ITC 4/54

Information Technology and Control
Vol. 54 / No. 4/ 2025
pp. 1189-1205
DOI 10.5755/j01.itc.54.4.41549

Patch-Based ECG Segmentation for Arrhythmia Classification Using Hybrid Deep Learning

Received 2025/05/15 | Accepted after revision 2025/07/24

HOW TO CITE: Seenivasan, D., Sakthivel, K. (2025). Patch-Based ECG Segmentation for Arrhythmia Classification Using Hybrid Deep Learning. *Information Technology and Control*, 54(4), 1189-1205. https://doi.org/10.5755/j01.itc.54.4.41549

Patch-Based ECG Segmentation for Arrhythmia Classification Using Hybrid Deep Learning

D. Seenivasan*

Department of Computer Science and Business Systems, M. Kumarasamy College of Engineering, Karur, 639113, India.

K. Sakthivel

Department of Computer Science and Business Systems, K. S. Rangasamy College of Technology, Tiruchengode, 637215, India.

Corresponding author: dseenivasanres1@outlook.com

The human heart abnormality can be detected using electrocardiogram (ECG) signal which is an electrical activity waveform with P, QRS and T waves. ECG abnormality normally analyzed using shape, wave form peaks and time duration of waves etc. Traditional techniques require manual interpretation to recognize the hear abnormality, hence time consuming process. In this research, automated deep learning model that understand each patches in ECG wave for accurate abnormality classification is implemented. We propose novel patch segmentation with a 180-timeframe timestamps for best feature engineering and learning of ECG waveforms in the MIT-BH dataset. The segmented features support the model to focus on main features, which helps to improve the learning efficiency and prediction accuracy of machine learning and deep learning algorithms. Waves are segmented as patches in 180 time stamps, and features are extracted for learning and classification. A machine learning model random forest with SMOTE and a hybrid deep learning model CNN-LSTM with SMOTE are used to train and test the patched ECG waveform. Further to test the generalizability of model, we tested the PTBXL dataset from physio net to MIT trained CNN-LSTM model. The result shows that our proposed model on random forest acquired a maximum of 98.3 % accuracy in binary classification and 99.2% in multiclass classification. CNN-LSTM with patch segmentation achieves 96.5% accuracy. For PTBXL dataset testing we achieved 91% of accuracy.

KEYWORDS: ECG Classification, SMOTE, Machine Learning, Deep Learning, Random Forest, Patch Segmentation, CNN-LSTM.

1. Introduction

ECG signal analysis is essential for identifying heart disorders due to its ability to analyse the functioning of the heart. With the help of ECG signals, medical practitioners can locate irregularities like arrhythmias, ischemia, or other abnormalities that are not immediately apparent during an assessment. Finding these defects at an early stage, and especially knowing when to treat them, can significantly reduce the likelihood of serious consequences, such as a heart attack, stroke, or sudden cardiac arrest. Due to the improvement in deep learning methods, ECG analysis has become less time-consuming and more precise, and it has reduced the amount of human error, automating diagnoses and improving overall clinical outcomes [12].

Handling data imbalance in medical datasets is challenging, with SMOTE being a standard balancing method. Combining SMOTE with deep learning improves ECG arrhythmia prediction by enhancing feature learning and classification accuracy [19]. CNNs play a key role in analyzing ECG signals, enabling noninvasive and accurate heart disease detection. The following sections explore current methods, challenges, and future directions in deep learning-based ECG analysis [28].

Convolutional Neural Networks (CNNs) [20] have become highly used in ECG arrhythmia detection because of their competency in processing high volumes of data and identifying intricate patterns in ECG signals. Research has provided ample evidence of the capability of CNNs to classify ECG signals into different levels of arrhythmia with remarkable accuracy. The accuracy in predicting cardiac arrhythmias has risen with the application of VGG16 models [1]. Signal Pretreatment and Augmentation Functions of Data Enhancements, including noise elimination and data augmentation, are essential to achieve optimal model performance. Approximately 38% of studies focus on noise-echocardiogram filtering (ECG) signal enhancement, and 28% utilise augmentation to balance and strengthen the dataset [2, 3] Performance is assessed through intra-patient and inter-patient evaluation paradigms. A significant decrease in performance is noted within the inter-patient paradigm, signifying a further need for cross-training diverse population models. One primary challenge is a model's ability to cross-generalize for different patients. The performance gaps widen in intra-patient paradigms, requiring stronger algorithms for feature extraction and segmentations on various ECGs.

Although deep learning techniques provide a high degree of accuracy, there is often a lack of accessibility to understanding the reason for a prediction, which hinders clinical settings requiring explanatory frameworks [4]. Even though more techniques are presented with a success rate, applying deep learning models in clinical practice still undergoes various results. This is because of problems like large annotated dataset requirements, model explainability, and the development of appropriate frameworks for non-technical healthcare staff [27].

Hybrid deep learning architectures, like CNNs paired with RNNs, show promise for improving arrhythmia detection and classification [31]. Affordable wearable ECG devices with real-time monitoring could enable timely diagnosis and intervention Collaborative data-sharing approaches may enhance model accuracy and generalizability, though challenges in generalization and model interpretability remain. The major research gaps from existing studies are identified as they are not close to local patterns observation. Moreover, segmentation is not widely concentrated by ECG classification models. This research focuses on local pattern extraction by patch based 180-time stamp segmentation technique which is unique contribution of this research.

The contribution of the research is as follows,

- MIT-BIH is essentially an imbalanced dataset. We have balanced it using SMOTE techniques at various data computation levels. Class level imbalances are performed.
- Novel Patch-based data segmentation with 180 time stamps is used for efficient ECG feature extraction. This prepares the data for model fit and improves the model learning and prediction rates.
- The machine learning and hybrid deep learning models are used to predict the arrhythmia in ECG waveform.

Remaining section of this article is arranged with five sections. Section 2 discusses various previous research works, Section 3 describes the proposed processing steps. Section 4 shows experimental evaluation and section 5 shows conclusion of the research.

2. Related Studies

Recent deep learning advancements in ECG analysis improve automation, accuracy, and customization, aiding in diagnosing conditions like myocardial infarction. The paper [6] presents a machine learning approach for early detection of congenital heart disease through ECG signal analysis. Challenges include the diversity of the datasets, computational needs, methodological scalability, and broader application usability.

Deep learning techniques, such as CNNs and RNNs, are employed for ECG anomaly detection to enhance accuracy and efficiency in diagnosing heart conditions. Challenges include dataset diversity and limited computing resources. The SMOTE technique addresses dataset imbalance in detecting congestive heart failure, improving prediction accuracy by extracting additional features from various data sources [16]. Deep learning models, along with correlation-aware SMOTE, have been shown to enhance predictions of cardiovascular diseases, with CNNs and LSTMs performing better when handling imbalanced datasets. However, the high computational demands of these models and their generalizability across different datasets remain significant challenges.

This study enhances heart failure prediction accuracy using SMOTE methods and integrates machine learning techniques [23]. It also proposes a focal-based deep learning model for multi-class arrhythmia detection, combining SMOTE variants and dynamic time warping (DTW) for improved real-time ECG analysis [8].

This research [11] develops hybrid CNN-BLSTM models for ECG classification, using SMOTE to handle imbalance and evaluating performance on the MIT-BIH dataset. It shows high precision and recall, supporting end-to-end ECG analysis. Additionally, it reviews deep learning applications in heart sound (PCG) analysis for automated cardiac disease detection [9].

This study explores using Random Forest [15] combined with Neural Networks for ECG signal and heart disease analysis, highlighting the synergy between deep learning and ensemble methods to enhance diagnostic accuracy. A comparative review [14] analyzes deep learning techniques for arrhythmia classification across various ECG datasets, evaluating CNNs, RNNs, and hybrid architectures. Key challenges include high computational costs, dataset variability, and ensuring scalability in real-world applications.

The study in [33] uses sample generation and an augmented attention module to apply deep representation learning for balanced ECG data classification. It introduces resampling to equalize class distributions. Additionally, it advances medical imaging diagnosis by proposing a novel LVH grading method using deep learning and Explainable AI, leveraging VGG16 and ResNet with echocardiography and patient data.

