

# Deep Convolutional Bidirectional LSTM Model for identifying Ventricular Tachyarrhythmia using ECG Signal Variability

MANOJ KUMAR OJHA (✉ [manojjha15@gmail.com](mailto:manojjha15@gmail.com))

Madhav Institute of Technology and Science <https://orcid.org/0000-0003-0557-4743>

**Sulochna Wadhvani**

Madhav Institute of Technology and Science

**Arun Kumar Wadhvani**

Madhav Institute of Technology and Science

**Anupam Shukla**

Indian Institute of Information Technology Pune

---

## Research Article

**Keywords:** Tachyarrhythmia, HRV, CNN and BiLSTM

**Posted Date:** December 29th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1194607/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

An electrocardiogram (ECG) signal is used widely to detect ventricular tachyarrhythmia (VTA) and to diagnose heart disease. Deep learning models and large ECG data have made the diagnosis of VTA an attractive task to demonstrate the power of artificial intelligence in clinical applications. One of the life-threatening complications of VTA is cardiac arrest (CA). The VTA is divided into two categories: ventricular fibrillation (VF) and ventricular tachycardia (VT). Abnormal electrical activity in the ventricle causes VT, which leads to CA, whereas the chaotic electrical activity in the ventricle leads to VF. To improve the clinical diagnostic system and to help cardiologists, it is essential to identify the risk of VTA at an early stage. The goal of this paper is to develop an end-to-end (E2E) deep learning model that uses a convolution neural network (CNN) and a bidirectional long-short term memory network (BiLSTM) to classify VT and VF arrhythmias from multiple ECG databases. The CNN extracts features from ECG signals, and BiLSTM learns information. The ECG signals are acquired from the MIT-BIH malignant ventricular arrhythmia database (VFDB) and the Creighton University VTA database (CUDB). These ECG signals indicate that heart rate variability is a fast and dynamic event. Before the method's implementation, ECG signals are windowed at a fixed size according to annotation information and then normalized within each window. In terms of accuracy and sensitivity, the proposed CNN-BiLSTM deep learning model outperforms existing state-of-the-art methods. These results made it possible to obtain a relatively higher average accuracy (AC) of 99.37%, precision (PE) of 97.12%, a sensitivity (SE) of 98.15%, and F-score (FS) of 98.43%, and an overall accuracy of 99.07%, respectively.

## 1 Background

Many patients suffer from cardiac arrest in hospitals every year, an estimated 209,000 people in the United States alone, of whom only one in four survive (Kibos et al. 2013). Cardiac arrest occurs when the heart ceases pumping blood to the brain and other vital organs (Theuns et al. 2019). Cardiac arrests in hospitalised patients are due to VF and VT without pulsation, collectively called VTA, as opposed to systole or electrical activity without pulsation (Chen et al. 2020). VT and VF are fatal heart rhythms that need to be corrected for the patient to survive. In VT, the ventricles beat so fast that the heart is not able to effectively fill its chambers and pump blood throughout the body. In VF, the ventricles quiver rather than a pump, requiring abortive treatment. If either condition is maintained, the brain dies from lack of blood flow and oxygen deprivation. Because the cardiac disease has such a high mortality rate, detecting VTA at an early contact as a boon is a patient treatment (Benjamin et al. 2017). The ECG signal, which shows the electrical activity of the heart over time using electrodes, is the most popular tool for identifying tachyarrhythmia. ECG leads, which measure the electrical potential of the heart from multiple angles and positions, are used to detect abnormalities in the signal types (Kleiger et al. 1987). The characteristics of the ECG have become one of the most important diagnostic tools for predicting heart disease. It contains a wealth of information, not only about the heart structure but also about the function of the electrical conduction system (Pandey et al. 2021). The one ECG cycle consists of five different waves. The P wave corresponds to atrial depolarization, the QRS complex occurs as a result of ventricular depolarization, and

the T wave corresponds to the polarization of the ventricle. The RR intervals are measured as the time between consecutive R waves. Thus, the variability of the RR interval is very important for measuring heart function. The analysis of variations in this R-R interval is known as HRV analysis (Pandey et al 2021). Many methods exist when it comes to classifying VTAs. Automated algorithms are highly desirable for more accurate determination of VT and VF states for timely diagnosis of VTA. Deep learning and machine learning model have been combined to classify VT and VF. Atienza et al. (2013) proposed a strategy of extraction and selection of the most appropriate features from the ECG signal for the identification of VT and VF. They extracted 13 ECG features that were input to an SVM to identify VT and VF rhythms with a SE of 92%. Lee et al. (2013) extracted fourteen optimal features from ECG signals using a genetic algorithm. The SVM classifier was used to classify different episodes of ventricular arrhythmias based on extracted features with an AC of 96.3%. Panda et al. (2020) proposed a method for detecting VT and VF using an ECG database. A fixed frequency range empirical wavelet transform filter bank for multiscale analysis was introduced. The evaluated ECG signals were used as inputs to CNN to detect VT and VF. The proposed CNN includes four convolutions, two poolings, and four dense layers. ECG signals from various public databases were used to evaluate CNN. The results show that the proposed approach has achieved an AC of 81.25% for the classification of VT and VF classes using 8-second ECG frames with a 10-fold cross-verification. Alwan et al. (2017) classified VT and VF using the leakage characteristics of the VF filter and spectral parameters. A set of high-dimensional functions preserves more up-to-date information about the ECG dataset and provides an AC of detection of 73.5% using a support vector classifier. Plesinger et al. (2018) extracted five features from 3-second blocks of ECG signal using analysis of ECG amplitude measurement, spectra, derivatives, and autocorrelation. These features were fed into a logistic regression method viewing the probability of the VT and VF. This method was trained on CUDB, consisting of 393 blocks, and tested on VFDB, showing a SE of 95%. In this paper, the deep learning model is composed of CNN-BiLSTM for identifying the VT and VF rhythms from the ECG signal. The ECG databases have been obtained from two publicly available databases, namely CUDB and VFDB. CNN automatically learns the features of multiple five-second ECG windows through multilayer nonlinear transformation and it is better than conventional features extraction methods. Deep learning can quickly learn from training data and automatically obtain valid features, mining the information hidden behind ECG large datasets. Various hidden factors are often associated with complex non-linear approaches, but deep learning can separate these factors. The CNN structure of deep learning makes it highly expressive and learning. It is especially good at extracting features and making VTA classification simple and effective. Following the extraction of features from the CNN model, they are input into the BiLSTM, which subsequently maintains information on long-distance dependencies. LSTM method, which uses recurrent neural network techniques to learn from sequential data, is well-suited for sequential data. Because the data is structured in a sequential manner, it is possible to derive significant conclusions in both the forward and backward directions. As a result, we used BiLSTM in this study, which is suitable for processing ECG window data in both directions and classifying them into NSR, VT, and VF. According to our experimental results, as mentioned in the results section, BiLSTM can enhance the classification rate.

