Computers in Biology and Medicine 120 (2020) 103726

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

Application of deep learning techniques for heartbeats detection using ECG signals-analysis and review

Fatma Murat [a], Ozal Yildirim [b], Muhammed Talo [c], Ulas Baran Baloglu [d], Yakup Demir [a], U. Rajendra Acharya [e,f,g,*]

[a] Department of Electrical and Electronics Engineering, Firat University, Elazığ, 23000, Turkey
[b] Department of Computer Engineering, Munzur University, Tunceli, 62000, Turkey
[c] Department of Software Engineering, Firat University, Elazig, Turkey
[d] Department of Computer Engineering, University of Bristol, Bristol, UK
[e] Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore
[f] Department of Bioinformatics and Medical Engineering, Asia University, Taichung, Taiwan
[g] International Research Organization for Advanced Science and Technology (IROAST), Kumamoto University, Kumamoto, Japan

ARTICLE INFO

Keywords:
Arrhythmia detection
Deep learning
ECG classification
CNN
LSTM

ABSTRACT

Deep learning models have become a popular mode to classify electrocardiogram (ECG) data. Investigators have used a variety of deep learning techniques for this application. Herein, a detailed examination of deep learning methods for ECG arrhythmia detection is provided. Approaches used by investigators are examined, and their contributions to the field are detailed. For this purpose, journal papers have been surveyed according to the methods used. In addition, various deep learning models and experimental studies are described and discussed. A five-class ECG dataset containing 100,022 beats was then utilized for further analysis of deep learning techniques. The constructed models were examined with this dataset, and results are presented. This study therefore provides information concerning deep learning approaches used for arrhythmia classification, and suggestions for further research in this area.

1. Introduction

Arrhythmias are an important group of cardiovascular disorder. An arrhythmia may occur on its own or in conjunction with other cardiovascular diseases [1]. Because of the high mortality rates in heart disease, early diagnosis and definitive differentiation of arrhythmias are important to patient treatment [2]. The most commonly used solution for arrhythmia detection is with the recording of the electrocardiogram (ECG), which displays the electrical activity of the heart over time from electrodes placed on the skin. The ECG leads, which capture the electrical potential of the heart from different angles and positions, can be used to indicate disease state via abnormalities in waveforms or rhythms [3]. The ECG is a record of the electrical characteristics of the heartbeat and has become one of the most important tools in the diagnosis of heart disease. It is crucial to diagnose a broad spectrum of abnormalities, from arrhythmia to acute coronary syndrome [4]. It contains much information not only about the structure of the heart, but also concerning the function of the electrical conduction system [5]. Different types of arrhythmias correspond to different patterns that can be represented by different ECG waveforms [6]. These patterns contain information about heart function and condition. Therefore, monitoring and recognition of ECG signals is an important issue in biomedicine [7].

Arrhythmia can be represented by a slow, rapid, or irregular heartbeat, and can be grouped as life-threatening versus non-life-threatening. According to the association for the advancement of medical instrumentation (AAMI), non-life-threatening arrhythmias can be divided into five main classes: non-ectopic (N), supraventricular ectopic (S), ventricular ectopic (V), fusion (F), and unknown (Q) [8]. Automatic arrhythmia detection based on the ECG provides great convenience as it does not require physicians to manually analyze the signals, and also helps people monitor heart conditions using wearable devices. Automated accurate arrhythmia detection requires machine assistance in the treatment of cardiovascular diseases [3].

Advances in machine learning have enabled the efficient development of computer-based diagnostic (CAD) systems and their application to many areas. The development of intelligent systems in the field of

* Corresponding author. Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore.
E-mail address: aru@np.edu.sg (U.R. Acharya).

https://doi.org/10.1016/j.compbiomed.2020.103726
Received 3 February 2020; Received in revised form 16 March 2020; Accepted 21 March 2020
Available online 8 April 2020

health, the processing of large amounts of raw data, and obtaining meaningful results from these data are of great interest [7]. Computer-aided interpretation has become increasingly essential in the field of healthcare since it recognized more than 50 years ago [4]. With the emergence of these systems, the workload of cardiologists has diminished, and the computational effectiveness and accuracy of disease detection have increased. In order to minimize visual errors and to compensate for manual interpretation, researchers began developing CAD systems to assist in the diagnosis using the ECG [5]. An effective CAD system requires a powerful pattern classifier, as well as a salient feature extractor that can extract significant information from the hidden layers of raw data [9]. Conventional methods require the use of handcrafted features for signal preprocessing, waveform detection, feature extraction, and classification. The encoded features are generally designed and selected by trial-and-error or experience. Therefore, these systems require more specific expertise in various domains and hence obtaining useful features is a time-consuming process. Deep learning techniques have been developed to overcome these difficulties and to provide improved detection rate without the use of fixed coded features [10].

