


Classification of cardiac arrhythmia using a convolutional neural network and bi-directional long short-term memory

Digital Health
Volume 8: 1–13
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076221102766
journals.sagepub.com/home/dhj


Shahab Ul Hassan¹ , Mohd S Mohd Zahid¹ , Talal AA Abdullah¹
and Khaleel Husain²

Abstract

Cardiac arrhythmia is a leading cause of cardiovascular disease, with a high fatality rate worldwide. The timely diagnosis of cardiac arrhythmias, determined by irregular and fast heart rate, may help lower the risk of strokes. Electrocardiogram signals have been widely used to identify arrhythmias due to their non-invasive approach. However, the manual process is error-prone and time-consuming. A better alternative is to utilize deep learning models for early automatic identification of cardiac arrhythmia, thereby enhancing diagnosis and treatment. In this article, a novel deep learning model, combining convolutional neural network and bi-directional long short-term memory, is proposed for arrhythmia classification. Specifically, the classification comprises five different classes: non-ectopic (N), supraventricular ectopic (S), ventricular ectopic (V), fusion (F), and unknown (Q) beats. The proposed model is trained, validated, and tested using MIT-BIH and St-Petersburg data sets separately. Also, the performance was measured in terms of precision, accuracy, recall, specificity, and f1-score. The results show that the proposed model achieves training, validation, and testing accuracies of 100%, 98%, and 98%, respectively with the MIT-BIH data set. Lower accuracies were shown for the St-Petersburg data set. The performance of the proposed model based on the MIT-BIH data set is also compared with the performance of existing models based on the MIT-BIH data set.

Keywords

Arrhythmia, convolutional neural network, accuracy, bi-directional long short-term memory, precision, electrocardiogram, classification.

Submission date: 29 November 2021; Acceptance date: 8 May 2022

Introduction

Cardiovascular disease (CVD) is the leading cause of death globally. As per the World Health Organization (WHO), over 17.9 million humans around the world have died as a result of CVD diseases.^{1,2} Early diagnosis of CVD is critical to preventing sudden death from a heart attack or cardiac arrest. Cardiac arrhythmias refer to a group of disorders in which the heart's electrical impulse is abnormal, resulting in a quicker or slower beat than usual.³ A thorough investigation of the electrocardiogram (ECG) segment offers structural instruction about cardiac patients, widely employed in clinical procedures for arrhythmia identification. Usually, the ECG signs of cardiac disease do not

appear within a short ECG recording period. They require a prolonged recording and monitoring of more than one day. This lengthens and complicates cardiologists' interpretation of ECG charts. Thus, numerous advancements in recent years have been made to ECG signals to decrease

¹Department of Computer and Information Sciences, Universiti Teknologi PETRONAS, Malaysia

²Institute of Health and Analytics, Universiti Teknologi PETRONAS, Malaysia (Until August 2021)

Corresponding author:

Mohd S Mohd Zahid, Department of Computer and Information Sciences, Universiti Teknologi PETRONAS, Perak, Malaysia.
Email: msoperi.mzahid@utp.edu.my



mortality and assist cardiologists in making timely, efficient, and accurate decisions. There are two processes for assessing ECG characteristics in the traditional manual method. The first stage involves extracting ECG features, while the second stage categorizes ECG based on the retrieved characteristics.⁴ The process is cumbersome and error-prone for the cardiologists, and there is a need for automated ECG classification. Therefore, early recognition of cardiac arrhythmia is critical to effective investigation and treatment.

Several machine learning-based techniques for extracting heart characteristics and training models for arrhythmia identification have recently been developed.⁵ Linear predictive coding,⁶ wavelet entropy,⁷ synchro-squeezing wavelet transform,⁸ k-nearest neighbor and support vector machine models have all been utilized in predicting arrhythmia. While these non-deep learning techniques perform well, they suffer from various constraints, including poor classification performance for large data.

Thus, several deep learning methods have been applied recently to address a variety of difficult problems across all disciplines of health care research, including ECG classification.^{9,10} Deep learning methods transcend the limitations of traditional disease diagnosis, enhancing performance and generalization by reducing pre-processing and feature extraction.^{11,12} In this context, only a few studies on convolutional neural networks (CNNs),^{13,14} recurrent neural networks (RNNs) such as long short-term memory (LSTM),¹⁵ and bi-directional long short-term memory (Bi-LSTM)¹⁶ are used for heart categorization and found significant improvement. In recent years, end-to-end training of CNN has been the dominating technique for health care image analysis.¹⁷ Additionally, because of its capacity to record position and shift-invariant modes, CNN is used to analyze the morphology of clinical information. Even when the input signal is noisy, CNN may also be able to retrieve valuable data.¹⁸ These performance characteristics are mirrored in the network structure built layer by layer. As the network's layers increase, features are learned and expressed more abstractly and concisely.

Moreover, LSTM is a kind of artificial RNN,^{19,20} which is suitable for classifying sequences and time-series data. LSTM only preserves the previous data because the only inputs it has received are from the past. The Bi-LSTM is a variant of the traditional LSTM capable of learning from both past and future states. It enables the network to learn representations of the characteristics and the temporal connection between the features.

This work evaluates a deep learning model that combines CNN and Bi-LSTM on two data sets (MIT-BIH and St-Petersburg data sets) for autonomously detecting arrhythmia illness from ECG signals. The model is capable of detecting and classifying five different types of cardiac arrhythmias such as non-ectopic (N), supraventricular ectopic (S), ventricular ectopic (V), fusion (F), and

unknown (Q) beats. The contributions of this article are as follows:

- An overview of the state-of-the-art work related to categorizing multiple classes from ECG signals using different data sets.
- Proposed a novel deep learning model for categorizing five classes of cardiac arrhythmia from ECG signals using the MIT-BIH and ST-Petersburg data sets.
- Evaluate the optimum hyper-parameters of conv1D in terms of Kernel size, number of filters, activation function and number of layers.
- We combined the Bi-LSTM technique of size 32 based on factorial cross-entropy following the Adam optimizer with an evaluated optimum conv1D.

This approach is the first to detect cardiac arrhythmia in the way described above to the best of our knowledge. The results indicate that the detection accuracy, sensitivity, specificity, and precision are promising.

The rest of the article's organization is as follows. The "Background" section covers the basic concepts of CNN, LSTM, and Bi-LSTM algorithms and their key parameters required to understand the topic. The "Methodology" section covers the methodology, explaining the workflow of the performance analysis process. The "Results" section gives the results, such as model performance, training and validation outcomes, and the "Discussions" section discusses the proposed model. Finally, the "Conclusions" section offers the conclusion.

Related work

There has been significant research on using deep learning models to classify arrhythmias. Study by Acharya et al.²¹ describes using a CNN method to identify the various ECG signals automatically. The method comprises an 11 layers of deep CNN with a four-neuron output layers that indicates typical atrial fibrillation (AFIB), Nsr, atrial flutter (AFL), and Vfib ECG classifications. The deep learning one-dimensional (1D)-CNN model was proposed in the work of Yıldırım²² which was capable of correctly identifying cardiac arrhythmias (17 diagnostic classifications including normal sinus rhythm, the rhythm of a pacemaker, and 15 other rhythm abnormalities) from the assessment of ECG signal.