## 2 Methods

The workflow process for identifying VTA from normal sinus rhythms (NSR) using the E2E deep learning model based on CNN-BiLSTM is represented in Figure 1. In order to classify the NSR, VT, and VF classes, the CNN-BiLSTM model has been designed using self-learning CNN features and BiLSTM-generated classes, which eliminates the need for a manual feature extraction and selection process. The continuous ECG signal is subdivided into five-second windows, with each window being normalized using the Z score approach. The windows database has been divided into five different train-test ratios, which are as follows: 50%-50%, 60%-40%, 70%-30%, 80%-20%, and 90%-10%.

### 2.1 ECG Databases

All of the ECG databases applied in this study were acquired from the publicly available physionet bank, which includes the CUDB and VFDB databases, among other databases (Goldberger et al. 2000). In both databases, cardiologists have annotated each ECG record that was recorded by different patients over time. All records from the CUDB and VFDB were evaluated and classified into two different VTA classes: VT and VF from NSR.

Table1: ECG databases details:

| Data Set | Number of ECG records | Sampling Frequency in Hz | Duration of ECG Signal in minute |
|----------|-----------------------|--------------------------|----------------------------------|
| CUDB     | 35                    | 250                      | 8                                |
| VFDB     | 22                    | 250                      | 30                               |

The CUDB has 35 ECG records of eight minutes in length from different patients who noted different VTA rhythms. Each recording has a sampling frequency of 250 Hz, and all of them have been annotated by the cardiologist as VF and VT. Similarly, the VFDB contains 22 recordings of 30 minutes of two-channel signals at a frequency of 250Hz, each labeled with VT, VF, NSR, and other rhythms. The typical ECG sample signals of NSR, VT, and VF rhythms are shown in figure 2.

The ECG records of both databases have been framed into fixed-size window lengths. In this study, we used a window length of 5 seconds (1250 samples) for each record. Each recording has a varied portion of the normal ECG signal, as well as portions that contain VT and VF signals. Table 2 shows the distribution of the number of ECG windows in each class.

Table 2  
The number of windows with duration of 5 seconds from both CUDB and VFDB.

| Class | Samples of Five Second Window in database |      |       |
|-------|---|------|-------|
|       | CUDB                                      | VFDB | Total |
| VT    | -   | 1202 | 1202  |
| VF    | 791                                       | -    | 791   |
| NSR   | 2744                                      | 7618 | 10362 |
| Total | 3535                                      | 8820 | 12355 |

All ECG records have been windowed into non-overlapping events. As a result, the total number of events is listed in Table 2 with a five-second window. Each window has been labelled as VF, VT, or NSR using the database's annotation file. Each window has been normalized using the Z- score method. This method eliminates the offset effect and addresses the issue of amplitude scaling before being fed into the CNN-BiLSTM model for the training and testing phases. Z-scores are calculated from the normalized signal value using the following equation:

$$Signalvalues = \frac{signalvalues - meanofsignalvalues}{standarddeviationofsignalvalues} \quad (1)$$

## 1.2 CNN Model

CNN model is composed of artificial neurons that have been designed with weights, biases, and activation functions to perform various tasks. These artificial neurons act as a bridge between the input (I/P) layer and the output (O/P) layer (Pandey et al. 2019). CNN model is made up of five different types of layers namely I/P layer, convolutional layer, pooling layer, fully connected layer, and the O/P layer. The ECG window data is the content of I/P layer in the proposed CNN-BiLSTM model. This data from the I/P layer is then passed on to the next convolution layer, which is also known as the feature learning layer since the features of the ECG signal data are learned through convolutional operations on the dataset from the I/P layer. This model also includes the rectifier linear unit (ReLU) function, which is used to make all negative values null. In the usual sequence, the O/P of the convolutional layer will be used as I/P for the pooling layer. To decrease the spatial volume of the I/P data, it is necessary to employ a pooling layer. It comes after the convolutional layer. If we didn't use the pooling layer, it would be computationally expensive to run the model. As a result, max-pooling is the only approach to decrease the spatial amount of the I/P data.

## 2.3 Bidirectional-LSTM Network

BiLSTM is a convolutional recurrent neural network that has been specifically designed to overcome the distinct limitations of convolutional recurrent neural networks (Jain et al. 2020). These network's neural states are divided into two categories, namely forwarding and backward states, which results in two distinct forward recurrent neural networks and backward recurrent neural networks, as shown in Figure 1. Both networks must be connected to the same output layer in order to produce an output. It is possible to evaluate future and past situations of sequential and one-time entries using this structure without waiting for the results to come in. When compared to the standard LSTM network, two different types of LSTM have been constructed specifically for sequential inputs in the BiLSTM architecture.

## 1.3 Proposed CNN-BiLSTM Model

The CNN-BiLSTM model consists of twelve layers, each of which is described below in Table 3. It has three convolutional layers, three max-pooling layers, two BiLSTM layers, flatten, and three dense layers, as illustrated in Table 3.

Table 3  
Detailed description of the proposed CNN-BiLSTM model for this task

| Layer            | Output Neuron | Kernel Size    | Activation Function |
|------------------|---------------|----------------|---------------------|
| Convolution -I   | 1216 X 4      | 35             | ReLU                |
| MaxPooling -I    | 608 X 4       | -              | -                   |
| Convolution -II  | 574 X 6       | 25             | ReLU                |
| MaxPooling -II   | 287 X 6       | -              | -                   |
| Convolution -III | 278 X 3       | 10             | ReLU                |
| MaxPooling -III  | 139 X 3       | -              | -                   |
| BiLSTM           | 139 X 8       | dropout = 0.25 | ReLU                |
| BiLSTM           | 139 X 16      | dropout = 0.25 | ReLU                |
| Flattern         | 2224          | -              | -                   |
| Dense -I         | 300           | -              | ReLU                |
| Dense-II         | 50            | -              | ReLU                |
| Dense-III        | 3             | -              | SoftMax             |

CNNs have the feature that they automatically emphasize and learn from the best-fit features in the input data by using the convolution operation. It is more concerned with local features and their placement, among other things. However, when using a CNN, the weights of neurons for the same feature map remain constant, resulting in parallel network learning, which reduces the amount of time spent learning.