Traditional neural network methods [11–17] and kernel-based classifiers [18–20] were among the most commonly used methods for arrhythmia data classification. These methods generally use inputs obtained by feature extraction instead of raw input signals. Since the desired high performance of raw input data could not be achieved, prior research focused on feature extraction methods rather than network structure. Time, frequency, statistical and non-linear properties are obtained by such approaches as the wavelet transform, Fourier transform, and higher order spectra (HOS) [21–24,67–70]. Due to the high dimensionality of feature vectors resulting from transformation methods, the size of these feature vectors is reduced with statistical approaches or techniques such as PCA. Deep learning techniques, which have recently become popular in machine learning, provide an effective means for knowledge gathering without need of feature engineering [25]. Deep learning structures using sufficient ECG input and dataset training have the potential to learn all previously important manual features, as well as previously unknown features [4]. In the field of machine learning, efficient use of multi-layer networks has been achieved due to both the introduction of effective approaches to solve optimization problems, as well as hardware advances such as implementation of graphical processing units. Innovative approaches for error propagation and developing techniques such as batch normalization, residual connections, and depthwise separable convolution, have facilitated the training of networks with many layers [26]. This area, a new sub-branch of machine learning called deep learning, has rapidly proliferated, leading to successful applications to process the ECG [27].

In this study, we examined the studies in the literature which have utilized deep learning methods for processing ECG signals. The contributions of these studies to the field are emphasized, and the methods proposed by them are analyzed. We also presented several applications on a heartbeat dataset containing 100,022 beats in five classes, for the purpose of evaluating commonly used deep learning techniques. This paper provided a comprehensive information on the classification of ECG signals using deep learning techniques which is the state-of-art techniques by implementing various models. We have also reviewed many related articles, identified the current challenges, suggested possible solutions, and delineated the popular trends with critical recommendations.

2. Material and methods

We employed ECG data from five different classes, containing 100,022 beats obtained from the MIT-BIH arrhythmia database, to evaluate deep learning techniques commonly used in the literature. The results were analyzed by applying various applications - from basic models to more complex models.

2.1. ECG dataset

We have used 100,022 ECG beats from PhysioNet MIT-BIH Arrhythmia public database [28] for the evaluation of deep learning models. The MIT-BIH arrhythmia database includes N, S, V, F and Q main classes with each class having many sub-classes. We have used few groups in these classes for this work. These groups are chosen as they are widely used in the literature. The beats in our used dataset consist of five classes: normal beats (N), atrial premature beats (APB), left bundle branch block (LBBB), right bundle branch block (RBBB) and premature ventricular contraction (PVC). ECG data from modified limb lead II signals were organized into segments with 260 samples. Segments (one single beat) of continuous beats of 48 half-hour records of 47 patients were used for this work. The signals in each segment consist of 99 samples before the R peak, and 160 samples after the pulses. Beats tags were annotated by multiple cardiologists. Table 1 exhibits the classes, the number of pulses, and waveform examples of these classes in the arrhythmia data.

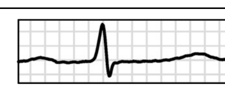
2.2. Experimental setup

In the application of deep learning techniques, we have used the Keras with TensorFlow backend Deep Learning Library. Raw ECG signals were first scaled in the range of 0–1, and then standardized. The scikit-learn library was used for pre-processing. An early stopping technique was utilized to determine how long the learning process would continue. With this technique, loss values are monitored, and the training process is stopped when model begins to overfit. Thus, learning was stopped so that overfitting problems do not occur for each network. Few common hyper-parameter adjustments of models are determined for learning rate value of 0.001, and batch size of 128. Optimizers and other parameters are varied depending on the examined networks. The related adjustments are presented separately in each experimental study. The computer used in the experimental studies has an Intel Core i7-7700HQ 2.81 GHz CPU, 16 GB memory and 8 GB NVIDIA GeForce GTX 1070 graphics card. To ensure consistency in all experimental studies, the data was divided into 80% training, 10% validation and 10% testing, and the same datasets were used in all proposed models. In order to take into account the imbalanced data distributions in the classes during the training of the models, a class weight was assigned to each class using the scikit-learn library. Accuracy, Sensitivity, Specificity, Precision, and F-Score performance metrics were used to evaluate the results obtained for the test data.