The study [13] classifies Parkinson's disease using SMOTE-ENN and fusion-based feature selection with classifiers like Random Forest and SVM, achieving high accuracy. Paper [21] introduces multi-modal emotion detection with EEG and ECG signals for improved emotion recognition. Additionally, discusses AdaBoost and XGBoost for cardiovascular disease detection, while [17] enhances early coronary artery detection using deep learning and ECG signals.

To detect cardiovascular diseases, the CNN-BiLSTM model on ECG images is used [18]. It discusses the application of CNN to feature extraction and bidirectional long-short-term memory (BiLSTM) networks for temporal dynamics, focusing on high classification accuracies. It also applies learning from deep patterns [30] for cardiovascular disease (CVD) detection with classification and feature extraction methods. Finally, it discusses the application of deep learning algorithms to enhance the efficiency of the diagnosis. Architecture can be integrated with security systems for data privacy [6, 24, 25].

This study presents a compact deep learning feature encoder for ECG signals using separable convolutions. It focuses on arrhythmia detection with low-resource encoding [10] and uses Bi-LSTM and

CNN for classification. Additionally, [5] explores identifying personality traits from ECG spectrograms using ResNet-18 and ViT.

The research gap in ECG waveform analysis relies on generalizability. Most models are developed to handle specific problems, like improving accuracy and identifying abnormalities at a high cost. However, significant techniques fail to cover the generalizability of the feature engineering of ECG abnormality detection. This research focuses on generalisable solutions in feature engineering and learning models.

3. Methodology

The data handling and preparing the data for training the model play a vital role in providing optimal accuracy. ECG is a waveform, where every single wave movement needs to be considered. Medical data diagnosis becomes sensitive; even a tiny misdiagnosis report leads to wrong treatment procedures. This section explains data preparation, handling, feature extraction and classification in detail. The overall architecture is given in Fig. 1. The database signals are extracted, and noises are removed. Further, the novel 180 segmentation ideology is studied from various segmentation theories and implemented in this research to increase the optimal outcome. A segmented image is validated, and it's found that normal data is large and abnormal data is much less when compared with normal data. Based on various data sizes, intensity and image clarity, this research chose the 180 segmentation technique for optimal outcome.

The methodology involves preprocessing ECG signals, extracting relevant features using 180 segments of the single waveform, and employing a hybrid deep learning model that integrates CNN with Long Short-Term Memory (LSTM) networks for heartbeat classification. This study uses a Random Forest Classifier with preprocessing techniques like filtering, normalization, and segmentation to extract heartbeats. It combines CNN for spatial feature extraction and LSTM for temporal dependencies, enhancing accuracy with additional features. SMOTE-balanced data ensures fair representation, with performance evaluated using precision, accuracy, recall, and F1-score for heartbeat classification.

3.1. Dataset Details

We retrieved the ECG recording raw dataset from the following link https://physionet.org/content/mitdb/1.0.0/.

The MIT-BIH Arrhythmia Database includes 48 half-hour excerpts of two-channel ambulatory ECG recordings arranged chronologically from 47 subjects by the BIH Arrhythmia Laboratory between 1975 and 1979. Twenty-three recordings were randomly selected from a collection of 4000 24-hour ambulatory ECG recordings gathered from a mixed inpatient (approx 60%) and outpatient (approx 40%) population at Boston's Beth Israel Hospital. The remaining 25 recordings were chosen from the same set to include less common clinically significant arrhythmias that would not be well represented in a small random sample. The recordings were digitised at 360 samples per second per channel and 11-bit resolution over a 10 mV range. Each record underwent independent annotation by two or more cardiologists, with inconsistencies being resolved. For every beat, the computer-readable reference annotations (which total approximately 110,000 annotations) included with the database underwent refinement so that they could be labelled for each beat included with the database.

The dataset was classified into two subcategories: 80% of the data for training and 20% for testing the model. The regular and abnormal ECG data are well-balanced before the model is trained and fitted to the hybrid convolutional neural network model. This makes our system more appropriate with supervised learning techniques and generalisation skills. In this study, we developed novel data handling and segmentation techniques for addressing the essential features in the ECG signals. Finally, features are extracted and processed using the CNN-LSTM model for better accuracy. The arrhythmia class distribution is shown in Table 1.

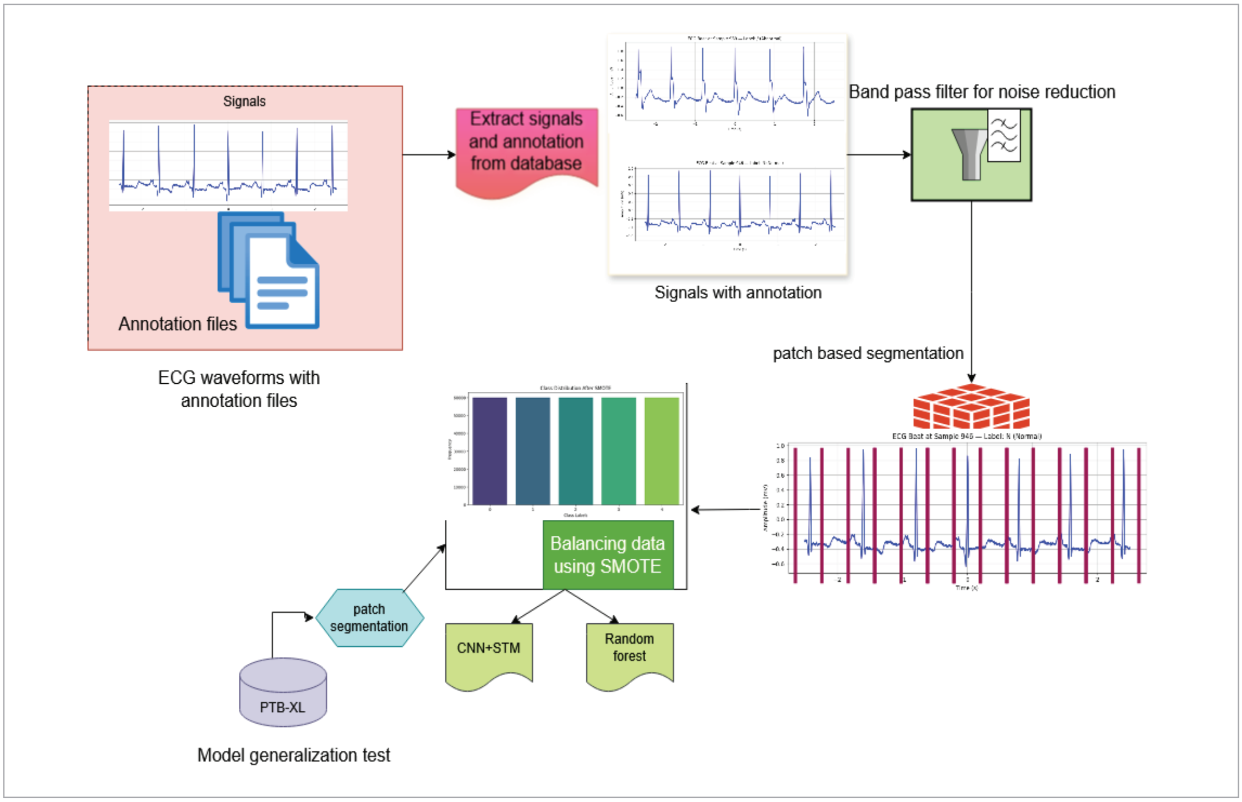
Table 1

Classes of arrhythmia classification used in this research.

Symbol	definition	class level
N	Normal beat	Class 0
L	Left bundle branch block beat	Class 1
R	Right bundle branch block beat	Class 2
A	Atrial premature beat	Class 3
V	Ventricular premature beat	Class 4

Figure 1

Proposed Patch based CNN-LSTM for ECG Classification.