Most researchers combined CNN with LSTM and achieved satisfactory results in classifying arrhythmia from ECG signals. Work of Oh et al.²³ proposed an automated system utilized a combination of CNN-LSTM algorithms for diagnosing normal beats, left bundle branch block, right bundle branch block, atrial premature beats and premature ventricular contraction from ECG signals. The deep CAE-LSTM method was proposed in²⁴ for recognizing arrhythmic heartbeat. Specifically, the CAE

technique was utilized to shorten ECG waves and gather low-dimensional digitized information from individual ECG records. The coded signals were then classified using an LSTM network model. They operated 100,022 fragmented data representing five distinct ECG pulse patterns. Work of Zheng et al.²⁵ developed a technique for categorizing arrhythmias by combining a CNN with LSTM, which was then applied to eight ECG data sets, one of which had a typical normal heart rhythm. Deep learning-based technique for automatically classifying six kinds of ECG signals: AFIB, normal sinus rhythm (N) segments, pacing rhythm (P), ventricular bigeminy (B), sinus bradycardia, and atrial flutter (AFL) were presented in.²⁶ This framework processed ECG recordings and related RR duration from the well-known data set MIT-BIH using a multi-input structure.

In the work of Yao et al.,²⁷ an attention-based time-incremental CNN model was developed to categorize arrhythmia using wide-ranging ECG data. This model collected information from ECG data in two stages: spatial information fusion using CNN, temporal information fusion using LSTM cells and an attention module that classified eight classes of arrhythmias and sinus rhythms. In the work of Rai and Chatterjee,²⁸ the authors suggest a hybrid CNN-LSTM model and ensemble technique for identifying myocardial infarction using ECG signals. They achieved satisfactory results by classifying all non-MI beats (N, S, V, F, Q) as regular beats and myocardial infarction beats as MI beats. An automatic diagnosis of cardiac arrhythmias from large-scale ECG data has been suggested²⁹ utilizing a feature extraction approach based on the multi-layer probabilistic neural network (MPNN) classifier to classify two distinct classes of arrhythmia.

Although there has been considerable work on detecting cardiac arrhythmia using deep learning techniques, an effective CNN-Bi-LSTM technique for arrhythmia classification is still lacking. Recently, Bi-LSTM has been used in different areas and has shown good performance.³⁰ In this work, we combine CNN with Bi-LSTM models to evaluate and classify five different ECG segments, comprising N, S, V, F, and Q.

Background

An ECG is a non-invasive diagnostic tool that visually captures the heart's pulse and electrical activity over time.³¹ Arrhythmia is described as the irregular electrical activity of any set of diseases affecting the heart, resulting in a fast or slow pulse.³² DNNs have been utilized in the categorization of ECGs recently. A DNN model's features are more extensive than those manually retrieved with adequate training data. The CNN and LSTM architectures are commonly utilized for prediction, detection, and recognition in most artificial intelligence applications using sequential data or signals of one dimension. As a result of

its local and parameter sharing capabilities, CNN is an efficient technique for extracting features. Also, LSTM is a frequently used technique for processing time-series signals. Both CNN and LSTM are more successful in detecting multi-class arrhythmias.³³ The following section provides an in-depth discussion on the CNN, LSTM, and Bi-LSTM techniques.

Convolutional neural network

CNN is a subset of artificial neural networks that are extensively utilized for image processing,³⁴ feature extraction and categorization of time series data. CNN has been one of the most often utilized artificial intelligence methods due to its superior capacity to identify characteristics automatically. CNN models are constructed using convolutional layers and a fully connected layer with related weights, and pooling layers.³⁵ Moreover, 1D-CNNs are highly suitable for real-time applications because of their minimal processing requirements.

The basic CNN architecture inside the network comprises a series of layers, each of which is responsible for a distinct function. Multiple convolutional filters work concurrently to extract outcomes and display them as activation. Several convolutions enhance activations, resulting in a feature map for the associated input.³⁶ The following equation applies to single convolutional processing of a signal $x_1^0 = [x_1, x_2, x_n]$ in which n denotes the total number of points, h is the activation function, layer index is l , b is the j th feature map's bias, W_m^j is the feature map's weight, kernel size is M , and m th is the filter index.¹ The local characteristics of the ECG segments are extracted using CNNs. The general conv1d network design is depicted in Figure 1.

$$C_i^{lj} = h(b_j + \sum_{m=1}^M W_m^j x_{i+m-1}^0) \quad (1)$$

Long short-term memory

LSTM technique is based on RNNs. A fundamental LSTM unit cell consists of an input gate, an output gate, and a forget gate. The LSTM enables the processing of a variety of different types of data in adjacent time steps, thereby producing an internal feedback state that helps the network to comprehend the concept of time and high variability of the exhibited data.³⁷ By incorporating gate control, the LSTM network integrates resolving gradient disappearing up to a specific extent.³⁸

While an LSTM framework can retain and comprehend previous data, it cannot correctly use new data to assist final prediction. As a result, a Bi-LSTM³⁹ with two-way capability was built using the LSTM network. Bi-LSTM networks outperform other RNN, and LSTM structures in sequence prediction modeling, particularly in the areas of machine translation and voice or handwriting detection.⁴⁰

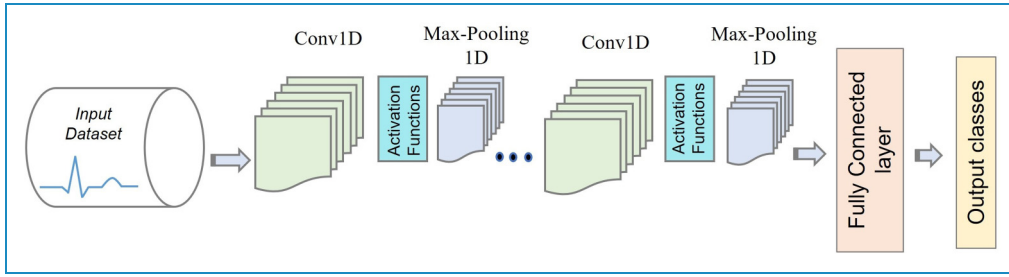


Figure 1. General architecture of Conv1D.