According to Equation 2, the first, third, and fifth convolution layers in this model are convolutional with kernel sizes of 35, 25, and 10 in the first, third, and fifth layers, respectively.

$$Y_n = \sum_{i=0}^{N-1} x_i f_{n-i} \quad (2)$$

Here, Y - indicates the output vector, whereas x- represents the values of the signal, f - denotes the filters, and N- represents the number of elements.

Following each convolution layer, a second max-polling layer is applied in order to reduce the size of the feature map. The activation function of ReLU is used. In order to learn about the temporal information, the BiLSTM layer is used to learn about the feature maps. In terms of temporal analysis, the features learned from the convolution and pooling procedures were broken down into sequential parts and components, which were then fed to all recurring BiLSTM units for analysis. Following that, the O/P of the previous phase, i.e., the last step of BiLSTM, is fed into the fully connected layers for the purpose of predicting VTA. According to the BiLSTM, the activation function is represented by Equation (3).

$$f_t = \alpha \left( W[x_t, h_{t-1}, C_{t-1}] \right) + b_f \quad (3)$$

Here,  $x_t$  = input sequence,  $b_f$  = bias vector,  $W$  = weight, and  $\alpha$  = activation function,  $h_t$  = output of previous state,  $C_t$  = previous LSTM memory.

The output neurons in the three dense layers are 300, 50, and 3 in number, respectively. Finally, the final dense layer (12) is utilized by the SoftMax activation function, which produces the use of Softmax to classify the output class into three classes: VT, VF, and NSR. Using this strategy, the CNN-BiLSTM model was trained in this study. The training parameters are as follows: the learning rates are in the.001 range, and the regularisation is 0.01. The CNN-BiLSTM model employs a batch size of 34 for its calculations. The categorical cross-entropy function is used to calculate the loss parameter in this case. The CNN-BiLSTM model has been trained and tested in an iterative fashion. The number of iterations was determined to be 100 in this case. Finally, the overall performance of all records is used to calculate statistical performance. The BiLSTM model used for the classification of VTA consists of two layers. The seventh and eighth layers are made up of BiLSTM blocks that are 8 units and 16 units in size, respectively, and are inserted into the model. A dropout rate of 0.25 is introduced into the seventh and eighth layers, which helps to mitigate the problem of overfitting the model. After the BiLSTM layer, three dense layers with a total of 300, 50, and 3 neuronal units each are put on top of each other. To automatically classify NSR, VT, and VF, the last dense layer makes use of the SoftMax activation function in the previous dense layer. The learning rate of 0.001, the total number of epochs of 25, and the batch size of 32 were the additional parameters for training. As a result, the BiLSTM model divides the ECG signal into three classes: NSR, VT, and VF, each of which is determined by the features learned from the five-second windows.

### 3 Results

The CNN-BiLSTM model was trained and tested on an Intel i5 PC with 8 GB of RAM. The Keras libraries were used in the development of this model, which was written in the Python programming language. The experiment was carried out on two different types of ECG records, namely, CUDB and VFDB, and the results were compared. In accordance with the annotation file that was provided in the database, all records were framed into a five-second window length. An overall number of 12355 windows were obtained from both ECG records, with the specifics listed below in Table 2. In order to optimize model performance, the database has been randomly divided into the following ratios: 70-30, 60-40, 80-20, and 90-10 training test ratios. As shown in Table 4, the confusion matrices of the test dataset for the training-test ratios of 80-20 are shown, and we have built a confusion matrix for each of the training-test ratios in a similar manner.

Table4: A best result actual and predict class is shown using an 80-20 train-test ratio in the database.

|              |     | Predicted Class |     |     |
|--------------|-----|-----------------|-----|-----|
|              |     | NSR             | VT  | VF  |
| Actual Class | NSR | 2059            | 8   | 5   |
|              | VT  | 4               | 33  | 3   |
|              | VF  | 2               | 1   | 155 |
| Total        |     | 2072            | 240 | 158 |

The AC, SE, PR, and FS of the CNN-BLSTM model have been evaluated against the performance of the model, which has been measured using the equations 2, 3, 4, and 5 given below.

|   |
|---|
| $\text{OverallAccuracy} = \frac{\text{TruePositive}_1 + \text{TruePositive}_2 + \text{TruePositive}_3}{\text{Totaldatabase}} \quad (2)$ |
| $\text{Sensitivity}(SE) = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}} \quad (3)$                             |
| $\text{Precision}(PE) = \frac{\text{TruePositive}}{\text{TotalPredictedPositive}} \quad (4)$  |
| $F - \text{score}(FS) = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$                        |

Table 5-9 presents a summary of the classification abilities of the CNN-BiLSTM model based on statistical measurement results.

The 50-50 training–test ratio is used to randomly split the dataset in Table 5, and then the applied CNN-BiLSTM model for training and performance is measured on the test dataset. In this table, the train-to-test ratio is considered to be 50-50; the results of this experiment are shown in the table below.

Table 5  
The results of the CNN-BiLSTM model are provided by 50% of the testing datasets.

| <b>Class</b> | <b>Accuracy (%)</b> | <b>Precision (%)</b> | <b>Sensitivity (%)</b> | <b>F-Score (%)</b> | <b>Overall Accuracy (%)</b> |
|--------------|---------------------|----------------------|------------------------|--------------------|-----------------------------|
| NSR          | 98.45               | 99.21                | 99.32                  | 99.13              | 98.19                       |
| VT           | 98.58               | 90.42                | 96.01                  | 93.46              |                             |
| VF           | 99.32               | 95.12                | 95.08                  | 95.21              |                             |
| Total        | 98.78               | 94.91                | 96.80                  | 95.93              |                             |

A randomly chosen partitioning of the dataset in Table 6 is performed, this time using different training–test ratios, and then the applied CNN-BiLSTM model for training and performance is tested on the test dataset. Using this table, the train-to-test ratio is assumed to be 60-40, with the resulting results presented in the following table. Again, in this table, the CNN-BiLSTM model achieved better performance on the dataset.