3.3. Preprocessing Techniques

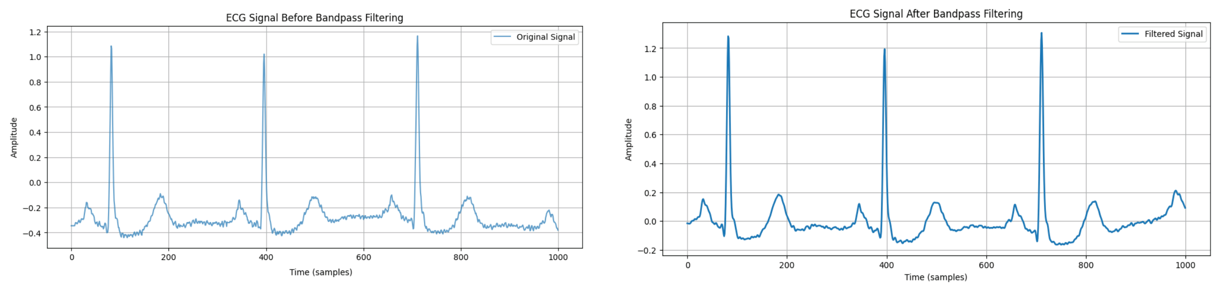
To make sure the data is clean, we need to perform preprocessing. Band pass filtering, normalization, and patch segmentation were the three preprocessing methods used in this Research investigation. A Butterworth bandpass filter with a passband of 0.5–50 Hz was used to eliminate undesired frequency components and noise from the raw cardiac signal. High-frequency artefacts like powerline interference (50/60 Hz) and baseline drift (low-frequency noise) are reduced with this filtering technique. The Butterworth filter was selected because it ensures minimal signal distortion while successfully maintaining crucial cardiac properties due to its maximum flat frequency response in the passband.

Normalizing ECG signals using a Standard scalar requires that the data be transformed to have a mean of zero and a standard deviation of 1. This step is essential for ECG signal transformation because features with different scaling are prevented from contributing equally to the model. The Standard scalar method was employed to standardise the signal amplitude variations. This technique transforms the data into a zero mean and unit variance, ensuring that different signals are comparable in magnitude. Normalisation is particularly useful in machine learning applications, where varying scales across features can negatively impact model performance. This procedure is essential in standardising ECG signals because it prevents features with different scales from contributing inequitably to the model.

Individual heartbeats were extracted from the signal using R-peak annotations. A fixed window centered around each R-peak was used to segment the complete waveform of each heartbeat. This segmentation process helps feature extraction by ensuring that each heartbeat is analyzed independently, improving model accuracy in detecting patterns related to heart disease.

Figure 2

(a) Signal with noise and (b) bandpass out come by removing noise.



(a) (b)

3.4. Data Balancing Using SMOTE

The MIT-BIH Arrhythmia Dataset is imbalanced, with normal beats far more common than other arrhythmias, leading to biased machine learning models. This class imbalance affects model performance, favoring the majority class. To address this, the Synthetic Minority Over-Sampling Technique (SMOTE) is used [3, 5], and a comparative analysis of class distribution before and after balancing is provided. Figure 3(a) shows original distribution and Figure 3(b) shows balanced outcome of smote.

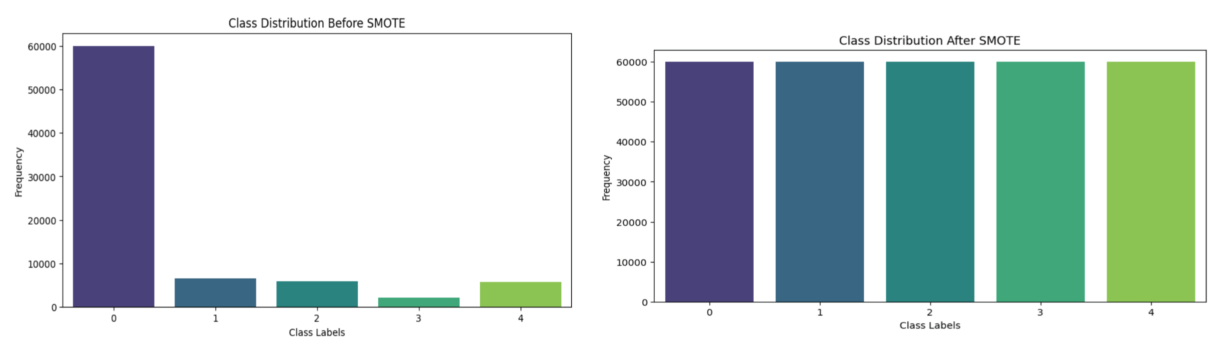
The binary classification includes normal and arrhythmia class. The multi classes have five classes: N (Normal), L (Left bundle block), R (Right bundle block), A (Arterial contraction), V (Ventricular contraction). The deep data analysis gives five classes of arrhythmia in MIT data. 1,10,000 annotation file has high class imbalances. Before applying SMOTE, the ECG dataset showed a significant class imbalance, with normal heartbeats 82,500 outnumbering other four abnormal ones like PVCs and ectopic beats. This imbalance can hinder model training, as the classifier may focus on common patterns and overlook rare ones. After applying SMOTE, the minority classes were synthetically augmented to achieve a more balanced representation with each class 60,000. Totally we have 3,000,00 data for training and testing. Train and test split of 80:20 which has 2,40,000 data for training and 60,000 samples for testing. In cross validation also, same count applies.

3.5. Proposed Patch segmentation with CNN-LSTM model

Feature extraction plays a crucial role in improving the performance of deep learning models by identifying key patterns in ECG signals. This study used

Figure 3

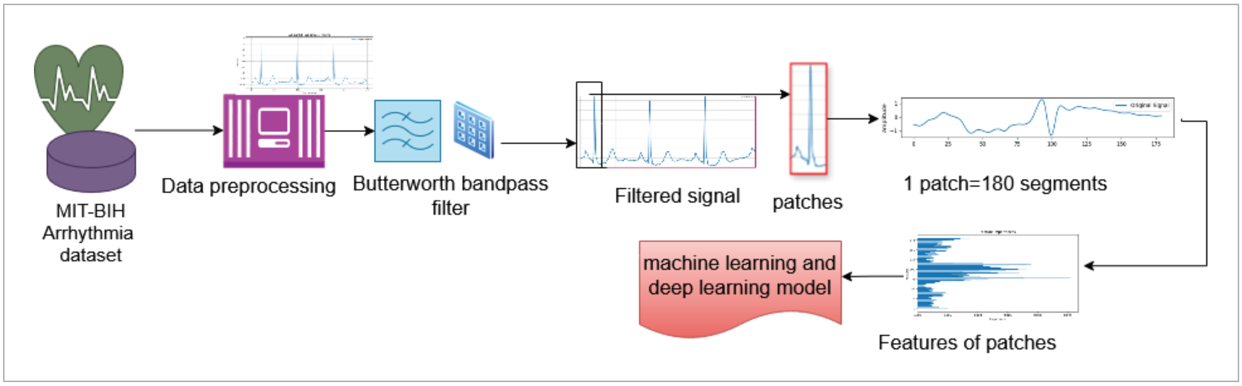
(a) Before Smote Balancing of five classes (b) Smote Balancing of five classes.



(a) (b)

Figure 4

Novel patch-based segmentation process for the feature engineering process.



heartbeat segmentation and visualization techniques to analyze and interpret heartbeat morphology effectively. While segmenting ECG signals, we consider optimal feature selection without missing a single waveform signal. Initially, we fix a 100-time stamp window for a single waveform. 1 millisecond for one feature is considered. While testing this, some features are missed and overlapped.

This is ECG data, where each second of data is sensitive for result prediction. To enhance the optimal outcome, we again tried with a 150-time stamp window, which still needs to increase the time stamp for the optimal results. Finally, the 180 timestamp window provides optimal features for the model. Table 1 below describes the results of various time stamps. Figure 4 shows the Novel patch-based segmentation for the feature engineering process.

In this research, after meaningful cross study in Table 2, features from ECG signals are extracted using 180 time stamp segments. Individual heartbeats were segmented using R-peak annotations. The R-peak, representing the most prominent point in

Table 2

Trial and error observation of various window size segments.