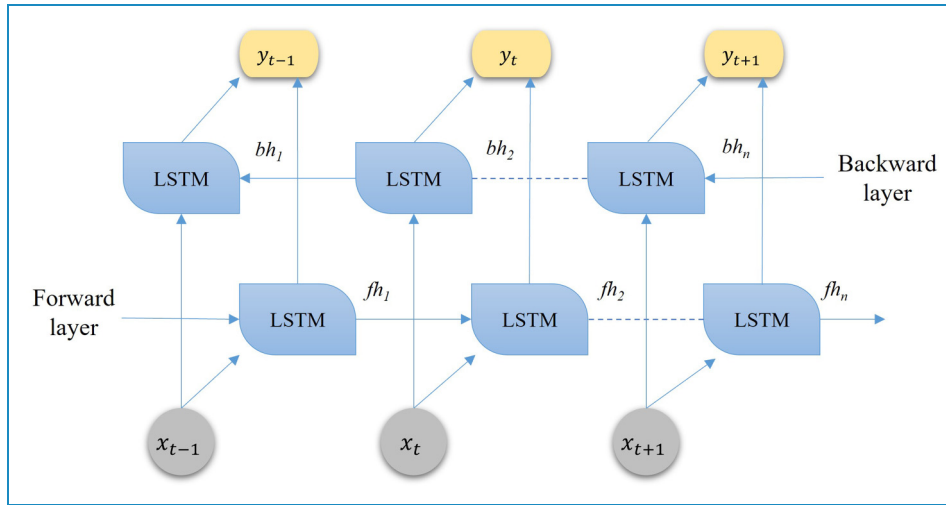


Figure 2. General architecture of bi-directional long short-term memory (Bi-LSTM) network.

Bi-directional long-short term memory

The Bi-LSTM comprises two separate LSTMs capable of summarizing data from both the backward and forward directions and then integrating the data from both ways. It is advantageous to access both previous and future contexts while doing sequence labeling tasks. Bi-LSTM proposes forwarding and reversing each sequence to two distinct hidden states to gather future and past information and then concatenating the two hidden states to produce the outcome.⁴¹ For every time t , the next LSTM generates the hidden vector fh_t using the prior hidden vector $fh_{(t-1)}$ and the input words encoding x_t , while the reverse LSTM produces the hidden vector bh_t using the opposite prior hidden vector $bh_{(t-1)}$. After that, the backward hidden vector bh and forward hidden vector fh are combined to form the Bi-LSTM model's last hidden vector. In the Bi-LSTM framework, the parameters of two opposing directions are distinct, although they share identical sentence embeddings. Figure 2 depicts the Bi-LSTM model's fundamental structure, where fh_1, fh_2, \dots, fh_n represents the forward hidden vector and bh_1, bh_2, \dots, bh_n representing the backward hidden

vector, respectively. h_n denotes the vector formed by the intersection of fh_n and bh_n .

MIT-BIH arrhythmia database

The MIT-BIH data set has grown in popularity in recent years. This data set has been utilized at over 500 locations globally since 1980 to conduct basic heart dynamics research.⁴² The proposed work incorporates experimental data from the MIT-BIH arrhythmia database, which is publicly available on PhysioNet.⁴³ The signals were digitized at a sampling rate of 360 samples per second per channel and a resolution of 11 bits throughout a 10 mV range. In total, 48 half-hour snippets of 24-hour two-channel ECG recordings from 48 recordings were used for this work at a sample rate of 360 Hz.⁴⁴

ECG signals in the MIT-BIH arrhythmia database were described by a binary file (.dat), a text header file (.hea), and a binary annotation file (.atr). The header file (.hea) is a small text file that contains information about the signals' details (includes the file's name, sample amount, recording's format, and comprehensive clinical data). Annotation files

have a collection of labels, each describing a property of one or more signals at a specific location in the record.⁴⁵

St-Petersburg arrhythmia database

St-Petersburg INCART 12 lead arrhythmia database includes 75 recordings that have been annotated. These samples were culled from 32 different Holter recordings. Each channel lasts thirty minutes and contains 12 standard leads, with a sampling rate of 257 Hz. The samples were collected from 17 men and 15 women, ranging from 18 to 80 years old, and 58 years old for men.⁴⁶ The .hea file contains information on the patient's age, gender, diagnosis, and an analysis of the ECG's characteristics for each record. Additionally, each .hea file has a patient's number (1–32) that uniquely distinguishes the source recordings; each patient's recording was generated from the identical Holter records.

Methodology

The flowchart for conducting the research is shown in Figure 3. Firstly, MIT-BIH and St-Petersburg data sets are prepared for model training by performing signal extraction, data normalization, and data balancing. The ECG signals have been split throughout the training, validation, and testing phases. We extract the ECG signals using the *WFDB* package and then normalize the data using *tf.Keras.utils.normalize*. The resampling method was used to balance the training data, with down-sampling for the majority class and up-sampling for the minority classes. The model development stage starts with building a CNN model with optimal number of layers, filters, kernel sizes, and activation function. The CNN model is then integrated with a Bi-LSTM model. The number of layers in Bi-LSTM is determined when the accuracy of categorizing output classes of arrhythmia has been satisfied at acceptable level.

The training, validation and testing phases of the model is performed separately for MIT-BIH and St-Petersburg data sets. Initially it is performed with MIT-BIH, then it is performed with St-Petersburg data set. The results are then compared as another measure to obtain a credible model.

Data preparation

The proposed model is evaluated utilizing MIT-BIH and St-Petersburg database. There are 48 individuals in the MIT-BIH database, with each subject's data including normal, abnormal, and non-beats. Moreover, the number of seconds of recording before and after the beat is set to three seconds. Finally, the data set is divided into subsets for training, validation, and testing. 60% patients' data were chosen randomly for training, and 20% for validation, while the remaining 20% patients' data were utilized for testing. Similarly, the St-Petersburg data set has 32 recordings, each lasting thirty minutes and consisting of 12

standard leads with normal, abnormal, and non-beats. The St-Petersburg data set has also been randomly divided into training, validation and testing phases of a time duration of 3 sec.

Data balancing

A balanced data set is one that has a similar amount of input samples for each output class or target class. The use of a resampling approach can accomplish balancing. It is critical to balance the data set, or at the very least, bring it near to being balanced. The primary purpose of this is to ensure that each class is given equal priority. The MIT-BIH and St-Petersburg arrhythmia data sets are imbalanced data sets. We used the down-sampling and up-sampling techniques in order to balance the data sets. After resampling the training data we fixed 30,000 samples for each output class. After data balancing, we implement the data into the proposed model to categorize cardiac arrhythmia.

Proposed model (CNN-Bi-LSTM)

The proposed deep learning model combined two neural networks, CNN and Bi-LSTM. The choice to integrate these two techniques was made with the understanding that their combination provides better outcomes.²⁵ The network created in this study has two Conv1D layers and two Bi-LSTM layers. An activation function follows each Conv1D layer for rectified linear units (ReLU). To boost their effectiveness in feature extraction, we then add two 1D max-pooling layers of size 2. The Conv1D layer was created with a kernel size of 5 and 32 filters in total, and then the Bi-LSTM layers having a size of 32 were used to link these layers. A fully connected layer with softmax activation is utilized after enhancing the temporal information between segments using Bi-LSTM. We also used a callback class to stop the model when reaching 0.999% accuracy.

We compute the output classification probabilities from the training and validation outcomes using performance matrices. We also apply `kernel_regularizer = regularizers.l2` of size 0.0003 to reduce information redundancy and prevent overfitting because most segments overlap. Furthermore, Table 1 shows the CNN-Bi-LSTM model summary, which includes information such as output shape, layer type, and the number of parameters. Moreover, we utilized `categorical_crossentropy` for the loss function, and the optimizer considered was "Adam" optimizer. Figure 4 explained the proposed model layer by layer detail as well as the optimum parameters which we achieved from experiments.