3. Applications and review

In this section, a comprehensive examination has been carried out for

Table 1
Classes, number of beats, and waveform examples in the arrhythmia dataset.

Beat Types	Number of Patients	Number of Beats	Waveform Sample
1. Normal Beats (N)	47	75020	
2. Atrial Premature Beat (APB)		2546	
3. Left Bundle Branch Block (LBBB)		8072	
4. Right Bundle Branch Block (RBBB)		7255	
5. Premature Ventricular Contraction (PVC)		7129	

arrhythmia detection from basic deep learning models and more complex network models. Under the deep learning techniques, the studies for arrhythmia analysis are detailed, and some of these techniques are applied on arrhythmia datasets, with the results then being evaluated.

3.1. Deep neural networks

Deep neural networks (DNN) are classical neural networks (NN) that are hierarchically bound and that contain many hidden layers [30]. In the arrhythmia classification problem, classical NN approaches and SVM classifiers have been replaced by DNN based classification models. DNN networks [31] input by raw ECG signals do not require preliminary feature extraction. The use of some temporal features in combination with raw signals has been shown to improve the performance of deep networks [32,33]. However, it should not be ignored that there is an additional cost in the stages of obtaining temporal features as RR interval. The denoising autoencoder (DAE) and stacked denoising autoencoder (SDAE) [33–36] approaches are frequently used to feed DNN classifiers with more suitable features, in contrast to the capability of shallow classifiers. The representations obtained from hidden layers of autoencoders are input to the Softmax layer, and classification operations are performed. Some studies have used the active learning structure to identify the most valuable beats for the DNN fine-tuning process [33–35]. In order to avoid overfitting during learning, effective solutions have been presented that fuse existing and previous Softmax outputs [33]. In addition, there are studies suggesting the use of one-dimensional ECG signals in model inputs by converting them into time-frequency images [36]. In studies using DNN, the effects of layer increments on classification have been an important parameter to address.

In our investigation, described herein, we studied a simple model with single hidden layer versus deeper models, to observe the behavior of end-to-end DNN structures, and the number of parameters useful for networks in arrhythmia classification. The inputs of these models consisted of raw ECG signals, and their output was composed of five ECG classes. Fig. 1 shows graphs of the performance of a classical NN with a single hidden layer of 128 units in the training phase, along with several parameters (activation function, optimizer, and loss function).

In the NN-1 model (see Fig. 1 (b)), the graph of loss values during the training phase showed that the model performance improved in small steps, and this process is time consuming. This result caused by the derivative, which becomes too small because the values of the selected sigmoid activation function are too high or too low. Also, it can lead to the vanishing gradient problem which is common in gradient-based learning methods. Training of the NN-1 network could be completed at approximately 2000 epochs. Each epoch lasted 2s on average. This means approximately 1 h of training would be needed with our existing GPU hardware. In CPU mode, considering that this time decelerates, the cost of training will increase considerably. When the ReLU activation function (see Fig. 1 (c)), which is frequently used in deep learning rather than the sigmoid function in the NN-2 network, was selected (see Fig. 1 (c)), and the gradient was updated with small values, so that the learning phase duration was approximately 1200 epochs. In the NN-3 and NN-4 networks, when the Adam optimizer (Fig. 1 (d)) is selected, unlike the NN-2 model (Fig. 1 (d)), it completes learning by decreasing to 0.005 values in a very short time period of approximately 11 epochs. The ReLU activation function and the Adam optimizer have improved both detection performance and time cost.

In Fig. 2, the graphs of training performances on the ECG data of the models with 2, 3, and 5 hidden layers are given, respectively, to observe the effect of layer increase on classification. These models are termed DNN-1, DNN-2 and DNN-3, respectively. The hidden layer units in these models are 128, and the activation functions are selected as ReLU and Adam optimizer.

The increase in the number of layers using the existing arrhythmia dataset had a positive effect on the training stage. On the other hand, with the increase in the depth of the layer, the performance of the model was improved. We have provided various performance criteria to evaluate the trained models on unseen test data. The performance values are given in Table 2.

