yıldırım, P. Pławiak, R. S. Tan, and U. R. Acharya, "Arrhythmia Detection Using Deep Convolutional Neural Network with Long Duration ECG Signals," Computers in Biology and Medicine, vol. 102, pp. 411-420, 2018.

[12] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam, "Application of Deep Convolutional Neural Network for Automated Detection of Myocardial Infarction Using ECG Signals," Information Sciences, vol. 415-416, pp. 190-198, 2017.

[13] R. J. Martis, U. R. Acharya, K. M. Mandana, A. K. Ray, and C. Chakraborty, "Cardiac Decision Making Using Higher Order Spectra," Biomedical Signal Processing and Control, vol. 8, no. 2, pp. 193-203, 2013.

[14] D. Li, J. Zhang, Q. Zhang, and X. Wei, "Classification of ECG Signals Based on 1D Convolution Neural Network," Proceedings of IEEE 19th International Conference on E-Health Networking, Applications and Services (Healthcom), pp. 1-6, 2017.

[15] J. A. Gutiérrez-Gnecchi, R. Morfin-Magaña, D. Lorias-Espinoza, A. del C. Tellez-Anguiano, E. Reyes-Archundia, A. Méndez-Patiño, et al., "DSP-Based Arrhythmia Classification Using Wavelet Transform and Probabilistic Neural Network," Biomedical Signal Processing and Control, vol. 32, pp. 44-56, 2017.

[16] N. K. Dewangan and S. P. Shukla, "ECG Arrhythmia Classification Using Discrete Wavelet Transform and Artificial Neural Network," Proceedings of IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology, pp. 1892-1896, 2016.

[17] C. K. Jha and M. H. Kolekar, "Cardiac Arrhythmia Classification Using Tunable Q-wavelet Transform Based Features and Support Vector Machine Classifier," Biomedical Signal Processing and Control, vol. 59, article no. 101875, 2020.

[18] M. Wu, Y. Lu, W. Yang, and S. Y. Wong, "A Study on Arrhythmia via ECG Signal Classification Using the Convolutional Neural Network," Frontiers in Computational Neuroscience, vol. 14, article no. 564015, 2021.

[19] S. M. Qaisar, A. Mihoub, M. Krichen, and H. Nisar, "Multirate Processing with Selective Subbands and Machine Learning for Efficient Arrhythmia Classification," Sensors, vol. 21, no. 4, article no. 1511, 2021.

[20] F. Murat, O. Yildirim, M. Talo, U. B. Baloglu, Y. Demir, and U. R. Acharya, "Application of Deep Learning Techniques for Heartbeats Detection Using ECG Signals-Analysis and Review," Computers in Biology and Medicine, vol. 120, article no. 103726, 2020.

[21] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep Convolutional Neural Network for the Automated Detection and Diagnosis of Seizure Using EEG Signals," Computers in Biology and Medicine, vol. 100, pp. 270-278, 2018.

[22] W. S. Admass and G. A. Bogale, "Retracted: Arrhythmia Classification Using ECG Signal: A Meta-Heuristic Improvement of Optimal Weighted Feature Integration and Attention-Based Hybrid Deep Learning Model," Biomedical Signal Processing and Control, vol. 87, no. B, article no. 105565, 2024.

[23] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG Heartbeat Classification: a Deep Transferable Representation," Proceedings of IEEE International Conference on Healthcare Informatic, pp. 443-444, 2018.

[24] A. H. Abdulaal, M. Valizadeh, M. C. Amirani, and A. F. M. S. Shah, "A Self-Learning Deep Neural Network for Classification of Breast Histopathological Images," Biomedical Signal Processing and Control, vol. 87, no. B, article no. 105418, 2024.

[25] A. H. Abdulwahhab, A. H. Abdulaal, A. H. T. Al-Ghrair, A. A. Mohammed, and M. Valizadeh, "Detection of Epileptic Seizure Using EEG Signals Analysis Based on Deep Learning Techniques," Chaos, Solitons & Fractals, vol. 181, article no. 114700, 2024.

[26] A. H. Abdulaal, M. Valizadeh, B. M. Albaker, R. A. Yassin, M. C. Amirani, and A. F. M. S. Shah, "Enhancing Breast Cancer Classification Using a Modified GoogLeNet Architecture with Attention Mechanism," Al-Iraqia Journal of Scientific Engineering Research, vol. 3, no. 1, pp. 47-63, 2024.

[27] 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.

[28] S. Sahoo, B. Kanungo, S. Behera, and S. Sabut, "Multiresolution Wavelet Transform Based Feature Extraction and ECG Classification to Detect Cardiac Abnormalities," Measurement, vol. 108, pp. 55-66, 2017.

[29] S. K. Pandey and R. R. Janghel, "Automatic Arrhythmia Recognition from Electrocardiogram Signals Using Different Feature Methods with Long Short-Term Memory Network Model," Signal, Image and Video Processing, vol. 14, pp. 1255-1263, 2020.

[30] Y. D. Daydulo, B. L. Thamineni, and A. A. Dawud, "Cardiac Arrhythmia Detection Using Deep Learning Approach and Time Frequency Representation of ECG Signals," BMC Medical Informatics and Decision Making, vol. 23, article no. 232, 2023.