Performance matrices

The performance of the developed model was calculated in terms of accuracy, recall, precision, and specificity. These performance metrics are described next:

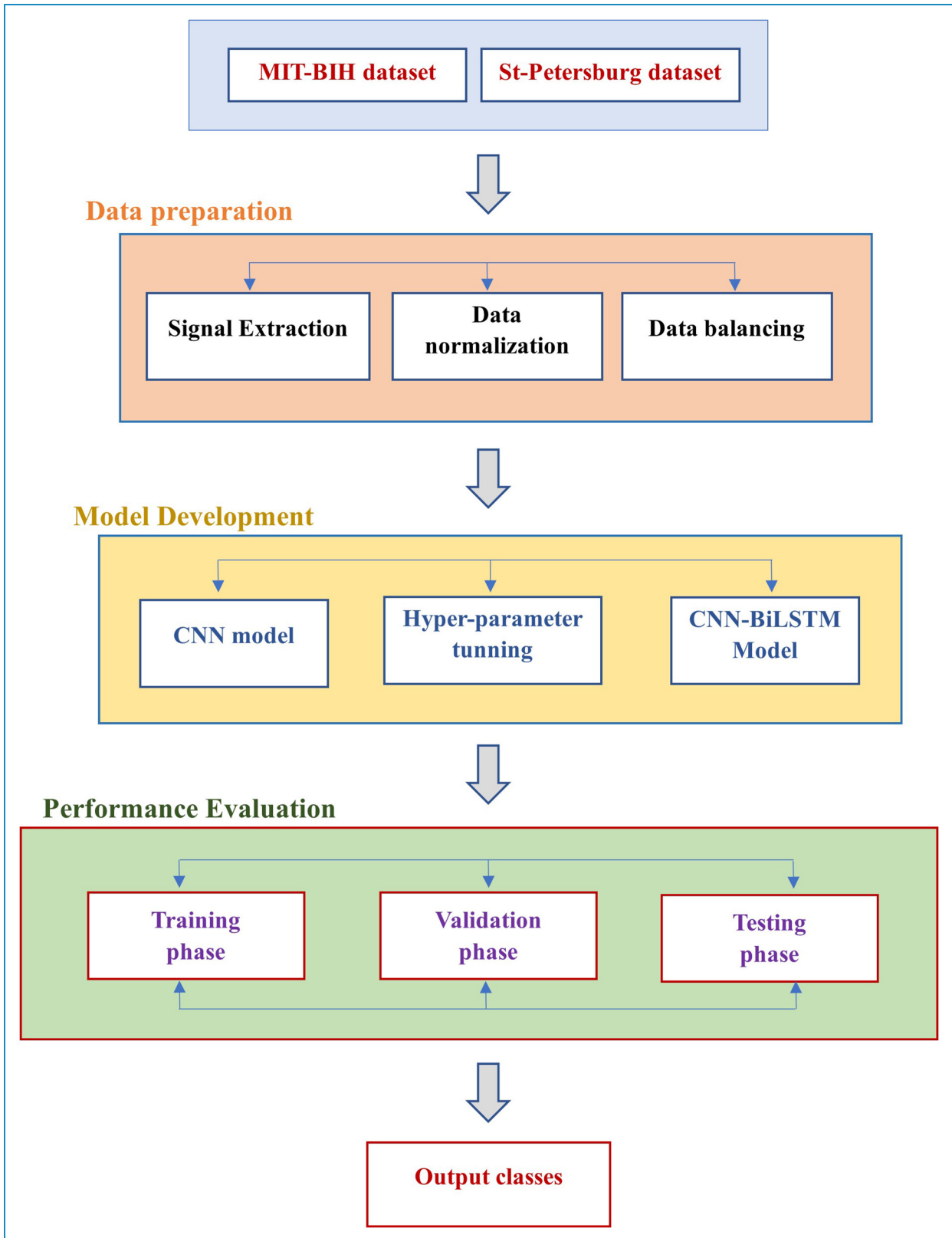


Figure 3. Flowchart diagram of the proposed work.

Table 1. Model summary of the convolutional neural network and bi-directional long short-term memory (CNN-Bi-LSTM) framework for MIT-BIH Arrhythmia data set highlighting the layer types, output shape and number of parameters.

Layer (type)	Output shape	Parameters
Conv1D_1	182, 32	192
Max-pooling1D	91, 32 32	0
Conv1D_2	87, 32	5152
Max-pooling1D	43, 32	0
Bidirectional_1	43, 64	16640
Bidirectional_2	43, 64	24832
Dense_1	16	1040
Dense_2	5	85
Total number of parameters:	47,941	
Trainable parameters:	47,941	
Non-trainable parameters:	0	

- **Accuracy:** is described as “the amount to which the outcome of a measurement correlates to the proper value or a norm”, and it relates to the proximity of a measure to its accepted values.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (2)$$

- **Recall:** also known as sensitivity, is the ability of a test to discover patients with a condition correctly. Sensitivity measures how accurate a test produces a positive result (also known as the “true positive” rate) for people who have tested the ailment.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

- **Precision:** is described as “performing precisely” and relates to the proximity of two or more measures, whether or not they are accurate. Precision measurements have a risk of being erroneous.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

- **Specificity:** refers to a test’s capacity to reliably produce a negative result for those who do not have the disease assessed (also known as the “true negative” rate).

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (5)$$

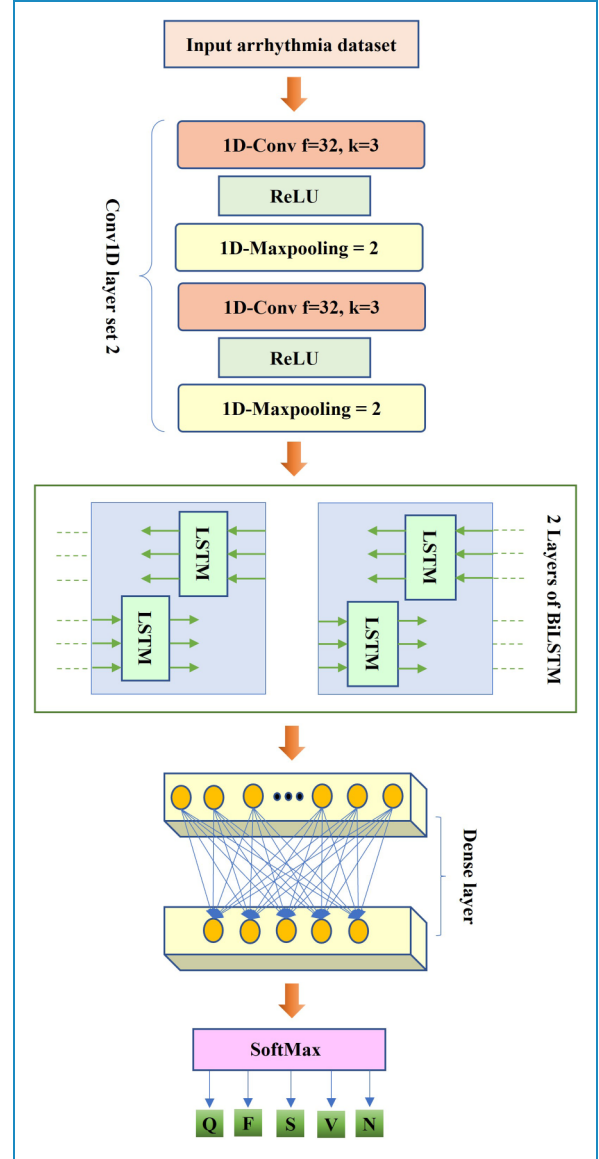


Figure 4. Proposed CNN-Bi-LSTM architecture.