The best accuracy performance of 99.11% was achieved with the DNN-3 model with five hidden layers. The sigmoid activation function used in the NN-1 model and the SGD optimizer led to a prolonged training period with very small changes in gradient. According to these

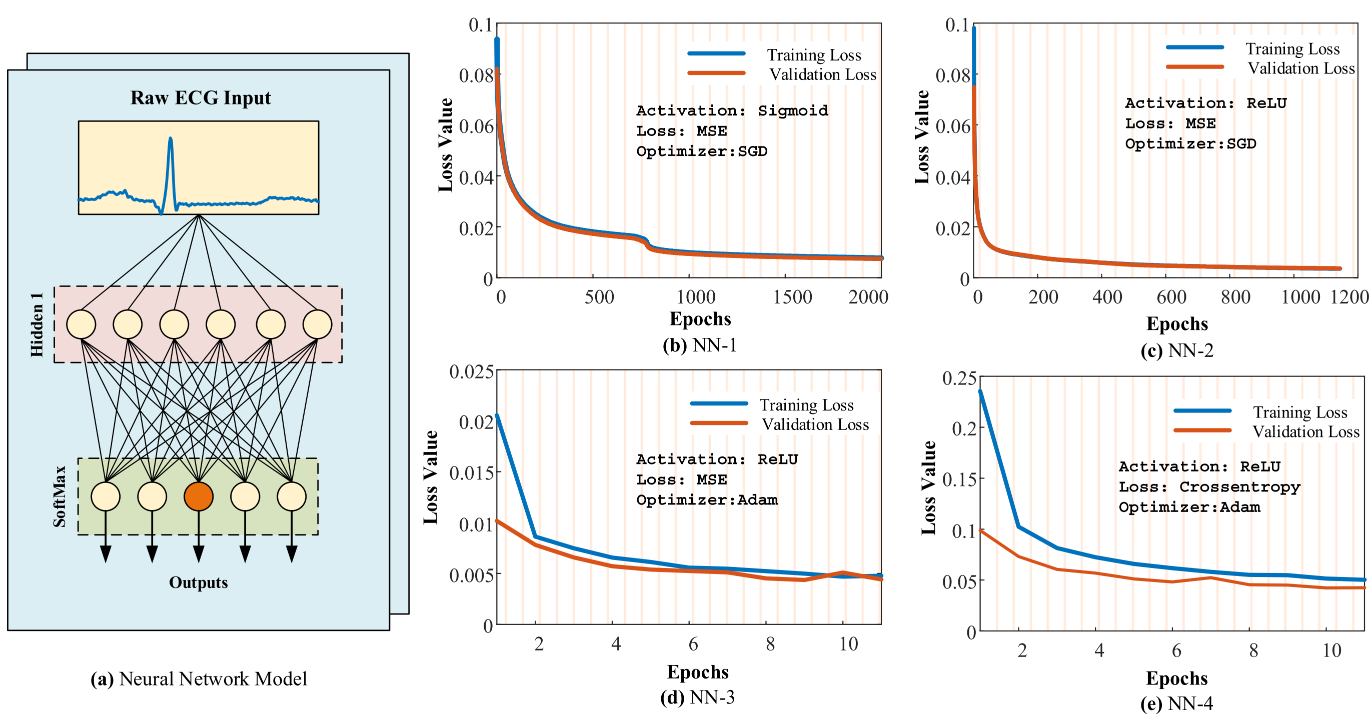


Fig. 1. An illustration of the effects of activation functions, optimizers and loss functions on learning for a single hidden layer network. b) Sigmoid, MSE loss, and SGD optimizer, c) ReLU, MSE loss, and SGD optimizer, d) ReLU, MSE loss, and Adam optimizer, e) ReLU, categorical cross entropy loss, and Adam optimizer.