Table 6  
The results of the CNN-BiLSTM model are provided by 40% of the testing datasets.

| <b>Class</b> | <b>Accuracy (%)</b> | <b>Precision (%)</b> | <b>Sensitivity (%)</b> | <b>F-Score (%)</b> | <b>Overall Accuracy (%)</b> |
|--------------|---------------------|----------------------|------------------------|--------------------|-----------------------------|
| NSR          | 98.75               | 99.04                | 99.12                  | 99.16              | 98.36                       |
| VT           | 98.77               | 94.20                | 94.23                  | 94.28              |                             |
| VF           | 99.21               | 93.42                | 95.32                  | 94.32              |                             |
| Total        | 98.91               | 95.55                | 96.22                  | 95.92              |                             |

Table 7 divides the dataset into 70–30 training–test ratios and measures performance using the CNN-BiLSTM model on the test dataset. The results obtained by the model are shown in the table below. At this train-test ratio, the model achieved better sensitivity while the training test ratio achieved better accuracy.

Table 7

The results of the CNN-BiLSTM model are provided by 30% of the testing datasets.

| <b>Class</b> | <b>Accuracy (%)</b> | <b>Precision (%)</b> | <b>Sensitivity (%)</b> | <b>F-Score (%)</b> | <b>Overall Accuracy (%)</b> |
|--------------|---------------------|----------------------|------------------------|--------------------|-----------------------------|
| NSR          | 98.46               | 99.32                | 99.43                  | 99.08              | 98.11                       |
| VT           | 98.70               | 93.46                | 93.26                  | 93.54              |                             |
| VF           | 99.06               | 94.01                | 93.08                  | 93.16              |                             |
| Total        | 98.74               | 95.59                | 95.25                  | 95.26              |                             |

In Table 8, the performance of the CNN-BiLSTM model is shown using the 80-20 training-test ratios of the dataset. The CNN-BiLSTM model in this table achieved better sensitivity and specificity than the others.

Table 8

The results of the CNN-BiLSTM model are provided by 20% of the testing datasets.

| <b>.Class</b> | <b>Accuracy (%)</b> | <b>Precision (%)</b> | <b>Sensitivity (%)</b> | <b>F-Score (%)</b> | <b>Overall Accuracy (%)</b> |
|---------------|---------------------|----------------------|------------------------|--------------------|-----------------------------|
| NSR           | 99.23               | 100                  | 99.01                  | 100                | 99.07                       |
| VT            | 99.35               | 96.13                | 97.18                  | 97.46              |                             |
| VF            | 99.55               | 95.24                | 98.28                  | 97.83              |                             |
| Total         | 99.37               | 97.12                | 98.15                  | 98.43              |                             |

Finally, in Table 9, the dataset is divided into 90-10 training test ratios and then the CNN-BiLSTM model is applied to evaluate the performance on the test dataset.

Table 9

The results of the CNN-BiLSTM model are provided by 10% of the testing datasets.

| <b>Class</b> | <b>Accuracy (%)</b> | <b>Precision (%)</b> | <b>Sensitivity (%)</b> | <b>F-Score (%)</b> | <b>Overall Accuracy (%)</b> |
|--------------|---------------------|----------------------|------------------------|--------------------|-----------------------------|
| NSR          | 99.12               | 100                  | 99.21                  | 99.13              | 98.96                       |
| VT           | 99.37               | 96.02                | 97.23                  | 96.34              |                             |
| VF           | 99.41               | 96.33                | 96.16                  | 96.19              |                             |
| Total        | 99.27               | 97.58                | 97.66                  | 97.28              |                             |

On both the databases (CUDB and VFDB) used in this study, we used five different training-test ratios, as shown in Tables 5, 6, 7 and 8. According to the given tables, we have been able to achieve the best results

by using the 20% of the test database, as illustrated in figure 4.

Comparing the proposed CNN-BiLSTM model, which achieved the best results on the CUDB and VFDB datasets, to other existing works presented in literature is shown in Table 8. As a result, our model has showed outstanding performance when compared to other existing models, as illustrated in Table 10.

## 4 Discussion

Based on the results in Table 10, the CNN-BiLSTM model achieves the highest level of accuracy. We compared the results of the newly proposed CNN-BiLSTM model with other methods in the literature that used the same database (CUDB and VFDB) to identify VTA rhythms using ECG signal analysis. We found that the results of the proposed model are significant compared to the excellent performance, indicating the effectiveness of the CNN-BiLSTM model in classifying NSR, VT, and VF rhythms. As can be seen in Table 4, we achieved the highest accuracy of 99.37% for three-class recognition, the highest accuracy ever achieved.

Table10. Compare this work with other existing published work.

| Author & Year        | Database            | Methods  | AC    | SE    |
|----------------------|---------------------|--|-------|-------|
| Othman et al. 2012   | NSRDB,CUDB          | Semantic   | -     | 96%   |
| Li et al. 2013       | CUDB,VFDB and AHADB | SVM and Genetic Algorithm                            | 96.30 | 96.2  |
| Tripathi et al. 2016 | MITDB,VFDB,CUDB     | VMD, Renyi and permutation entropy and Random forest | 97.23 | 96.54 |
| Xu et al. 2017       | MITDB,VFDB,CUDB     | Adaptive VMD and Boosted CART Classifier             | 98.29 | 97.32 |
| Acharya et al. 2018  | MITDB,VFDB,CUDB     | CNN  | 93.18 | 91.40 |
| Mohanty et al. 2018  | VFDB,CUDB           | DWT, Decision tree (C4.5)                            | 99.23 | 98.00 |
| Sabut et al. 2021    | VFDB,CUDB           | DNN  | 99.20 | 98.80 |
| Proposed 2021        | VFDB,CUDB           | CNN-BiLSTM   | 99.37 | 98.15 |