Window Size	Features Learned	Potential Issues
100,150	– It may capture only part of a heartbeat, missing essential features like the QRS complex, P-wave, or T-wave. – This could lead to the loss of critical diagnostic information, reducing classification accuracy.	– There is a risk of incomplete beat segmentation, affecting feature extraction and model performance. – Increased noise sensitivity due to fewer data points. – Reduced ability to differentiate between similar arrhythmias.
180	– Provides a balanced window that captures the whole cardiac cycle (including P-wave, QRS complex, and T-wave) for most heart rates. – Suitable for distinguishing between different heartbeat types. – Maintains enough temporal information for feature extraction and classification.	– Minimal risk as it effectively captures complete heartbeats while maintaining efficiency.
190,200	– Captures additional context before and after the heartbeat, which may be helpful in analysing arrhythmias	– Can introduce redundant information that is unnecessary for classification. – Higher computational cost due to increased input size. Overlapping beats may introduce unwanted interference.

Figure 5

180 patches of ECG signal segmentation sample output representation.

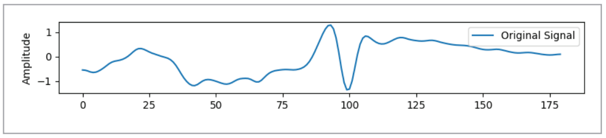
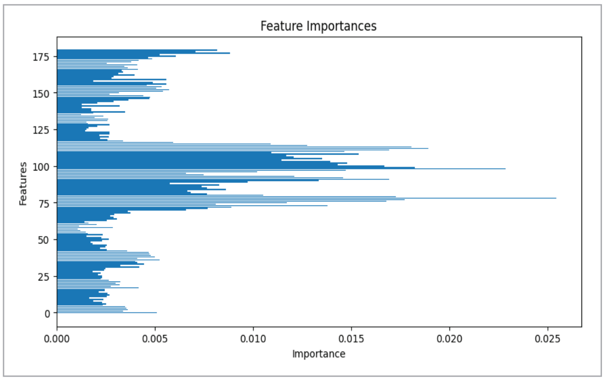


Figure 6

Feature representation each time second in 180 segments.



Algorithm: Patch Segmentation of ECG Signal

Input:

ecg_signal: List of raw ECG signal data

r_peak_annotations: List of indices representing R-peaks in the ECG signal

window_size: Length of each segment (default = 180)

Output:

segments: List of ECG signal patches centered around each R-peak

Begin

Initialize an empty list: segments ← []

For each r_peak in r_peak_annotations do

start_index ← r_peak - (window_size // 2)

end_index ← r_peak + (window_size // 2)

If start_index ≥ 0 and end_index < length of ecg_signal then

segment ← ecg_signal[start_index : end_index]

segments.append(segment)

End If

End For

Return segments

End

---

the QRS complex, is a reliable reference for heartbeat extraction. A fixed window (180) around each R-peak was used to capture the complete cardiac cycle, including the P wave, QRS complex, and T wave. This segmentation ensures that each heartbeat is analyzed separately, preserving essential morphological features required for classification. Proper segmentation is vital, as inaccuracies in extracting heartbeats may lead to incorrect classifications and affect model performance. Figure 5 and 6 shows patch based segmentation process.

3.6. CNN-LSTM MODEL

This research uses CNN for feature selection and LSTM for tracking long-term dependencies in ECG classification. The CNN automatically extracts features from 180-segmented ECG image data, focusing on key waveform patterns like P-wave, QRS complex, and T-wave. These wave patterns vary across patients, requiring effective feature extraction by CNN to account for differences in amplitude, duration, and shape.

The CNN's convolutional layer applies filters to raw ECG waveforms, identifying local features like edges, peaks, and patterns. These filters can detect key structures such as the R-peak of the QRS complex or the beginning of the T-wave, which occur within small time windows of the signal (time-domain).

Stride and filters convolve features to create feature maps, while padding is added to preserve important patterns at the signal's borders. The convolution process produces feature maps that highlight specific structures like the P-wave, QRS complex, and T-wave, capturing key local patterns within the ECG signal.

Advanced layers in the CNN-LSTM model extract and merge ECG features like P-wave, QRS complex, and T-wave to distinguish normal and abnormal rhythms. Conv1D captures local features, max pooling highlights key peaks, and LSTM models temporal dependencies. The dense layers learn high-level patterns, and the final layer classifies rhythms, supporting accurate arrhythmia and heart disease detection.

The LSTM layer processes sequential relationships between local features extracted by the CNN, capturing long-term dependencies in the ECG signal. It tracks dependencies between key features like the R, QRS, and T waves, using its gates to filter relevant information, improving the model's ability to handle time-dependent data and rhythm abnormalities.

3.7. Layers Architecture

This study uses a CNN-LSTM model to classify pre-processed and segmented ECG signals. Trained on a large dataset with the Adam optimizer (learning rate 0.0001), the model effectively identifies heart conditions. It uses Conv1D with 64 filters (3x3 kernel, ReLU) to extract features, and performance is evaluated using accuracy, precision, recall, and confusion matrix.

$$V_{i,j} = ReLU(\sum_{L=1}^{L} X_{i,j+L} \,.\, w_{j,L} + b_j, \tag{1}$$

where $V_{i,j}$ represents feature map, $X_{i,j+L}$ represents the input sequences, $w_{j,L}$ is weights of filter and bias b_j with kernel size L=3. Activation function represented as $ReLU$(m),

$$ReLU(Z) = \max!(0, Z) \tag{2}$$

Batch normalization layer normalize each convolutional layer output to ranges from 0 to 1.

$$\hat{a}_{i,j} = \frac{a_{i,j} - \mu_j}{\sqrt{\sigma_j^3 + e}} \;, \tag{3}$$

where $\hat{a}_{i,j}$ is a feature map, and j is the mean and variance of the feature map. e is considered as a sample constant to handle division by zero.

MaxPool1D- A Pooling operation by choosing the maximum values is performed in this layer. It reduces the dimensions of the feature map obtained from the previous convolution layer by choosing the maximum value in each feature map with window size 2.

$$Max_{pool}\left(Xp(i,j)\right) = \; max\left(a_i, 2_j, b_i, 2_j + 1\right), \tag{4}$$

where $Xp(i,j)$ is the outcome of max pooling operation.

The second block of Conv1D- this layer is composed of 64 filters with 3 kernel filter sizes. Another convolutional operation is similarly performed. The final feature map is fed as input to the LSTM model.

The layered architecture of the LSTM is explained below.

The LSTM layer captures the temporal dependencies of ECG signals, learning long-term patterns beneficial for sequential data. Our model uses 50 memory units. With "return sequences" set to true, it returns the entire sequence when stacking multiple LSTM layers. The convolutional and LSTM layers extract features, which are then classified by a fully connected dense layer. The output layer has five neurons for multiclass classification (normal, supraventricular ectopic beat, ventricular ectopic beat, fusion beat, and unknown beat) and one neuron for binary classification (normal vs. abnormal ECG).

The letters n. While Xt indicates the network's input, h_{t-1}, represents the unit's outputs at the current time and the time before, respectively.

Four gates such as forget, input, cell state, output, and hidden state are used by LSTM.

Figure 4 illustrates the internal organization of an LSTM memory block. The mathematical computation of features is as follows: Nt stands for the current and previous moment's neuronal states, respectively. h_{t-1} are the outputs of current time units and preceding time units respectively while Xt denotes input to the network?

The LSTM forget gate is O_t, which controls the forgotten information by means of the sigmoid function. it is the input gate In_t, which sets threshold value then applies tanh function to decide on the state of the neuron. Op_t is the output gate which reguletes the output information with sigmoid function. Figure 7 shows CNN-LSTM for ECG classification model. The model is easy to compute and it takes only 450sec to compute the data and it seems its light weight model.

Table 3 explains hyper parameter involved in architecture. Their Mathematical norms closer to, are:

$$O_t = sigmoid(j_o.\left[h_{t-1,} x_t\right] + V_o) \tag{5}$$

$$In_t = sigmoid(j_{in}.\left[h_{t-1,} x_t\right] + V_{in}) \tag{6}$$

$$Op_t = sigmoid\left(j_{op}.\left[h_{t-1,} x_t\right] + V_{op}\right) \tag{7}$$

$$N_t^{\cdot} = tanh\;(j_n.\left[h_{t-1,} x_t\right] + V_n) \tag{8}$$

Figure 7
CNN-LSTM for ECG classification.