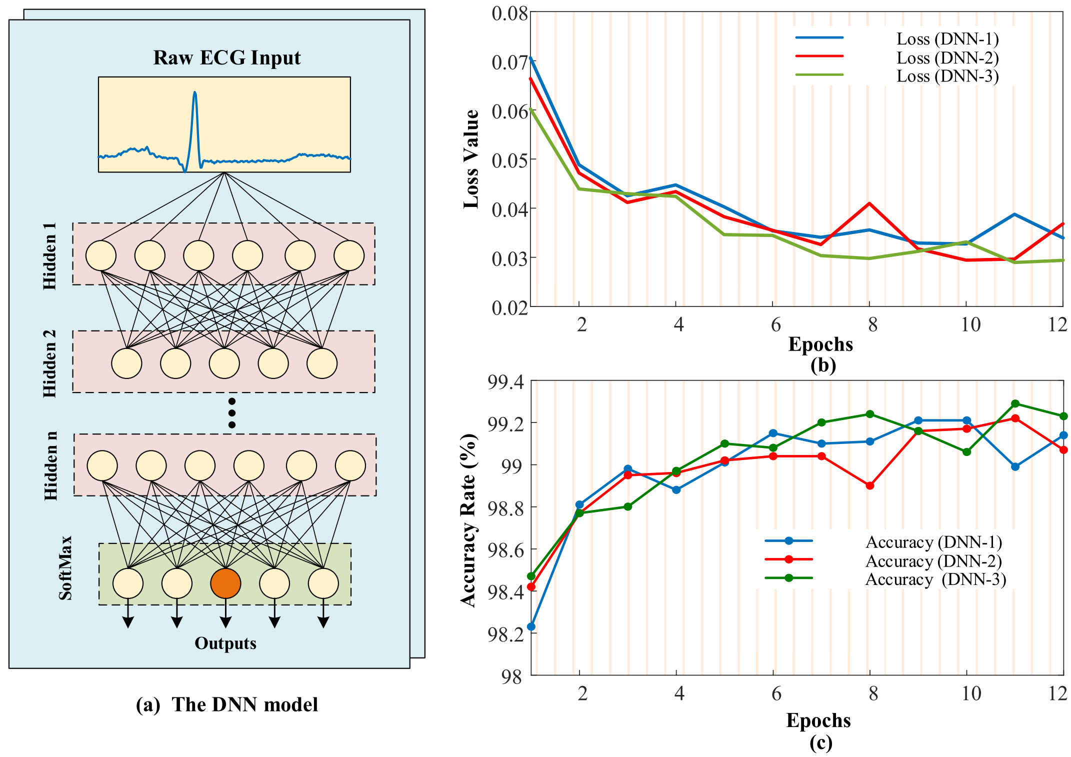


Fig. 2. DNN models designed in the study and some performance graphs of these models during the training stages. a) The DNN models, b) Loss graphs and c) Accuracy graphs.

Table 2
Performance values of DNN and NN models on arrhythmia test data.

Models	Total Training Time	Overall Sensitivity (%)	Overall Specificity (%)	Overall Precision (%)	Overall F-Score (%)	Overall Accuracy (%)
NN-1	4000 s	89.66	98.70	93.93	91.63	97.73
NN-2	2290 s	94.30	99.23	97.54	95.77	98.73
NN-3	22 s	93.02	99.13	97.73	95.02	98.56
NN-4	22 s	94.22	99.21	97.48	95.69	98.67
DNN-1	24 s	95.34	99.34	97.50	96.37	98.85
DNN-2	36 s	95.89	99.43	97.76	96.78	98.99
DNN-3	36 s	96.45	99.53	97.72	97.05	99.11

results, besides the increase in the number of layers, the ReLU activation function and Adam optimizer have a significant effect on the performance of DNN networks.

3.2. Convolutional neural networks

The process of learning differential representations to map input data to target outputs is the basic step of machine learning. Traditional machine learning methods use various hand-engineered features to obtain representations of input data. In the case of deep learning, there is an

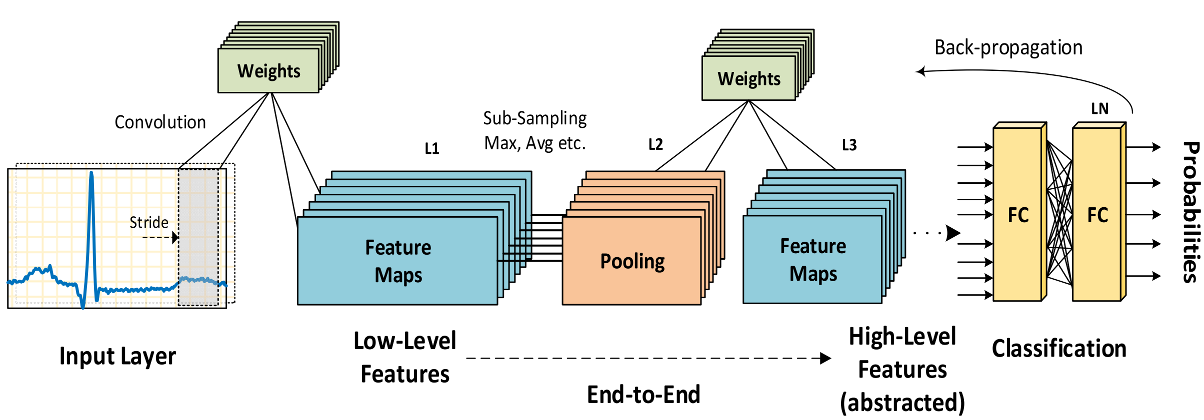


Fig. 3. A simple 1D convolutional neural network structure which has convolution, pooling and fully-connected layering.

automatic learning process from the low-level representations obtained from multiple layers, to the higher abstract representations in the training stage (see Fig. 3) [25]. Convolutional neural networks learn useful representations of input data in an end-to-end structure using the convolution operator. AlexNet's success in the ImageNet [37] competition in 2012 has made CNN applications popular, and their use in the medical field has become widespread.