Othman et al. (2012) used semantic mining to classify VT and VF episodes by identifying three significant parameters that were damping coefficient, input signal and frequency from the ECG signal. The 4-second ECG signals were analyzed using this method. The VT and VF signals were classified from NSR signals without false detection, with a recognition sensitivity of 96% and 94%, respectively. Li et al. (2013) proposed the support vector classifier for classified VF/VT. A total of 14 features were extracted from a 5-second window length of the ECG signal. The best combinations of factors were then chosen using a genetic algorithm. As training, test, and validation of ECG datasets, three annotated ECG databases (the American Heart Association database, VFDB, and CUDB) were used. A variety of window sizes, ranging from 1 to 10 seconds, were evaluated. On the training data in the sample with a window size of 5 seconds, AC 98.1% and SE 98.4% were obtained from training datasets. The test data was validated using 5 fold cross-validation and found AC 96.3% and SE 96.2%. Tripathi et al. (2016) were proposed that an ECG signal be used for detecting and classifying episodes of VF. The decomposition of an ECG signal is performed by the use of variation mode decomposition. The energy, renyi entropy, and permutation entropy of the first three modes were calculated and used as diagnostic features. Mutual information feature-based assessment was used to select the optimal set of diagnostic functions for the given situation. In order to evaluate the performance of diagnostic functions, a random forest classifier is used. The AC and SE of the mutual evaluation based on the information and the Random forest classifier are 97% and 96.54% respectively. Xu et al. (2018) presented a classification system for VT and VF; and Using an adaptive variable mode decomposition algorithm, six features of the ECG signal were extracted and analyzed. To distinguish between VT and VF in an ECG signal, the boosted classification and regression method was applied. Three publicly available ECG databases with annotations were used for the training and testing of classifiers. To evaluate the effectiveness of the method, tenfold cross-verification methods were used. The 98.29% of AC and 97.321% of SE were obtained respectively. Acharya et al. (2018) presented an integrated index process based on nonlinear features such as fractal dimension, approximate entropy, and sample entropy, which were extracted from the decomposed ECG signal using DWT to predict sudden cardiac death. The t-value and predicted SCD of the extracted features were used to rank the features, with 92.11% and 98.68% accuracy in the k-nearest neighbour and SVM algorithms, respectively. Sabut et al. (2021) proposed classifying VF and VT conditions, they were developing a deep learning model and comparing the results with those obtained using other standard machine learning algorithms. The signal is decomposed using the wavelet transform, empirical mode decomposition, and variable mode decomposition approaches, and twenty-four features are extracted to form a hybrid model from a 5-second window. The AC of the DNN classifier was 99.2% and the SE was 98.8%.

## 5 Conclusions

The proposed CNN-BILSTM model is an efficient method to identify VT and VF from ECG variability signals. The performance of this E2E deep learning model has been better than traditional machine learning techniques in ECG analysis in terms of accuracy and sensitivity. In addition, the results showed that this model can improve the accuracy of VTA prediction without the need for data cleansing,

preprocessing, or feature extraction methods. Several experiments were carried out with the objective of analyzing the impact of the CNN-BiLSTM on the accuracy of 99.37%, the sensitivity of 98.15%, the precision of 97.12%, and the F-score of 98.43% when compared to results obtained using advanced methods. We come to the conclusion that the model has been performing better and accurately classifies VTA on large ECG databases. This system can help cardiologists with the early detection of VTA in their patients.

## 6 Abbreviations

VTA

ventricular tachyarrhythmia

CA

cardiac arrest

VF

ventricular fibrillation

VT

ventricular tachycardia

E2E

end-to-end

CNN

convolution neural network

BiLSTM

bidirectional long-short term memory network

MIT-BIH

Massachusetts Institute of Technology-Beth Israel Hospital

VFDB

MIT-BIH malignant ventricular arrhythmia database

CUDB

Creighton University VTA database

AC

accuracy

PE

precision

SE

sensitivity

FS

F-score

ReLU

rectifier linear unit

I/P

input  
O/P  
input

## Declarations

### Conflicting Interests

Author Manoj Kumar Ojha declares that he/she has no conflict of interest. Author Dr Sulochna Wadhvani declares that he/she has no conflict of interest. Author Dr Arun Kumar Wadhvani declares that he/she has no conflict of interest. Author Dr Anupam Shukla declares that he/she has no conflict of interest.

### Consent for publication

Informed consent was obtained from all participants involved in the study.

### Ethics approval and consent to participate

This article does not contain any studies with human participants or animals performed by any of the authors.

### Availability of data and materials

All datasets for supporting the conclusions of this article are available from the public data repository at the website of [https:// physionet.org](https://physionet.org).

### Authors' contributions

Manoj Kumar Ojha designed, carried out experiments and data analysis and drafted the manuscript. Dr. Sulochna Wadhvani, Dr Arun kumar Wadhvani, and Dr anupam shukla conceived of the study, and participated in research coordination. The authors read and approved the final manuscript.

## References

Kibos, A. S., Knight, B. P., Essebag, V., Fishberger, S. B., Slevin, M., & Tintoiu, I. C. (Eds.). (2013). *Cardiac Arrhythmias: From Basic Mechanism to State-of-the-Art Management*. Springer Science & Business Media. URL <https://doi.org/10.1007/978-1-4471-5316-0>

Theuns, D. A., Bhagwandien, R. E., Szili-Torok, T., Zijlstra, F., & Yap, S. C. (2019). Evaluation of recurrent ventricular tachyarrhythmias in patients who survived out-of-hospital cardiac arrest due to ventricular fibrillation: eligibility for subcutaneous implantable defibrillator therapy. *Journal of Interventional Cardiac Electrophysiology*, 55(3), 317-323. URL <https://doi.org/10.1007/s10840-018-0490-4>

- Chen, T. M., Huang, C. H., Shih, E. S., Hu, Y. F., & Hwang, M. J. (2020). Detection and classification of cardiac arrhythmias by a challenge-best deep learning neural network model. *Iscience*, 23(3), 100886. URL <https://doi.org/10.1016/j.isci.2020.100886>
- Benjamin, E. J., Blaha, M. J., Chiuve, S. E., Cushman, M., Das, S. R., Deo, R., ...&Muntner, P. (2017). Heart disease and stroke statistics—2017 update: a report from the American Heart Association. *circulation*, 135(10), e146-e603. URL <https://doi.org/10.1161/CIR.0000000000000485>
- Kleiger, R. E., Miller, J. P., Bigger Jr, J. T., & Moss, A. J. (1987). Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *The American journal of cardiology*, 59(4), 256-262. URL [https://doi.org/10.1016/0002-9149\(87\)90795-8](https://doi.org/10.1016/0002-9149(87)90795-8)
- Pandey, S. K., &Janghel, R. R. (2021). Classification of electrocardiogram signal using an ensemble of deep learning models. *Data Technologies and Applications*. URL <https://doi.org/10.1108/DTA-05-2020-0108>
- Pandey, S. K., Janghel, R. R., & Gupta, V. (2021). Cardiac arrhythmia detection and classification from ECG signals using XGBoost classifier. *Machine Learning Algorithms and Applications*, 141-157. URL <https://doi.org/10.1002/9781119769262.ch8>
- Alonso-Atienza, F., Morgado, E., Fernandez-Martinez, L., García-Alberola, A., &Rojo-Alvarez, J. L. (2013). Detection of life-threatening arrhythmias using feature selection and support vector machines. *IEEE Transactions on Biomedical Engineering*, 61(3), 832-840. URL <https://doi.org/10.1109/TBME.2013.2290800>
- Li, Q., Rajagopalan, C., & Clifford, G. D. (2013). Ventricular fibrillation and tachycardia classification using a machine learning approach. *IEEE Transactions on Biomedical Engineering*, 61(6), 1607-1613. URL <https://doi.org/10.1109/TBME.2013.2275000>
- Panda, R., Jain, S., Tripathy, R. K., &Acharya, U. R. (2020). Detection of shockable ventricular cardiac arrhythmias from ECG signals using FFREWT filter-bank and deep convolutional neural network. *Computers in Biology and Medicine*, 124, 103939. <https://doi.org/10.1016/j.compbiomed.2020.103939>
- Alwan, Y., Cvetković, Z., & Curtis, M. J. (2017). Methods for improved discrimination between ventricular fibrillation and tachycardia. *IEEE Transactions on Biomedical Engineering*, 65(10), 2143-2151. URL <https://doi.org/10.1109/TBME.2017.2785442>
- Plesinger, F., Andrla, P., Viscor, I., Halamek, J., &Jurak, P. (2018, September). Fast Detection of Ventricular Fibrillation and Ventricular Tachycardia in 1-Lead ECG from Three-Second Blocks. In *2018 Computing in Cardiology Conference (CinC)* (Vol. 45, pp. 1-4). IEEE. URL <https://doi.org/10.22489/CinC.2018.037>
- Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M.,Ivanov, P. C., Mark, R. G.& Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex

physiologic signals. *circulation*, 101(23), e215-e220.

URL <https://doi.org/10.1161/01.CIR.101.23.e215>

Pandey, S. K., & Janghel, R. R. (2019). Automatic detection of arrhythmia from imbalanced ECG database using CNN model with SMOTE. *Australasian physical & engineering sciences in medicine*, 42(4), 1129-1139. URL <https://doi.org/10.1007/s13246-019-00815-9>

Jain, D., Kumar, A., & Garg, G. (2020). Sarcasm detection in mash-up language using soft-attention based bi-directional LSTM and feature-rich CNN. *Applied Soft Computing*, 91, 106198.

URL <https://doi.org/10.1016/j.asoc.2020.106198>

Othman, M. A., Safri, N. M., Ghani, I. A., & Harun, F. K. C. (2012). Characterization of ventricular tachycardia and fibrillation using semantic mining. *Computer and Information Science*, 5(5), 35. URL <https://doi.org/10.5539/cis.v5n5p35>

Tripathy, R. K., Sharma, L. N., & Dandapat, S. (2016). Detection of shockable ventricular arrhythmia using variational mode decomposition. *Journal of medical systems*, 40(4), 79. URL <https://doi.org/10.1007/s10916-016-0441-5>

Xu, Y., Wang, D., Zhang, W., Ping, P., & Feng, L. (2018). Detection of ventricular tachycardia and fibrillation using adaptive variational mode decomposition and boosted-CART classifier. *Biomedical Signal Processing and Control*, 39, 219-229. URL <https://doi.org/10.1016/j.bspc.2017.07.031>

Acharya, U. R., Fujita, H., Oh, S. L., Raghavendra, U., Tan, J. H., Adam, M., ... & Hagiwara, Y. (2018). Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network. *Future Generation Computer Systems*, 79, 952-959.

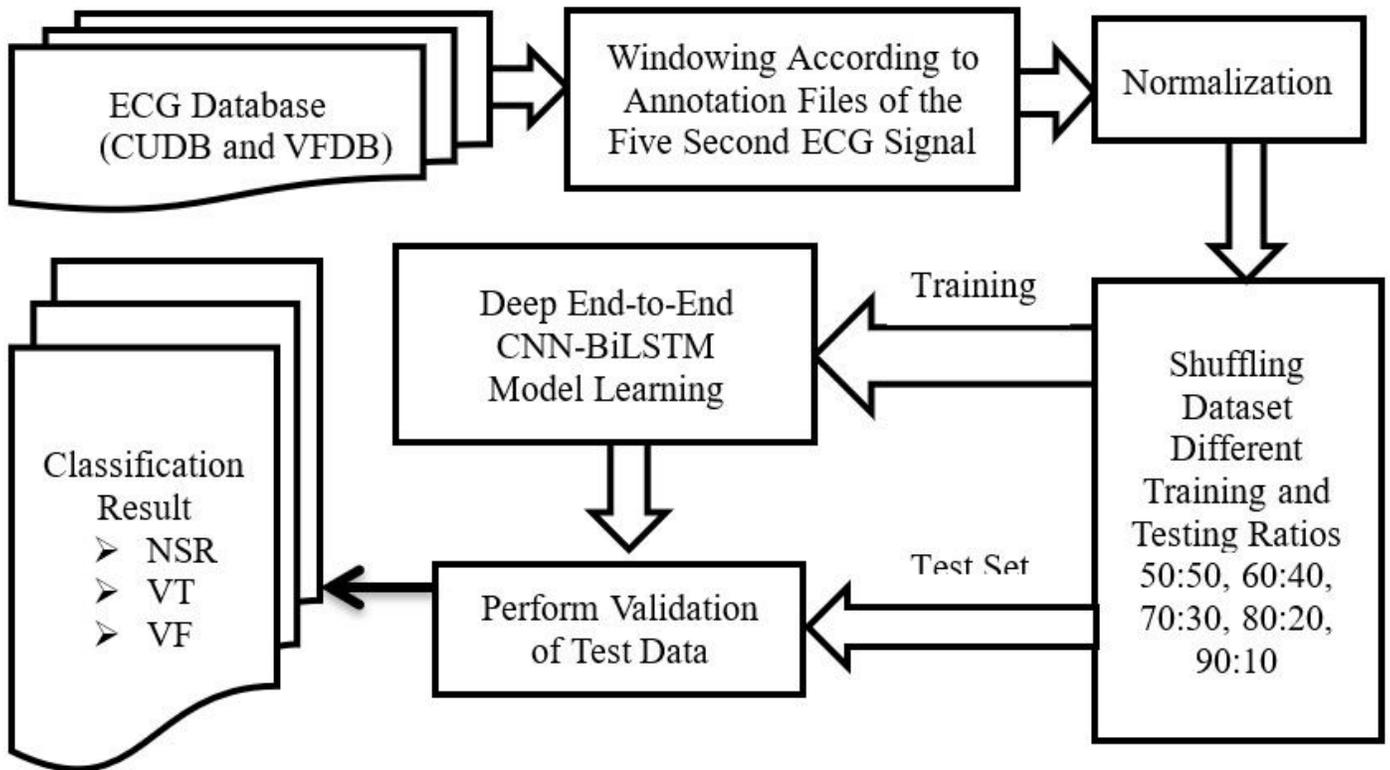
URL <https://doi.org/10.1016/j.future.2017.08.039>

Mohanty, M., Sahoo, S., Biswal, P., & Sabut, S. (2018). Efficient classification of ventricular arrhythmias using feature selection and C4.5 classifier. *Biomedical Signal Processing and Control*, 44, 200-208.