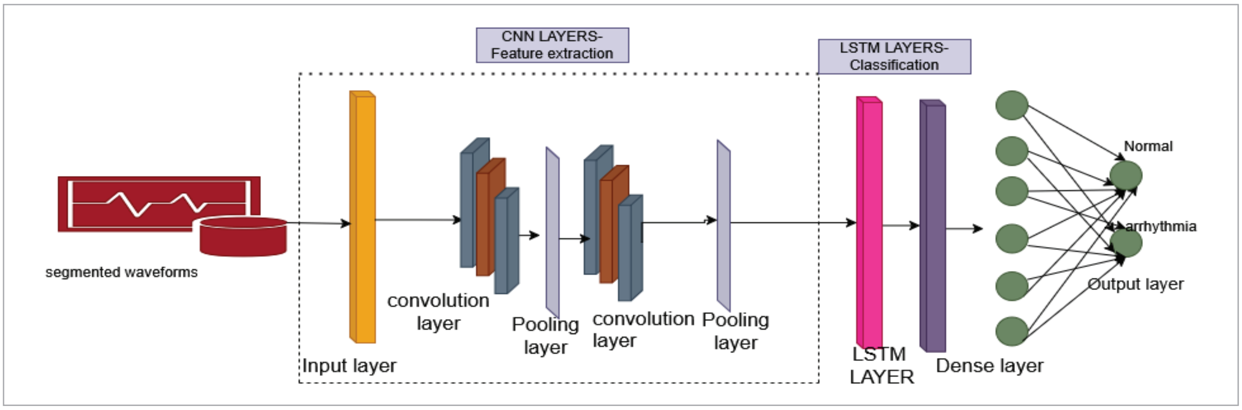


Table 3
Hyperparameters used for the deep CNN-LSTM hybrid architecture for the detection of heart conditions using ECG signals.

Layer Number	Layer Type	Layer Parameters	Output Shape	Number of Parameters
Layer 0	Input	No parameters	(None, 176, 1)	0
Layer 1	Conv1D	Number of filters = 32; Filter size = (3,); Strides = (1,); Activation = ReLU	(None, 176, 32)	192
Layer 2	MaxPooling 1D	Pool size = (2,); Strides = (2,)	(None, 88, 32)	0
Layer 3	LSTM	Number of units = 50; Activation = tanh	(None, 88, 50)	16,600
Layer 4	Flatten	No parameters	(None, 4400)	0
Layer 5	Dense	Number of hidden neurons = 100; Activation = ReLU	(None, 100)	440,100
Layer 6	Dense	Number of neurons = 5; Activation = Softmax or Sigmoid (depends on task)	(None, 5)	505
-	Dropout	0.3	-	-

Total Parameters: 457,399 (1.74 MB)
Trainable Parameters: 457,397 (1.74 MB)
Non-Trainable Parameters: 0
Optimizer Parameters: 2 (12 bytes)

3.8. Random Forest Classifier

Random Forest is an ensemble learning method that uses decision trees to classify arrhythmia. It builds multiple trees from bootstrapped samples, with each tree voting on the final classification. Key parameters such as the number of trees, tree depth, and leaf samples are tuned for optimal performance.

4. Experimental Setup

The model was developed using TensorFlow 2.3.0 and Python 3.8, with hardware support from an NVIDIA GeForce GTX 1060 GPU. Five-fold cross-validation [26] was employed for robustness testing, splitting the dataset into five sections for training and testing purposes. Model evaluation metrics—specificity, sensitivity, and accuracy—were averaged over five cycles, with a total training time of 20 minutes and an average testing time of 0.68 seconds per model.

We chose the Adam optimizer with backpropagation, set the learning rate of 0.001 for each round of training fold, trained for 60 epochs, and set the maximum mass size to 32. Table 4 provides Ablation of the research.

Table 4

Ablation of the research work.

MODEL	Accuracy%	Precision%	Recall%	F1 score%	AUC score%
RF	89.21	89.62	88.21	88.35	89.18
RF+SMOTE	93.51	93.26	92.78	92.43	92.94
Random forest Classifier+ Patch segmentation +SMOTE (2 categories)	98.36	98.14	96.91	97.8	97.99
RF Classifier + patch segmentation + SMOTE (5 categories)	98.96	98.95	98.96	98.7	99.80
CNN-LSTM	79.15	79.14	78.82	78.77	79.92
CNN	80.6	81.24	80.32	80.94	80.83
CNN-LSTM + SMOTE	82.37	82.46	81.73	82.03	82.26
CNN + LSTM with SMOTE + Patch segmentation	85.50	86.39	85.50	85.91	92.08
CNN + LSTM with SMOTE (applied for both test and train data) + patch segmentation	96.60	96.68	96.60	96.58	99.52

various models, including accuracy, precision, recall, F1 score, and AUC score for four classifications. The Random Forest with binary classifier achieved 98.36% accuracy, 98.14% precision, 96.91% recall, 97.8% F1 score, and a 97.99% AUC score, demonstrating the highest effectiveness in class separation.

The Random Forest Classifier (5 categories) performance is with an accuracy of 98.96%, precision of 98.95%, and recall of 98.96%. RF achieved an F1 score of 98.7 and an AUC score of 99.80 which marks better performance for multi-category prediction whilst retaining extremely high accuracy and precision.

The model with CNN and LSTM together with SMOTE outcome is discussed. The model combines these techniques while using SMOTE for class balancing. The model performs reasonably well with an accuracy of 85.50%, precision of 86.39%, recall of 85.50%, F1 score of 85.91, and AUC of 92.08, but as stated before, is not at the same level of the Random Forest models.

The final model greatly exceeds the previous CNN LSTM model, achieving an accuracy of 96.60% along with precision, recall, F1 score and AUC of 96.68, 96.60, 96.58, and 99.52 respectively. This indicates that applying SMOTE on both the train and test data produces better class outcomes. Using SMOTE on both train and test data suggests that the model is capable of being more generalized which results in improved classification capability throughout the data and overall performance of the model. The Random Forest Classifiers excel in most metrics, especially in multi-class problems, where they outperform the CNN + LSTM models. The application of SMOTE within the CNN + LSTM models does assist in boosting the performance, particularly when the model is tested on the training and test data. Still, the Random Forest models remain stronger, particularly in multi-class classification challenges.

Figures 8-9 show the accuracy and loss curve of the baseline vs proposed approach. The CNN-LSTM model in Figure 7 with automatic feature extraction and learning reaches accuracy of 90 -91% in 20 epochs, whereas loss rate is 0.20-0.25 which implies has some learning problem. The proposed patch based novel feature extraction with Figure 8 shows accuracy of training and validation up to 96% and loss rate 0.06-0.09 which implies proposed achieves better classification accuracy with patch based segmentation and feature engineering. Table 5 shows patch wise model performance. The 180 patches achieve less noise patterns which helps to gain 96% accuracy.

Figure 8

CNN-LSTM outcome without patch segmentation.

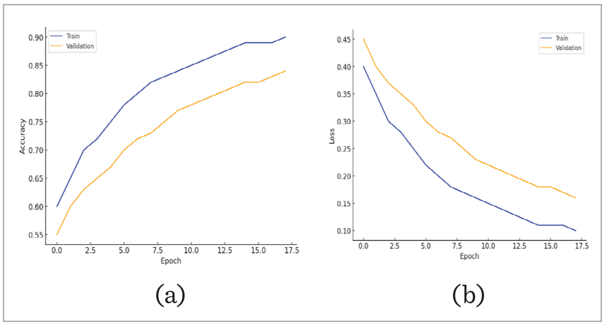


Figure 9

CNN-LSTM outcome with patch segmentation: (a) accuracy and (b) loss

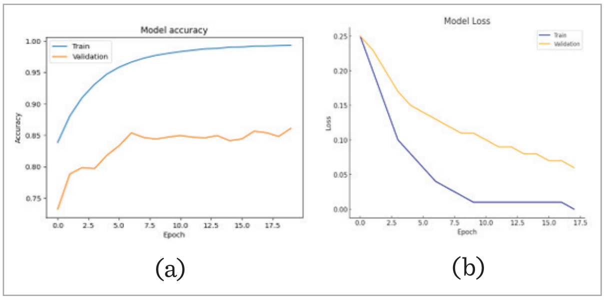


Table 6 shows the performance of the model using five-fold cross-validation, with average accuracy of 85.4%, precision of 90.8%, recall of 89.2%, F1 score of 90%, and specificity of 89.3%. After SMOTE resampling (Table 7), the average accuracy increased to 96.4%, precision to 97.2%, recall to 98%, F1 score to 98%, and specificity to 98%, demonstrating improved model efficiency on unseen data.