Results

The model was trained and tested to recognize arrhythmia disease using CNN and Bi-LSTM algorithms. The Jupyter Notebook application is used to build and evaluate our deep learning model.⁴⁷ The *pypi* package *wfdb* is used to import the ECG and annotation and then show all the annotation in order to visualize the distribution of the cardiac rhythm. Table 2 describes the achieved training, validation and testing overall performance parameters of the proposed technique on both data sets. The developed model comprises two layers of conv1D along with 1D max-pooling and two layers of Bi-LSTM. The number of filters is set at 32, and the kernel size is fixed at 5, which achieved higher training, validation and testing accuracies of

Table 2. Overall performance parameters of training, validation and testing values of the proposed convolutional neural network and bi-directional long short-term memory (CNN-Bi-LSTM) model using both data sets.

Performance parameters	MIT-BIH data set			St-Petersburg data set		
	Train values	Valid values	Test values	Train values	Valid values	Test values
Accuracy	100%	98.0%	98.0%	98.0%	95.0%	95.0%
Recall	100%	92.20%	91.0%	98.2%	92.4%	92.2%
Precision	100%	88.0%	88.20%	98.2%	75.0%	73.8%
Specificity	99.2%	92.02%	90.96%	97.6%	92.5%	92.3%
F1-score	100%	89.80%	89.80%	98.2%	81.0%	80.4%

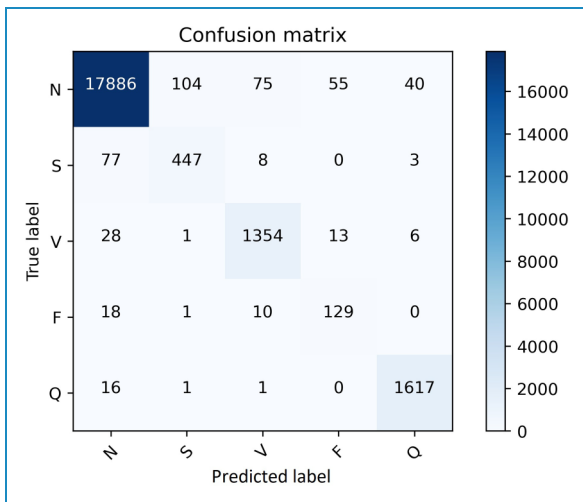


Figure 5. Validation-based confusion matrix for arrhythmia categorization using the MIT-BIH data set.

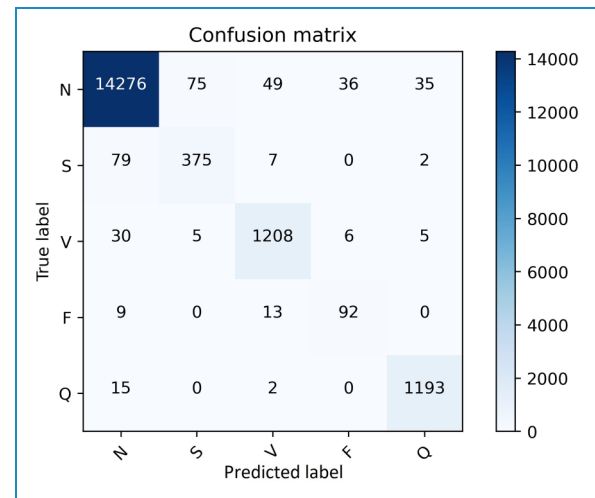


Figure 6. Testing-based confusion matrix for arrhythmia categorization using the MIT-BIH data set.

MIT-BIH and St-Petersburg data sets. A stratified random test was conducted by splitting the MIT-BIH and St-Petersburg data set into different parts. In total, 60% of the data was utilized for model training 20% for validation, and the remaining 20% was used to test the model. The procedure was repeated until all of the data had been examined.

Model performance

Prior to finding the optimal hyperparameter, we train our model with a variety of parameters, including the number of filters from 32 to 128, kernel size from 3 to 11, batch normalization, exponential Linear Unit (ELU), and a dropout rate of 0.25, until we find the optimal performance parameters, which are as follows: Adam optimizer, conv1d filter size 32, kernel size 5, kernel_regularizer L2, ReLU activation function, max-pooling size 2, Bi-LSTM 32, epochs 100, dense 16, and finally dense 5 with softmax function.

The confusion matrix for each class of validation and testing of the MIT-BIH data set is shown in Figures 5 and 6. While the confusion matrix of the St. Petersburg data set is shown in Figures 7 and 8 which indicates the number of correctly predicted samples against incorrectly predicted samples for each class.

Validation and training outcomes

The developed network performed well in investigations of rhythm epoch classification. After 100 epochs of training, the categorization accuracy values for the MIT-BIH data set were almost 100%, with a relatively flat curve. Whereas the valid set had an accuracy of between 96% and 98%, and the curve was stable (Figure 9). Consequently, the training set's loss curve was almost zero and very flat. In comparison, valid loss varied between 0 and 0.2 across epochs (Figure 10), suggesting

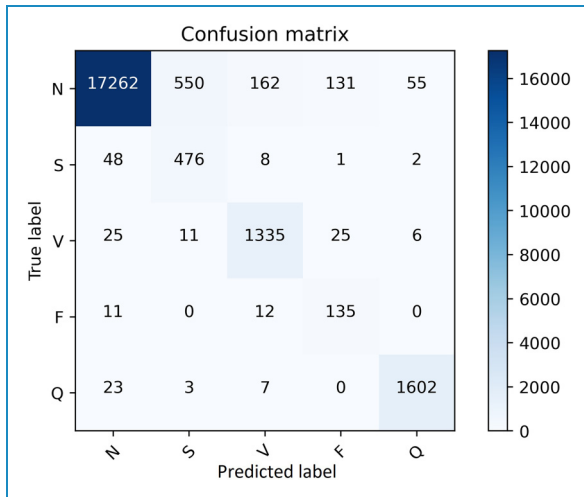


Figure 7. Validation-based confusion matrix for arrhythmia categorization using the St-Petersburg data set.