As the feature extraction process, which plays a critical role in ECG signal classification, is automated with convolutional neural networks, the use of CNN has become widespread in this field. These networks are used to classify patient-specific beats [6,38] and long duration ECG signals containing multiple rhythm classes [39,40,66], to detect different interval ECG segments [41], atrial fibrillation [41–48], and different types of ECG beats [8,49].

Models in CNN-based arrhythmia classification studies are prepared at depths reaching 9-layers [8], 11-layers [41], 16-layers [40] and 34-layers [39,42]. In these models, there are many hierarchically connected layers where the feature maps obtained with convolutional layers are sub-sampled with pooling layers and fully connected layers in the last stage of the model. In addition to these layers, regularized layers such as batch normalization and dropout are also employed [6,39,40,45]. These make the model more resistant to overfitting, so that the learning process is more effective. As layer size increases, optimization of the network becomes more difficult. Rajpurkar et al. [39] have employed residual network-like shortcut connections in 34-tier models to solve this problem in arrhythmia classification. Another difficulty in designing the networks is to determine the filter length and number of convolution layers. Filter length is usually selected in small sizes, such as 3×1 or 5×1. The main reasons for this are reduction of calculation cost and ability to distinguish signals with small differences between them [6]. Yet, it is seen that these filter lengths are chosen to be larger in networks designed for long duration signals. For example, Yildirim et al. [40] used a 50×1 size filter in the first convolution layer for 10 s ECG signals, and Rajpurkar et al. [39] used a filter length of 16×1 in convolution layers for 30 s. The filter numbers are generally selected as multiples of two. Lu et al. [49] showed that variable learning rate is more beneficial than constant rate learning. In addition, imbalance data in some arrhythmia datasets may yield misleading results in classifier performance [5,49]. Jiang et al. [5] discussed this problem in detail in heartbeat classification, and proposed three different solution methods.

Herein, in order to evaluate the performance of the operation of CNN networks on ECG input signals according to the number of layers, the models in Fig. 4 with different size layers are designed and developed.

While designing these models, only CNN base layers such as convolution, sub-sampling and dense layers were used. The convolution layer numbers of the models were increased and the layer parameters were designed to be the same. Our aim is to observe the impact of deepening CNN networks on existing data. The filter numbers are set to increase continuously to 32, 64, 128, and 256. Kernel sizes 5, 3, 5, and 3 were chosen, respectively. The CNN networks designed for the experimental study were trained on the arrhythmia data separately, and changes in accuracy and loss values for the training and validation sets are given in Fig. 5.

For the data analyzed, there were no significant differences between CNN performances during the training phase. However, it was found that the best performance was obtained for the CNN-4 model containing four convolution layers. Accordingly, it can be said that more distinctive representations of the input data are obtained due to the increase in number of layers. The values of some performance criteria on the test data of the models are given in Table 3. The F1-score and sensitivity