Table 5

Performance metrics of various patch size evaluation.

Patch size	Accuracy in %	Precision in %	recall	in %F1-score in %
100	93.2	92.1	92.67	92.46
150	90.8	94.1	93.33	94.68
180	96.6	96.68	96.6	96.58
190	94.7	93.52	93.85	93.11
200	93.3	90.48	90.59	91.36

Table 6

Five-fold cross-validation performance metrics.

K-Fold	Accuracy Per Fold in %	Precision in %	Recall (Sensitivity) in %	F1 Score in %	Specificity in %
1st Fold	85.72	0.9090	0.8928	0.9009	0.8888
2nd Fold	85.10	0.8888	0.8727	0.8807	0.875
3rd Fold	85.16	0.8990	0.8828	0.8909	0.8817
4th Fold	85.41	0.9189	0.9026	0.9107	0.9052
5th Fold	86.09	0.9285	0.9122	0.9203	0.9166
Average	85.49	0.908	0.892	0.90	0.893

Table 7

Five-fold cross-validation performance metrics after SMOTE resampling in test dataset.

K-Fold	Accuracy in %	Precision in %	Recall in %	F1 Score in %	Specificity in %	ROC AUC
1st Fold	96.66	0.972	0.98	0.976	0.98	0.99
2st Fold	92.6	0.966	0.946	0.973	0.986	0.987
3st Fold	92.45	0.874	0.985	0.899	0.974	0.982
4st Fold	94.89	0.94	0.99	0.99	0.971	0.97
5st Fold	96.4	0.97	0.983	0.987	0.97	0.992

Table 8

Statistical analysis.

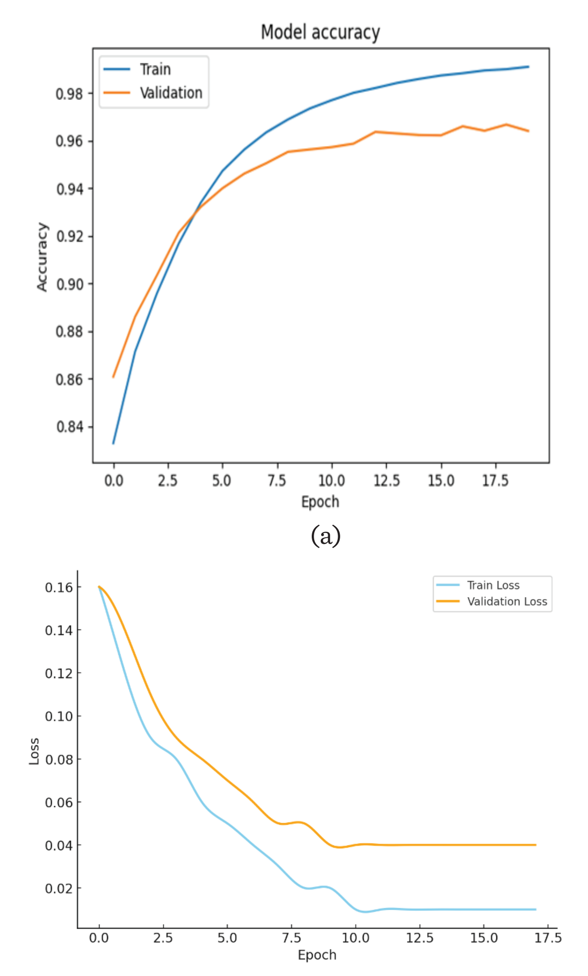
Statistical model	Statistical value	p-value:
Friedman Test	4.0000	0.4060
t-Test	1.0982	0.3523
Holm method	----	0.3523

The performance of the random forest and CNN-LSTM models was tested using statistical methods, as shown in Table 8. The Friedman test, comparing related groups, yielded a p-value of 0.4060, indicating no significant difference between the models. The t-test showed a p-value of 0.3523, which is greater than 0.05, suggesting no significant difference in mean accuracy. The Holm method confirmed no significant differences, with a p-value of 0.35. Therefore, both models perform equally well, with no substantial evidence that one outperforms the other

We balanced the test data during computation to improve the imbalances in the test data. SMOTE-based oversampling is done. Figures 10(a)-10(b) shows the model's accuracy and loss improvement over each epoch. Initially, overfitting occurs; after the 5th epoch, the model fits the training data, and validation accuracy improves over time.

Figure 10

(a) CNN-LSTM with smote balancing on test data.
(b) CNN-LSTM loss with smote balancing on test data.

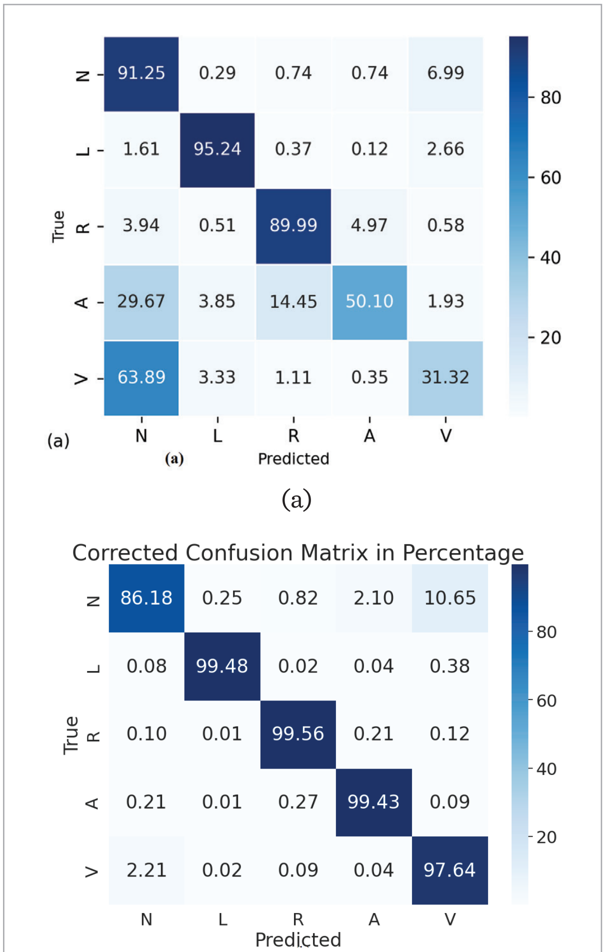


(a)

(b)

Figure 11

Confusion matrix for CNN-LSTM performance metrics
(a) before applying smote and segmentation
(b) with smote balancing on test data.