URL <https://doi.org/10.1016/j.bspc.2018.04.005>

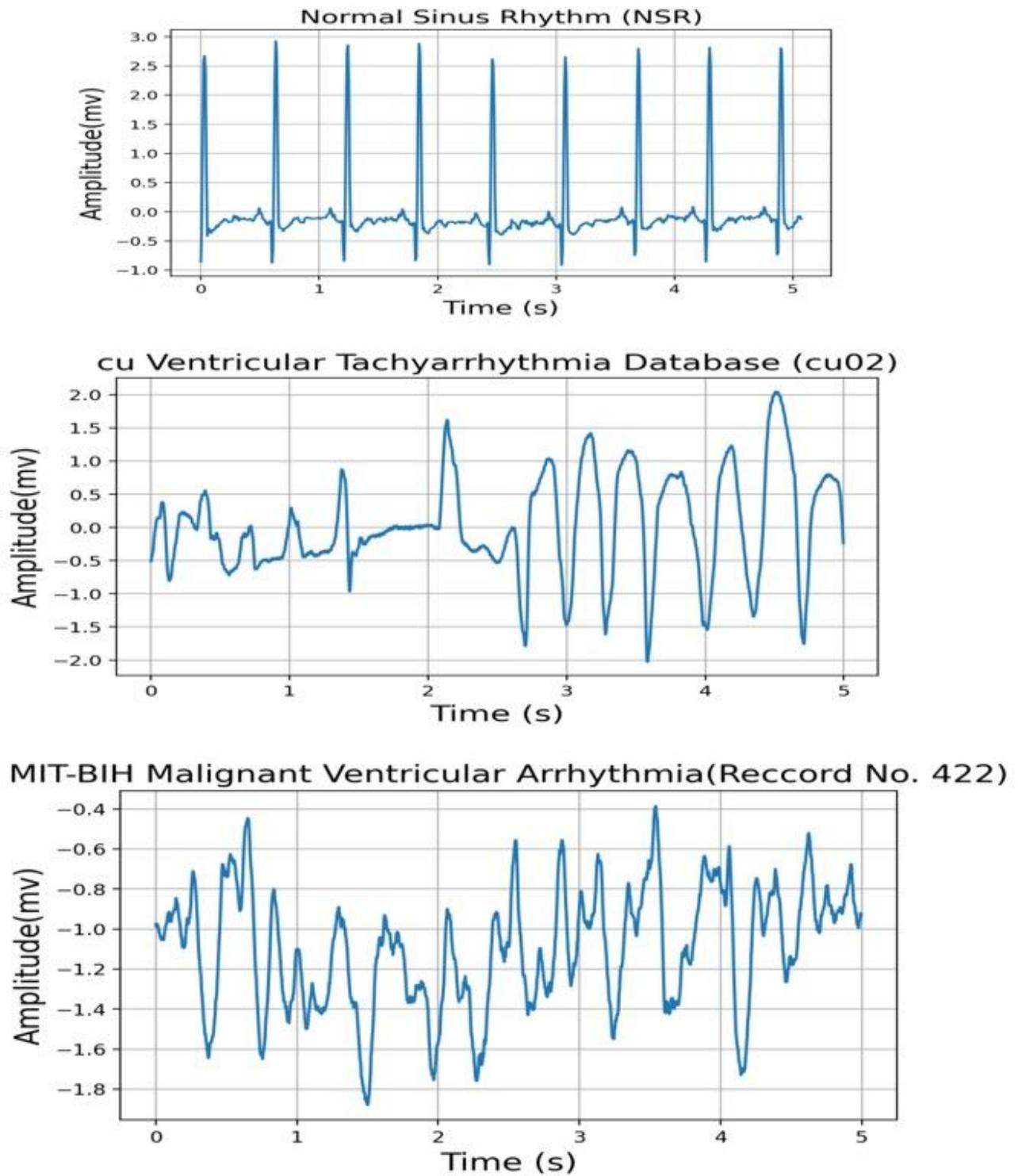
Sabut, S., Pandey, O., Mishra, B. S. P., & Mohanty, M. (2021). Detection of ventricular arrhythmia using hybrid time-frequency-based features and deep neural network. *Physical and Engineering Sciences in Medicine*, 44(1), 135-145. URL <https://doi.org/10.1007/s13246-020-00964-2>

## Figures



**Figure 1**

Work flow diagram of Classification model for prediction of NSR, VT and VF classes.