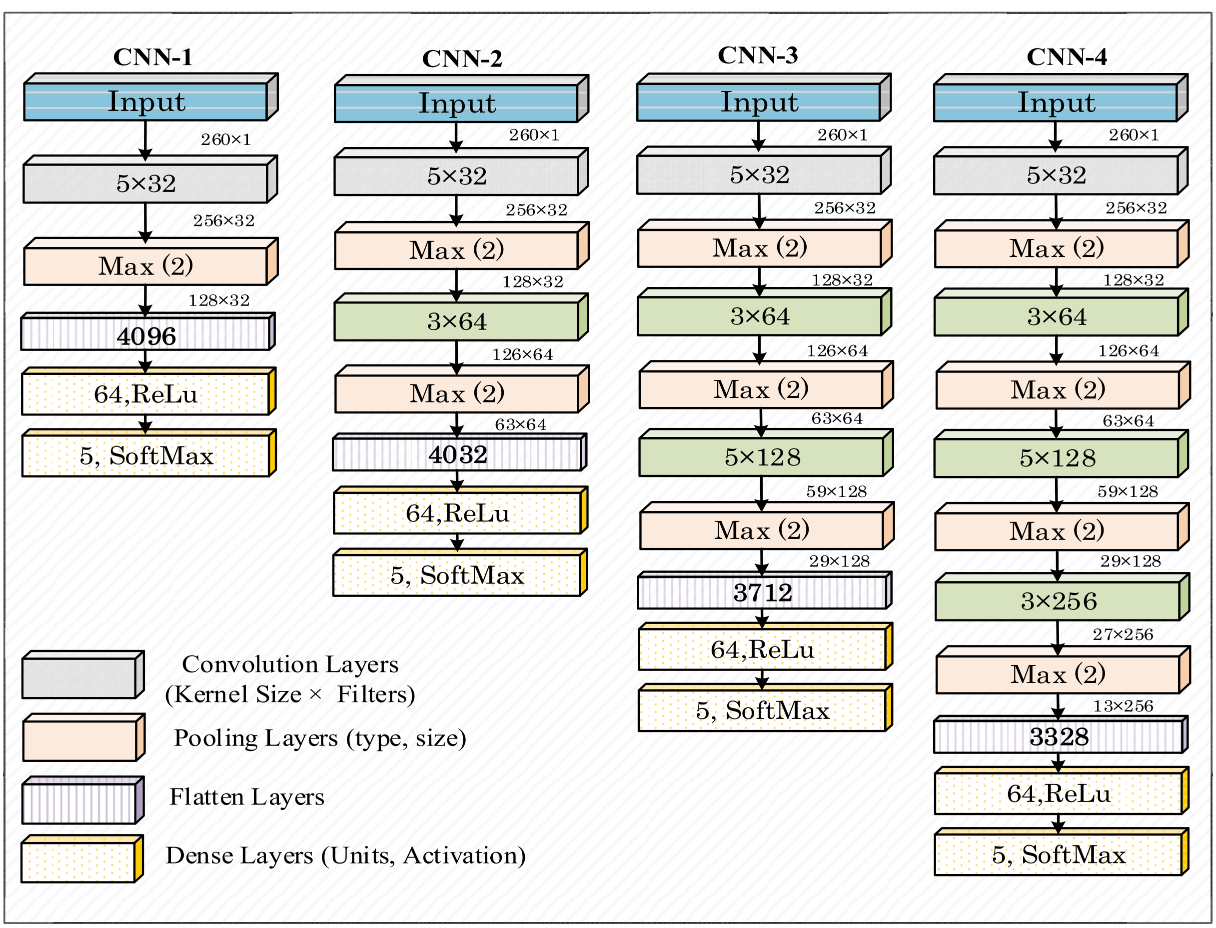
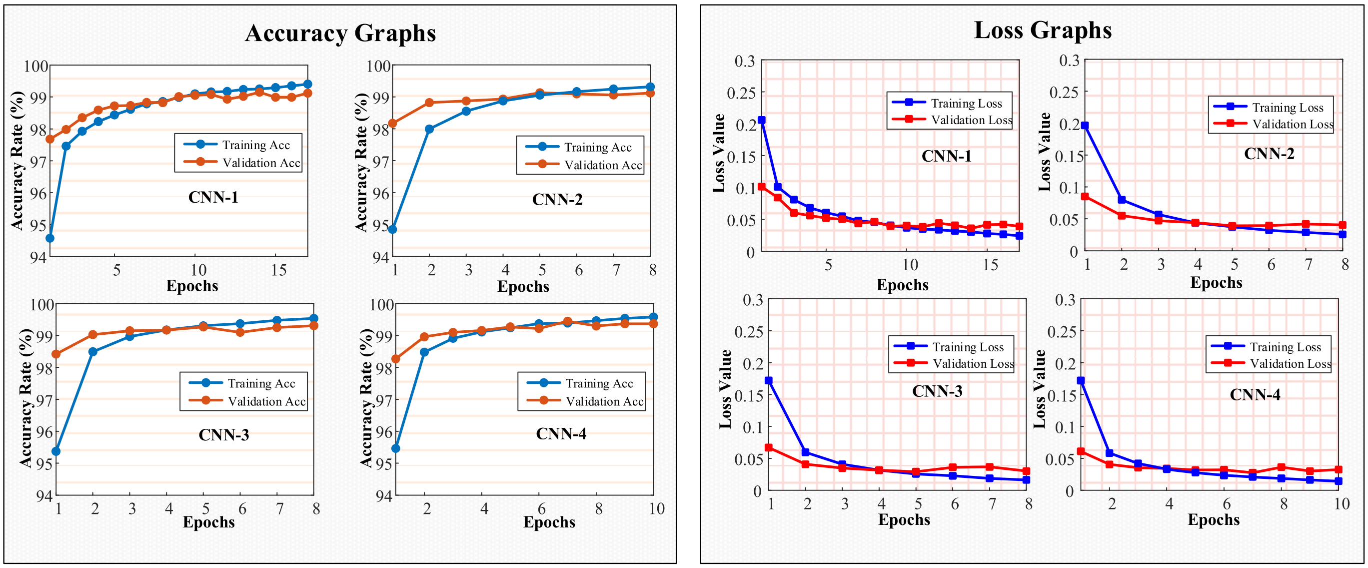


Fig. 4. Detailed layer representations of CNN models constructed for experimental studies.