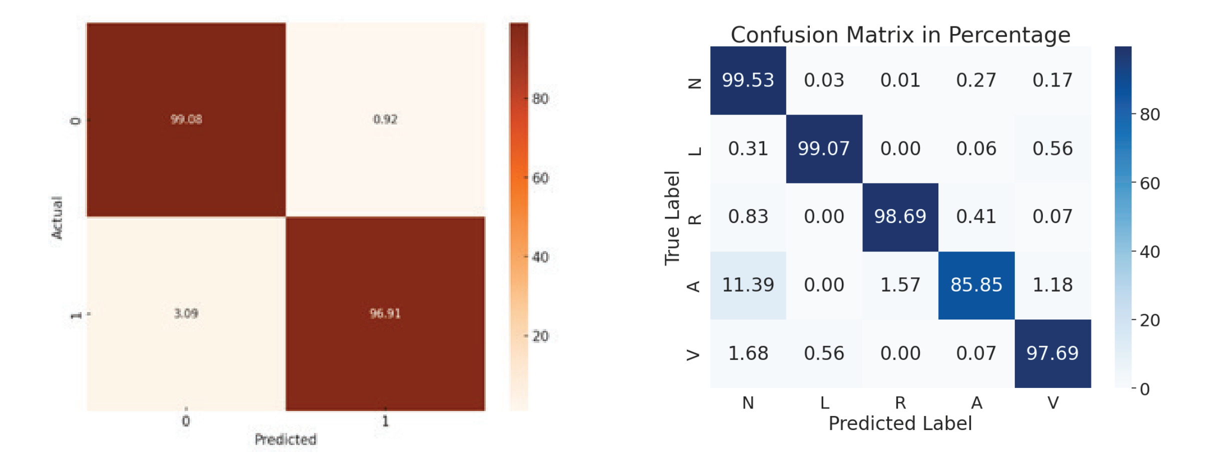


(a)

(b)

Figure 12

(a) confusion matrix performance of a Random Forest classifier on a binary classification task.
(b) five categorical classification performance in the confusion matrix for the random forest algorithm.



(a) (b)

From the confusion matrix in Figure 11(a), we can understand that the model is performing well with the "N", "L", and "R" categories, scoring 91.25%, 95.24%, and 89.99%, respectively, along the diagonal, indicating strong classification performance. However, there are some misclassifications in the ″A″ class where 29.67% of the samples are classified as ″N″ and in the ″V″ class where 63.89% are classified as ″N″. Despite the model being able to separate most classes, some values outlier from the diagonal value of zero indicating some confusion between classes; predominately, the "A" and "V" classes in Figure 11(b)

4.1. Performance Metrics by Random Forest Classifier

Figure 12 (a) shows a confusion matrix illustrating the performance of a Random Forest classifier on a binary classification task. It performs very well, with the model accurately predicting True Negatives (99.08%) and True Positives (96.91%). The model makes an error of classifying 0.92% of actual negatives as positives and 3.09% of actual positives as negatives.

This indicates that the classifier effectively distinguishes between the two classes.

Figure 12(b) confusion matrix illustrates a multiclass classification problem with five classes: N (Normal), L (Lidocaine), R (Rash), A (Abnormal), V (Ventricular). In this case, Random Forest achieved remarkable success with N (99.53%), L (99.07%), and R (98.69%), demonstrating high accuracy. Nevertheless, some other misclassifications do occur, particularly related to the A (Abnormal) category (11.39% misclassifications to L) and V (Ventricular) category (1.68% misclassified to N). Regardless, the model shows reasonable classification performance over all classes, although a better distinction concerning these misclassifications might be reached with some adjustments.

4.2. Testing External Data for Model Generalizability Using PTB-XL Image

To assess the model's robustness and generalizability, we used an independent dataset from the PTBXL repository for external validation, ensuring no overlap with the training set. The performance metrics from this validation highlight the model's ability to maintain accuracy and reliability across different datasets, boosting confidence in its practical applicability, as shown in Table 9.

Table 9

PTBXL data validation using proposed model

Metrics	Values in %
Accuracy	90.35
Precision	91.97
Recall	90.35
F1 score	90.92

Table 10
State of art comparison with proposed results

Reference	Datasets	Feature selection	Classification	Accuracy (%)
Atal et al. 2020 [7]	PTBXL	Wavelet with Gabor	Bat Rider Optimizer with Deep CNN	93%
Zhang et al. 2021 [32]	MIT-BIH	Rapid ramp with ECG segments	Artificial deep neural network with conv1D	94.7%
Liu et al 2024. [22]	MIT-BIH	Random horizontal flip	Lightweight CNN	92.4%
Saeed et al. 2024 [29]	MIT-BIH	Chebyshev function	ANN with 3 layers	93%
Wang et al. 2024 [30]	MIT-BIH	Residual attention	BI-LSTM	95%
Pal et al. 2024 [26]	MIT-BIH	Median +gaussion filter	Modified lightweight CNN	92%
Proposed model	MIT-BIH	Bandpass filter with patch-based segmentation	Random forest CNN-LSTM	99%

Table 9 presents a state-of-the-art comparison between the proposed approach and existing work.

5. Conclusion

This study presents a comprehensive and flexible framework for ECG waveform classification, utilizing machine learning and deep learning approaches. This research contributes to the field by integrating patch-based segmentation of ECG signals into 180-frame units, thereby increasing the temporal resolution of features. Localized feature extraction is essential for effective arrhythmia detection, thereby enhancing model generalizability. Furthermore, our approach offers scalable solutions for both binary and multi-class classification tasks, utilizing the MIT-BIH dataset. In addition, our methodology, which employs a Random Forest classifier in conjunction with a CNN-LSTM model, sequentially captures the tabular and temporal dependencies of ECG data, thereby augmenting the hybrid model literature in biomedical signal processing.

Practical Implications

As the estimation model is capable of reliable arrhythmia detection with an accuracy of 96 to 99 per cent, it poses minimal requirements for manual signal annotation in terms of segmentation. Furthermore, its capabilities enhance compatibility with worn ECG monitors as well as with bedside diagnostic tools that require temporally dependent automated decision support. From a clinical standpoint, these features are beneficial in the context of time-critical environments, such as emergency rooms or units with constrained resources. Additionally, privacy-preserving approaches are suitable within institutions that are subject to data-sharing constraints. Hence, relying less on annotated datasets alleviates the problem posed by the lack of shareable data. For model generalizability, ptbxl dataset from physionet is tested. It achieves an accuracy of 90-91% for the new dataset.

Research Limitations

In real time, the model needs to learn various diseases based on ECG patterns. We focused on abnormality, which may be limited to five arrhythmia classes. There may be more abnormality classes in real time.

Future Work

1. Pruning and lightweight compression strategies will be explored to enable deployment within real-time embedded clinical monitoring systems. Streaming telemetry systems specifically targeted towards remote cardiovascular supervision markers shall also benefit significantly from instantaneous responsiveness.
2. Evaluating the model across multiple ECG datasets and diverse healthcare settings would enhance its reliability and global applicability.