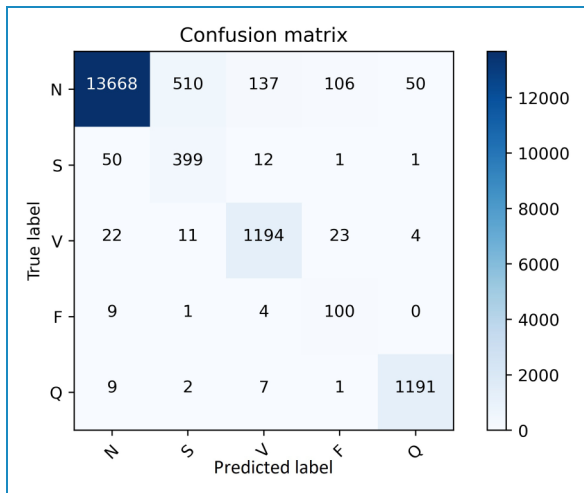


Figure 8. Testing-based confusion matrix for arrhythmia categorization using the St-Petersburg.

that the cross-entropy loss function has a high level of generalization. The five standard statistical measures were utilized to determine the effectiveness of each categorization algorithm: recall, accuracy, precision, specificity, and f1-score. The overall performance demonstrates that the MIT-BIH data set is always superior to the St-Petersburg data set in terms of five performance matrices. Figure 11 depicts our network’s receiver operating characteristic (ROC) curves on the validation data set. The dashed diagonal line indicates a random performance with a 0.5 ROC. Our network demonstrated a high ROC of 1.

Discussions

Doctors are frequently required to evaluate and diagnose cardiac problems using single-lead or multiple-lead ECG

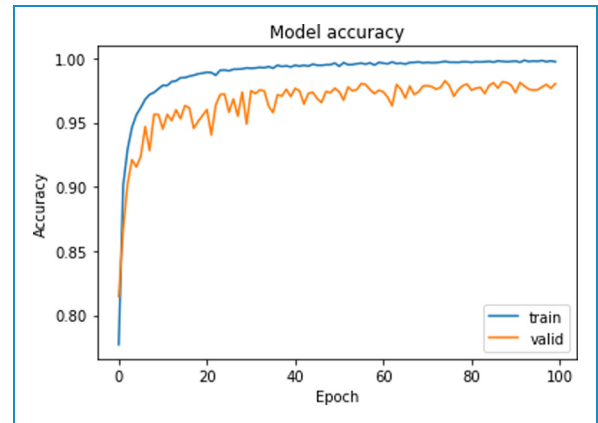


Figure 9. Model accuracy curve.

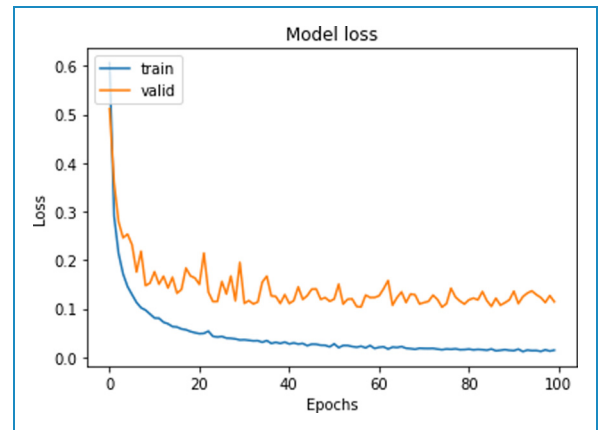


Figure 10. Model loss curve.

readings in clinical applications. However, an efficient and low-complexity automatic CVD identification model is required due to the heavy burden associated with a clinical diagnosis and the disparity in doctors’ expertise levels. A novel technique integrating CNN with Bi-LSTM explores how well this combination performs in detecting and classifying heartbeat occurrences.

There exist considerable work in the literature to detect ECG signals in the MIT-BIH arrhythmia database. Table 3 summarises some of these investigations. Rajendra et al.²¹ have achieved 92.50% accuracy performance for identifying four kinds of arrhythmia classes by applying 11 layers of CNN. Acharya et al.⁴⁸ designed a nine-layer CNN model capable of automatically classifying five distinct types of heartbeats in ECG readings achieving the accuracy of 94.03%. Shu Lih Oh et al.²³ reported 98.10% accuracy for the categorization of five classes by using the CNN-LSTM model. Li Guo et al.⁴⁹ used DenseNet and gated recurrent unit networks GRU to overcome the challenge of inter-patient ECG categorization. In this work they categorize five ECG classes from MIT-BIH

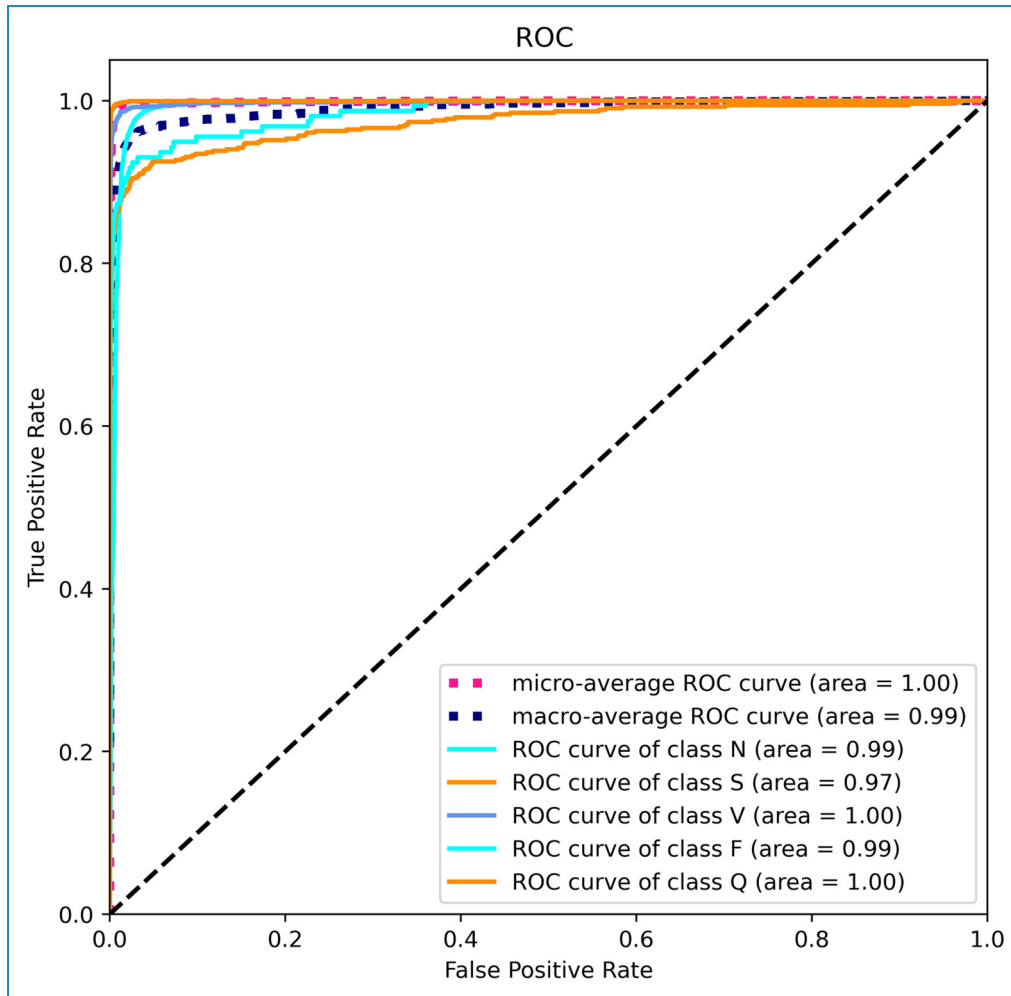


Figure 11. Receiver operating characteristic (ROC) curve of the proposed model.