**Figure 2**

Typical ECG sample 5 second signal of (a) NSR, (b) VT, and (c) VF

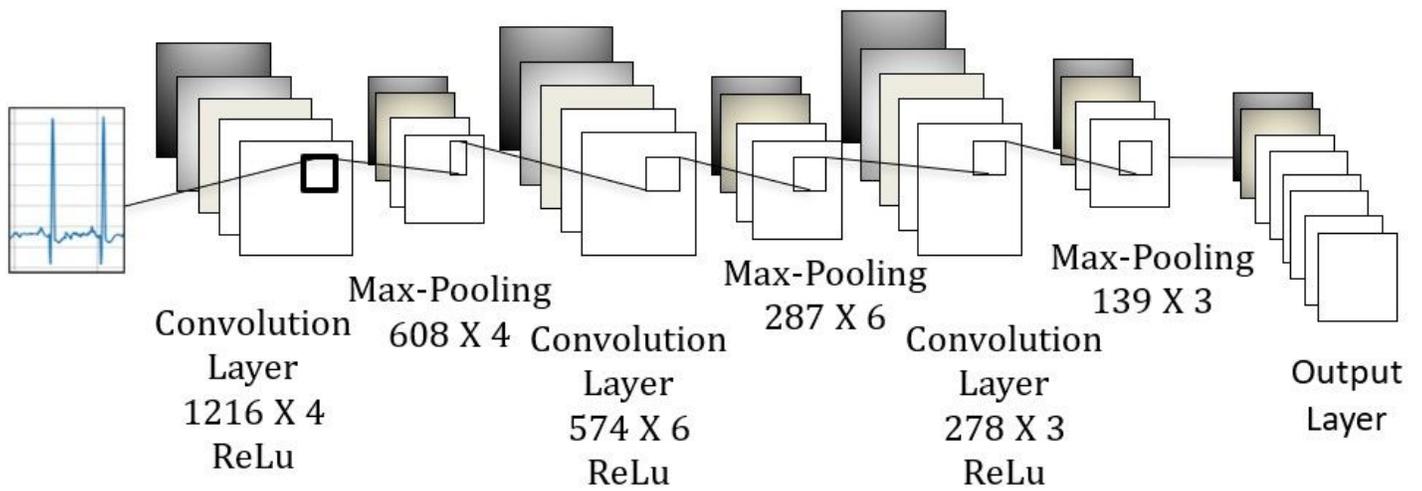


Figure 3

Fig 2.Convolution Neural Network(CNN) model.

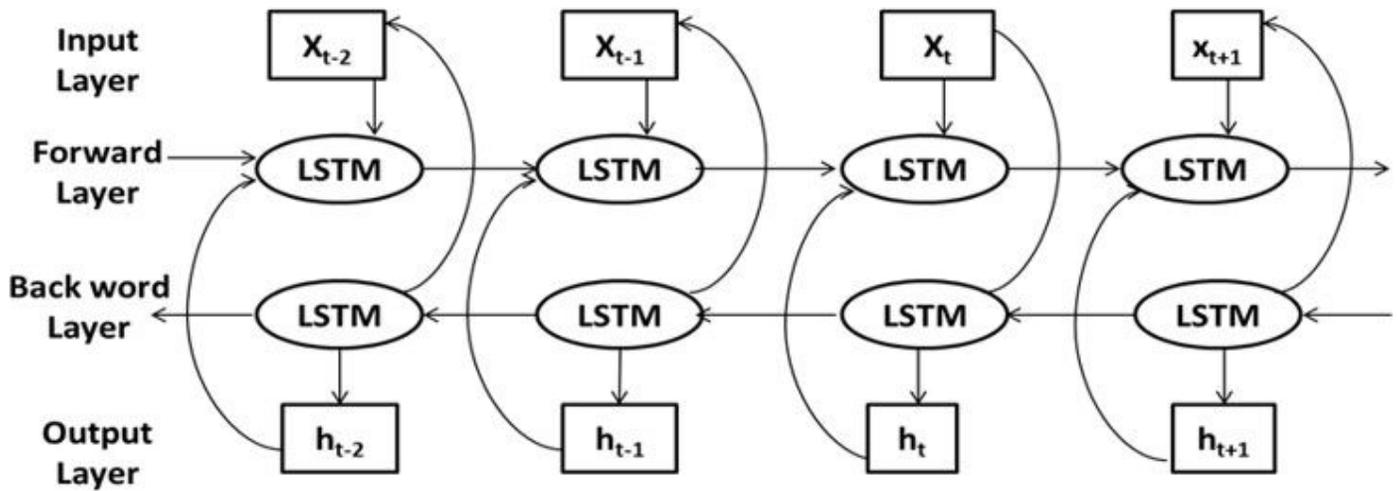
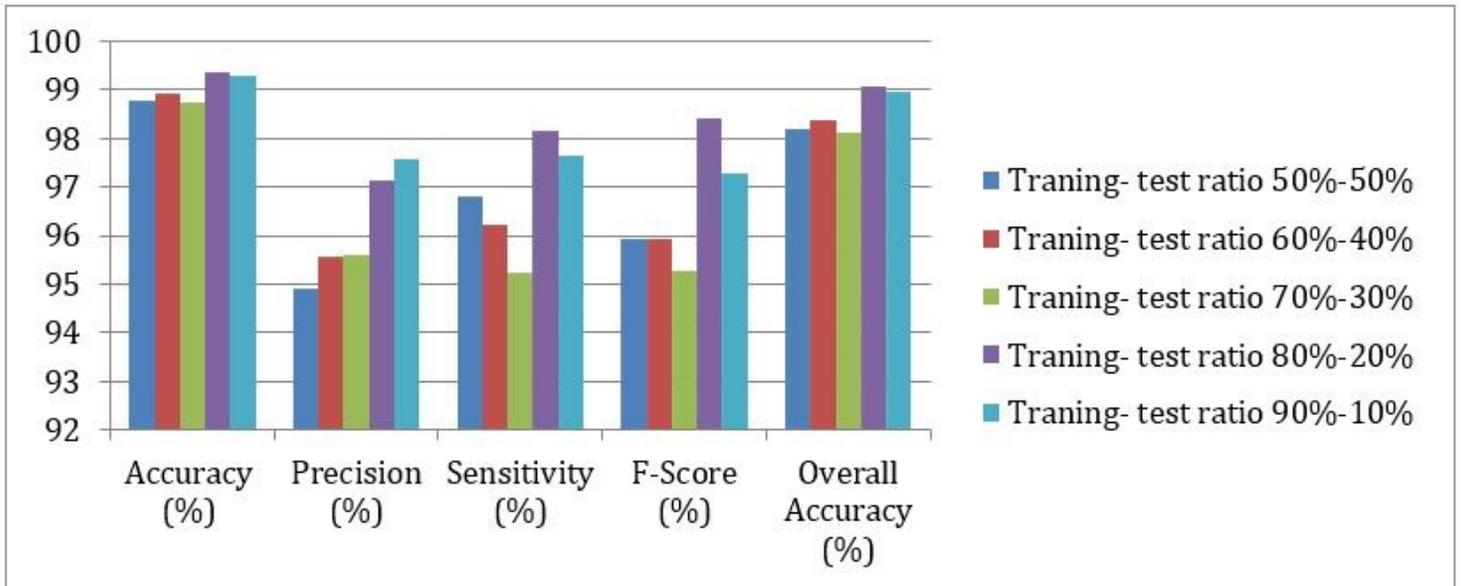


Figure 4

Fig.3 Bidirectional Long-short Term Memory (BiLSTM) model.

Here,  $x_t$ = Input vector,  $h_t$  = Output value.



**Figure 5**

Fig.4 CNN-BiLSTM model validation on different training-test data ratios.