References

1. Abbas, S., Ojo, S., Krichen, M., Alamro, M. A., Mihoub, A., Vilcekova, L. A Novel Deep Learning Approach for Myocardial Infarction Detection and Multi-Label Classification. IEEE Access, 2024. https://doi.org/10.1109/ACCESS.2024.3401744
2. Aghaomidi, P., Wang, G. ECG-SleepNet: Deep Learning-Based Comprehensive Sleep Stage Classification Using ECG Signals. 2024.
3. Alex, S. A. Imbalanced Data Learning Using SMOTE and Deep Learning Architecture with Optimized Features. Neural Computing and Applications, 2025, 37(2), 967-984. https://doi.org/10.1007/s00521-024-10481-y
4. Alreshidi, F. S., Alsaffar, M., Chengoden, R., Alshammari, N. K. Fed-CL-An Atrial Fibrillation Prediction System Using ECG Signals Employing Federated Learning Mechanism. Scientific Reports, 2024, 14(1), 21038. https://doi.org/10.1038/s41598-024-71366-7
5. Altaf, M. M., Khan, S. U., Majid, M., Anwar, S. M. Personality Trait Recognition Using ECG Spectrograms and Deep Learning. In 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2024, July, 1-4. IEEE. https://doi.org/10.1109/EMBC53108.2024.10782328
6. Ameen, A. H., Mohammed, M. A., Rashid, A. N. Enhancing Security in IoMT: A Blockchain-Based Cybersecurity Framework for Machine Learning-Driven ECG Signal Classification. Fusion: Practice & Applications, 2024, 14(1). https://doi.org/10.54216/FPA.140117
7. Atal, D. K., Singh, M. Arrhythmia Classification with ECG Signals Based on the Optimization-Enabled Deep Convolutional Neural Network. Computer Methods and Programs in Biomedicine, 2020, 196, 105607. https://doi.org/10.1016/j.cmpb.2020.105607
8. Baber, A., Ahmed, F., Rehman, A. U., Javaid, S., Alshammeri, M., Mir, A., Kumar, D. Optimizing Heart Failure Predictive Accuracy: An Effective Approach Using SMOTE Techniques. In 2024 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC), 2024, July, 1-7. IEEE. https://doi.org/10.1109/ETNCC63262.2024.10767552
9. Bhatia, S., Pandey, S. K., Kumar, A., Alshuhail, A. Classification of Electrocardiogram Signals Based on Hybrid Deep Learning Models. Sustainability, 2022, 14(24), 16572. https://doi.org/10.3390/su142416572
10. Bontinck, L., Fonteyn, K., Dhaene, T., Deschrijver, D. ECGencode: Compact and Computationally Efficient Deep Learning Feature Encoder for ECG Signals. Expert Systems with Applications, 2024, 255, 124775. https://doi.org/10.1016/j.eswa.2024.124775
11. Boulif, A., Ananou, B., Ouladsine, M., Delliaux, S. Feature Fusion for Multi-Class Arrhythmia Detection Using Focal-Based Deep Learning Architecture. In 2024 18th International Conference on Control, Automation, Robotics and Vision (ICARCV), 2024, December, 435-440. IEEE. https://doi.org/10.1109/ICARCV63323.2024.10821679
12. Chen, Z., Yang, D., Cui, T., Li, D., Liu, H., Yang, Y., Zhang, S., Yang, S., Ren, T. L. A Novel Imbalanced Dataset Mitigation Method and ECG Classification Model Based on Combined 1D_CBAM-Autoencoder and Lightweight CNN Model. Biomedical Signal Processing and Control, 2024, 87, 105437. https://doi.org/10.1016/j.bspc.2023.105437
13. Farhad, M., Masud, M. Automated LVH Grading: Integration of Deep Learning and Explainable AI for Accurate Diagnosis. In Proceedings of the 2024 8th International Conference on Medical and Health Informatics, 2024, May, 212-218. https://doi.org/10.1145/3673971.3674000
14. Gour, A., Wadhvani, R., Gupta, M. Comprehensive ECG Signal Analysis for Heart Disease Classification: Integrating Neural Networks with Random Forest. In 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2024, June, 1-7. IEEE. https://doi.org/10.1109/ICCCNT61001.2024.10726130
15. Hamza, M. F. A. B., Sjarif, N. N. A. A Comprehensive Overview of Heart Sound Analysis Using Machine Learning Methods. IEEE Access, 2024. https://doi.org/10.1109/ACCESS.2024.3432309
16. Hussain, L., Lone, K. J., Awan, I. A., Abbasi, A. A., Pirzada, J. U. R. Detecting Congestive Heart Failure by Extracting Multimodal Features with Synthetic Minority Oversampling Technique (SMOTE) for Imbalanced Data Using Robust Machine Learning Techniques. Waves in Random and Complex Media, 2022, 32(3), 1079-1102. https://doi.org/10.1080/17455030.2020.1810364
17. Ibrahim, S. S., Madhu, G., Kadali, J., Salivendra, N., Anitha, B. Improving Cardiovascular Disease Detection: A Comparative Approach with Boosting Algorithms and Imputed Features. In 2024 First International Con-

ference for Women in Computing (InCoWoCo), 2024, November, 1-6. IEEE. https://doi.org/10.1109/InCoWoCo64194.2024.10863208

18. Iqbal, J. M. A Novel Deep Learning Approach for Early Detection of Cardiovascular Diseases from ECG Signals. Medical Engineering & Physics, 2024, 125, 104111. https://doi.org/10.1016/j.medengphy.2024.104111
19. Jain, P., Gajbhiye, P., Tripathy, R. K., Acharya, U. R. A Two-Stage Deep CNN Architecture for the Classification of Low-Risk and High-Risk Hypertension Classes Using Multi-Lead ECG Signals. Informatics in Medicine Unlocked, 2020, 21, 100479. https://doi.org/10.1016/j.imu.2020.100479
20. Kumar, G. K., Anila, M., Manikyam, N. R. H., Thatha, V. N., Reddy, R. V. K., Papana, K. R. Heart Failure Detection Through SMOTE for Augmentation and Machine Learning Approach for Classification. Smart Factories for Industry 5.0 Transformation, 2025, 123-134. https://doi.org/10.1002/9781394200467.ch7
21. Li, Y., Zhang, H. EEG Signal Recognition of VR Education Game Players Based on Hybrid Improved Wavelet Threshold and LSTM. The International Arab Journal of Information Technology (IAJIT), 2025, 22(1), 170-181. https://doi.org/10.34028/iajit/22/1/13
22. Liu, Y., Liu, J., Tian, Y., Jin, Y., Li, Z., Zhao, L., Liu, C. Pruned Lightweight Neural Networks for Arrhythmia Classification with Clinical 12-Lead ECGs. Applied Soft Computing, 2024. https://doi.org/10.1016/j.asoc.2024.111340
23. Mahel, A. S. B., Alotaibi, F. M. G., Rao, N. The Role of Synthetic Data in Mitigating Imbalance in Deep-Learning-Based Arrhythmia Classification: A Comparative Study. In Sixth International Conference on Image, Video Processing, and Artificial Intelligence (IVPAI 2024), 2024, September, Vol. 13225, 120-125. SPIE. https://doi.org/10.1117/12.3046225
24. Mutlag, A. A., Abd Ghani, M. K., Mohammed, M. A., Lakhan, A., Mohd, O., Abdulkareem, K. H., Garcia-Zapirain, B. Multi-Agent Systems in Fog-Cloud Computing for Critical Healthcare Task Management Model (CHTM) Used for ECG Monitoring. Sensors, 2021, 21(20), 6923. https://doi.org/10.3390/s21206923
25. Mutlag, A. A., Ghani, M. K. A., Mohammed, M. A. A Healthcare Resource Management Optimization Framework for ECG Biomedical Sensors. In Efficient Data Handling for Massive Internet of Medical Things: Healthcare Data Analytics, 2021, 229-244. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-66633-0_10
26. Pal, A. S., Mohan, H., Agarwal, S., Agarwal, N. Advanced Noise-Resistant Electrocardiography Classification Using Hybrid Wavelet-Median Denoising and a Convolutional Neural Network. Sensors, 2024, 24(21), 7033. https://doi.org/10.3390/s24217033
27. Pan, H., Li, Z., Fu, Y., Qin, X., Hu, J. Reconstructing Visual Stimulus Representation from EEG Signals Based on Deep Visual Representation Model. IEEE Transactions on Human-Machine Systems, 2024, 54(6), 711-722. https://doi.org/10.1109/THMS.2024.3407875
28. Pan, H., Tong, S., Song, H., Chu, X. A Miner Mental State Evaluation Scheme with Decision-Level Fusion Based on Multidomain EEG Information. IEEE Transactions on Human-Machine Systems, 2025, 55(2), 289-299. https://doi.org/10.1109/THMS.2025.3538162
29. Saeed, M., Märtens, O., Larras, B., Frappe, A., John, D., Cardiff, B. ECG Classification with Event-Driven Sampling. IEEE Access, 2024. https://doi.org/10.1109/ACCESS.2024.3364115
30. Wang, Y. N., Zhou, G., Yang, C. Interpatient Heartbeat Classification Using Modified Residual Attention Network with Two-Phase Training and Assistant Decision. IEEE Transactions on Instrumentation and Measurement, 2022, 72, 1-15. https://doi.org/10.1109/TIM.2022.3232646
31. Yin, J., Qiao, Z., Han, L., Zhang, X. EEG-Based Emotion Recognition with Autoencoder Feature Fusion and MSC-TimesNet Model. Computer Methods in Biomechanics and Biomedical Engineering, 2025, 1-18. https://doi.org/10.1080/10255842.2025.2477801
32. Zhang, J., Liu, A., Liang, D., Chen, X., Gao, M. Interpatient ECG Heartbeat Classification with an Adversarial Convolutional Neural Network. Journal of Healthcare Engineering, 2021, 2021(1), 9946596. https://doi.org/10.1155/2021/9946596
33. Zhu, C. Computational Intelligence-Based Classification System for the Diagnosis of Memory Impairment in Psychoactive Substance Users. Journal of Cloud Computing, 2024, 13(1), 119. https://doi.org/10.1186/s13677-024-00675-z