data set. Xue Xu et al.³⁰ suggested a model for arrhythmia classification based on deep learning. Their investigation used a one-dimensional CNN and Bi-LSTM model to classify ECG data into five categories and found that this strategy achieved an accuracy of 95.90%. Our proposed CNN-Bi-LSTM technique outperforms the existing deep learning techniques for arrhythmia classification by achieving superior accuracy 98.0% and sensitivity 91.0% performance.

Conclusion

Arrhythmia is a severe CVD that can be predicted via ECG segment processing. Arrhythmia must be accurately diagnosed and prevented early to reduce cardiac disease. Our proposed system model met the study's primary goal of assisting doctors in swiftly determining the kind of ECG or verifying their diagnostics in a medical context while maintaining a high level of precision and cost. In this work, a CNN-Bi-LSTM model is proposed to categorize

five categories of ECG fragments to construct an effective and resilient autonomous computer-aided diagnosis system. The developed network achieved maximum accuracies of 100%, 98.0%, and 98.0% of training, validation, and testing using MIT-BIH data set. In comparison, the St-Petersburg data set achieved 98.0%, 95.0%, and 95.0% accuracies of training, validation, and testing in identifying arrhythmia.

This research showed many advantages, including its ability to help clinicians reliably make ECG recording-related clinical decisions. Moreover, it was intended to be as simple as possible while delivering the most significant performance. The described method is straightforward for health professionals and does not involve signal modification or feature extraction. Additionally, this research focused only on one kind of CVD, namely, arrhythmia, whereas the manifestations of cardiac disease are often complex and varied. As a result, more types of ECG data will need to be added to broaden the scope of the planned network.

Table 3. Summary of the proposed network comparing with state-of-the-art methodologies.

Work	Architecture	Preprocessing	Classes	Accuracy (%)	Specificity (%)	Recall (%)
Rajendra et al. ²¹	11 layers CNNs	Yes	4	92.50	93.10	98.06
Acharya, et al. ⁴⁸	9 layers CNN	Yes	5	94.03	91.54	94.03
Shu Lih et al. ²³	CNN-LSTM	Yes	5	98.10	98.70	97.50
Li Guo et al. ⁴⁹	CNN	No	5	93.71	94.77	91.25
Chen Chen ²⁶	CNN + LSTM	Yes	6	96.62	96.80	95.40
Alqudah et al. ¹¹	CNN	Yes	5	97.80	99.40	97.80
W. Jung et al. ⁵⁰	WKNN	Yes	4	96.12	99.97	96.12
Shraddha et al. ⁵¹	RNN-LSTM	No	2	88.10	92.40	83.35
Vandana et al. ⁵²	Ensemble SVMs	Yes	4	94.05	92.96	92.84
Our work	CNN-Bi-LSTM	No	5	98.00	90.96	91.00

SVM: support vector machine; CNN-Bi-LSTM: convolutional neural network and bi-directional long short-term memory; CNN: convolutional neural network; RNN: recurrent neural network.

Contributorship: SH and MS contributed to conceptualization. SH and KH conducted literature review and formulated the study. SH wrote the manuscript's first draft. TA contributed to data preparation and code implementation. The manuscript was reviewed and edited by all authors, and the final version was approved.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding: The research in this article is funded by: The Yayasan Universiti Teknologi PETRONAS Fundamental Research Grant (YUTP-FRG), [Grant 015LC0-244].

Guarantor: MS

ORCID iDs: Shahab Ul Hassan  <https://orcid.org/0000-0003-0995-4866>

Mohd S Mohd Zahid  <https://orcid.org/0000-0001-5042-5793>

References

- Alkhodari M and Fraiwan L. Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings. *Comput Methods Programs Biomed* 2021; 200: 105940.
- Direito A, Rawstorn J, Mair J, et al. Multifactorial e-and mhealth interventions for cardiovascular disease primary prevention: protocol for a systematic review and meta-analysis of randomised controlled trials. *Digital Health* 2019; 5: 2055207619890480.
- Tse G. Mechanisms of cardiac arrhythmias. *J Arrhythmia* 2016; 32(2): 75–81.
- He R, Liu Y, Wang K, et al. Automatic cardiac arrhythmia classification using combination of deep residual network and bidirectional LSTM. *IEEE Access* 2019; 7: 102119.
- Spencer R, Thabtah F, Abdelhamid N, et al. Exploring feature selection and classification methods for predicting heart disease. *Digital Health* 2020; 6: 2055207620914777.
- Ediriweera DS, Kasturiratne A, Pathmeswaran A, et al. Mapping the risk of snakebite in sri lanka-a national survey with geospatial analysis. *PLoS Negl Trop Dis* 2016; 10: e0004813.
- Zheng Y, Guo X and Ding X. A novel hybrid energy fraction and entropy-based approach for systolic heart murmurs identification. *Expert Syst Appl* 2015; 42: 2710–2721.
- Varghees VN, Ramachandran K and Soman K. Wavelet-based fundamental heart sound recognition method using morphological and interval features. *Healthc Technol Lett* 2018; 5: 81–87.
- Alqudah AM, Albadameh A, Abu-Qasmieh I, et al. Developing of robust and high accurate ecg beat classification by combining gaussian mixtures and wavelets features. *Australas Phys Eng Sci Med* 2019; 42: 149–157.
- Al-Hiyali MI, Yahya N, Faye I, et al. Identification of autism subtypes based on wavelet coherence of bold fmri signals using convolutional neural network. *Sensors* 2021; 21: 5256.
- Alquran H, Alqudah AM, Abu-Qasmieh I, et al. Ecg classification using higher order spectral estimation and deep learning techniques. *Neural Netw World* 2019; 29: 207–219.
- Al-Hiyali MI, Yahya N, Faye I, et al. Autism spectrum disorder detection based on wavelet transform of bold fmri

- signals using pre-trained convolution neural network. *Int J Integr Eng* 2021; 13: 49–56.
13. Xiao B, Xu Y, Bi X, et al. Heart sounds classification using a novel 1-d convolutional neural network with extremely low parameter consumption. *Neurocomputing* 2020; 392: 153–159.
 14. Al-Hiyali MI, Yahya N, Faye I, et al. Classification of bold fmri signals using wavelet transform and transfer learning for detection of autism spectrum disorder. In *2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*. IEEE, pp. 94–98.
 15. Raza A, Mehmood A, Ullah S, et al. Heartbeat sound signal classification using deep learning. *Sensors* 2019; 19: 4819.
 16. Husain K, Mohd Zahid MS, Ul Hassan S, et al. Advances of ecg sensors from hardware, software and format interoperability perspectives. *Electronics* 2021; 10: 105.
 17. Zhang F, Cai N, Wu J, et al. Image denoising method based on a deep convolution neural network. *IET Image Process* 2018; 12: 485–493.
 18. Qian Y, Bi M, Tan T, et al. Very deep convolutional neural networks for noise robust speech recognition. *IEEE/ACM Trans Audio Speech Lang Process* 2016; 24: 2263–2276.
 19. Hochreiter S and Schmidhuber J. Long short-term memory. *Neural Comput* 1997; 9: 1735–1780.
 20. Hassan SU, Zahid MSM and Husain K. Performance comparison of cnn and LSTM algorithms for arrhythmia classification. In *2020 International Conference on Computational Intelligence (ICCI)*. IEEE, pp. 223–228.
 21. Acharya UR, Fujita H, Lih OS, et al. Automated detection of arrhythmias using different intervals of tachycardia ecg segments with convolutional neural network. *Inf Sci (Ny)* 2017; 405: 81–90.
 22. Ö Yildirim, P Pławiak, al Tan RS, et al. Arrhythmia detection using deep convolutional neural network with long duration ecg signals. *Comput Biol Med* 2018; 102: 411–420.
 23. Oh SL, Ng EY, San Tan R, et al. Automated diagnosis of arrhythmia using combination of cnn and LSTM techniques with variable length heart beats. *Comput Biol Med* 2018; 102: 278–287.
 24. Yildirim O, Baloglu UB, Tan RS, et al. A new approach for arrhythmia classification using deep coded features and LSTM networks. *Comput Methods Programs Biomed* 2019; 176: 121–133.
 25. Zheng Z, Chen Z, Hu F, et al. An automatic diagnosis of arrhythmias using a combination of cnn and LSTM technology. *Electronics* 2020; 9: 121.
 26. Chen C, Hua Z, Zhang R, et al. Automated arrhythmia classification based on a combination network of cnn and LSTM. *Biomed Signal Process Control* 2020; 57: 101819.
 27. Yao Q, Wang R, Fan X, et al. Multi-class arrhythmia detection from 12-lead varied-length ecg using attention-based time-incremental convolutional neural network. *Inf Fusion* 2020; 53: 174–182.
 28. Rai HM and Chatterjee K. Hybrid CNN-LSTM deep learning model and ensemble technique for automatic detection of myocardial infarction using big ecg data. *Appl Intell* 2021; 52(5): 1–19.
 29. Rai HM and Chatterjee K. A novel adaptive feature extraction for detection of cardiac arrhythmias using hybrid technique mrdwt & MPNN classifier from ecg big data. *Big Data Res* 2018; 12: 13–22.
 30. Xu X, Jeong S and Li J. Interpretation of electrocardiogram (ecg) rhythm by combined cnn and BILSTM. *IEEE Access* 2020; 8: 125380.
 31. Sannino G and De Pietro G. A deep learning approach for ecg-based heartbeat classification for arrhythmia detection. *Future Gener Comput Syst* 2018; 86: 446–455.
 32. Rai H, Trivedi A, Chatterjee K, et al. R-peak detection using daubechies wavelet and ecg signal classification using radial basis function neural network. *J Inst Eng (India): Series B* 2014; 95: 63–71.
 33. Ö Yildirim. A novel wavelet sequence based on deep bidirectional LSTM network model for ecg signal classification. *Comput Biol Med* 2018; 96: 189–202.
 34. Radzi SA and Khalil-Hani M. Character recognition of license plate number using convolutional neural network. In *International Visual Informatics Conference*. Springer, pp. 45–55.
 35. Alqudah AM, Qazan S, Al-Ebbini L, et al. Ecg heartbeat arrhythmias classification: a comparison study between different types of spectrum representation and convolutional neural networks architectures. *J Ambient Intell Humaniz Comput* 2021; : –.
 36. Schmidhuber J. Deep learning in neural networks: An overview. *Neural Netw* 2015; 61: 85–117.
 37. Rai HM, Chatterjee K and Mukherjee C. Hybrid CNN-LSTM model for automatic prediction of cardiac arrhythmias from ecg big data. In *2020 IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*. IEEE, pp. 1–6.
 38. Reddy BK and Delen D. Predicting hospital readmission for lupus patients: An RNN-LSTM-based deep-learning methodology. *Comput Biol Med* 2018; 101: 199–209.
 39. Liu G and Guo J. Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing* 2019; 337: 325–338.
 40. Rao G, Huang W, Feng Z, et al. LSTM with sentence representations for document-level sentiment classification. *Neurocomputing* 2018; 308: 49–57.
 41. Na SH, Kim H, Min J, et al. Improving LSTM CRFS using character-based compositions for korean named entity recognition. *Comput Speech Lang* 2019; 54: 106–121.
 42. Moody GB and Mark RG. The impact of the mit-bih arrhythmia database. *IEEE Eng Med Biol Mag* 2001; 20: 45–50.
 43. Moody GB, Mark RG and Goldberger AL. Physionet: a web-based resource for the study of physiologic signals. *IEEE Eng Med Biol Mag* 2001; 20(3): 70–75.
 44. Rai HM and Chatterjee K. A unique feature extraction using mrdwt for automatic classification of abnormal heartbeat from ecg big data with multilayered probabilistic neural network classifier. *Appl Soft Comput* 2018; 72: 596–608.
 45. Rai HM, Trivedi A and Shukla S. Ecg signal processing for abnormalities detection using multi-resolution wavelet transform and artificial neural network classifier. *Measurement* 2013; 46: 3238–3246.
 46. Tihonenko V, Khaustov A, Ivanov S, et al. St.-petersburg institute of cardiological technics 12-lead arrhythmia database, 2007. doi:10.13026/C2V88N.
 47. Kluyver T, Ragan-Kelley B, Pérez F, et al. *Jupyter Notebooks-a publishing format for reproducible computational workflows.*, volume 2016. 2016.
 48. Acharya UR, Oh SL, Hagiwara Y, et al. A deep convolutional neural network model to classify heartbeats. *Comput Biol Med* 2017; 89: 389–396.

-
49. Guo L, Sim G and Matuszewski B. Inter-patient ecg classification with convolutional and recurrent neural networks. *Biocyber Biomed Eng* 2019; 39: 868–879.
 50. Jung WH and Lee SG. An arrhythmia classification method in utilizing the weighted knn and the fitness rule. *IRBM* 2017; 38: 138–148.
 51. Singh S, Pandey SK, Pawar U, et al. Classification of ecg arrhythmia using recurrent neural networks. *Procedia Comput Sci* 2018; 132: 1290–1297.
 52. Singh V, Reddy US and Bhargavia GM. A generic and robust system for automated detection of different classes of arrhythmia. *Procedia Comput Sci* 2020; 167: 1801–1810.
-