(a) (b)

Fig. 5. Performance graphs of CNN networks on the arrhythmia dataset during the training phase. a) Accuracy graphs (training and validation), b) Loss values (training and validation).

Table 3
The performance values of deep models on arrhythmia test data.

Model	Classes	Performance Values					
		Acc (%)	Sen (%)	Spe (%)	Prec (%)	F-Score (%)	Overall Acc (%)
CNN-1	APB	99.37	79.83	99.87	94.28	86.46	98.93
	LBBB	99.84	99.10	99.91	98.97	99.04	
	N	99.08	99.61	97.44	99.18	99.39	
	RBBB	99.88	99.70	99.90	98.70	99.20	
	PVC	99.64	97.37	99.82	97.77	97.57	
CNN-4	APB	99.57	86.29	99.91	96.39	91.06	99.16
	LBBB	99.86	98.97	99.94	99.35	99.16	
	N	99.33	99.76	98.01	99.36	99.56	
	RBBB	99.88	99.56	99.91	98.84	99.20	
	PVC	99.63	97.09	99.83	97.90	97.50	

values for APB class were at low levels as compared to other classes. The main reason for this may be that this class has the least amount of data.

Although the widespread use of CNN networks is based upon end-to-end classification, the feature maps obtained from the intermediate layers of these models are rich in information. Therefore, many of the traditional machine learning methods are of interest for the feature extraction stage. By using appropriate feature selection and size reduction methods on attributes obtained from low-to-high level by convolution, useful input sets for shallow or deep classifiers can be obtained [5,48–51]. In addition, some handcrafted features are added to the feature set in order to improve classification performance [49,50]. In a scenario as presented in Fig. 6, the features obtained from a convolution layer are combined by means of the fusion process, and the useful features and classifier inputs selected by various approaches can analyze these features.

Pourbabaee et al. [48] used a CNN network as a feature extractor for the detection of paroxysmal atrial fibrillation. They obtained better results than the end-to-end CNN model by training the features of fully connected layers with the K-NN classifier. Lu et al. [49] classified the random forest classifier by fusing the CNN and PQRST features for arrhythmic signals. He stated that these fused features gave better results than CNN features. Golrizkhatami et al. [50] used arrhythmia detection with three sub-classifier systems using some handcrafted features, along with features from different CNN layers. Fan et al. [43] utilized two 13-layer CNN networks in parallel, and concatenated the final pooling layer features, then performed the learning process of the model in fully connected layers. An illustration of the obtained convolutional feature maps for the CNN-4 model is displayed in Fig. 7.

Since the ECG is 1-dimensional, 1D CNN network models can be used on these data without any conversion. The widespread use of 2D CNN architecture for image problems has led to the emergence of models that work effectively on large datasets. Thus, instead of difficulties encountered in designing new models such as layer and parameter setting, existing models are adapted to existing problems. The conversion of ECG signals to 2-D representative images and the classification by application of known models (such as AlexNet, GoogleNet, and ResNet) have become widespread [1,44,46,49,52,53]. In these conversion processes, frequency spectra of signals and 2-D images of frequency and time functions are generally obtained. An illustration showing process steps is given in Fig. 8.

By converting ECG data into two-dimensional representations, many profound learning techniques applied on images can be used. The short-time Fourier Transform (STFT) approach is frequently used to obtain time-frequency representations. In this method, spectral changes are obtained as a function of time by applying the Fourier transform on all segments in the dedicated window size. These changes are used by converting to image information [52,54]. Apart from STFT, there are other techniques for converting ECG signals into 2D representations. Cao et al. [47] utilized sub-sampling at different scales by decomposing segmented samples to improve CNN performance in AF detection. Zhai et al. [52] employed the 2D-CNN structure with the dual-beat coupling matrix. Rajput et al. [53] incorporated both wavelet and STFT transformations and classified arrhythmia data by converting them to image data. It has also been shown in several studies that arrhythmia inputs converted to 2D representations have advantages over 1D CNN models [1,52].

3.3. Long short-term memory networks

CNNs are powerful in learning representations on input data. However, for sequential signals such as the ECG, it is important to consider long and short term dependencies. Since classical neural networks do not contain memory units, they are insufficient to learn these dependencies. With this in mind, recurrent neural network (RNN)

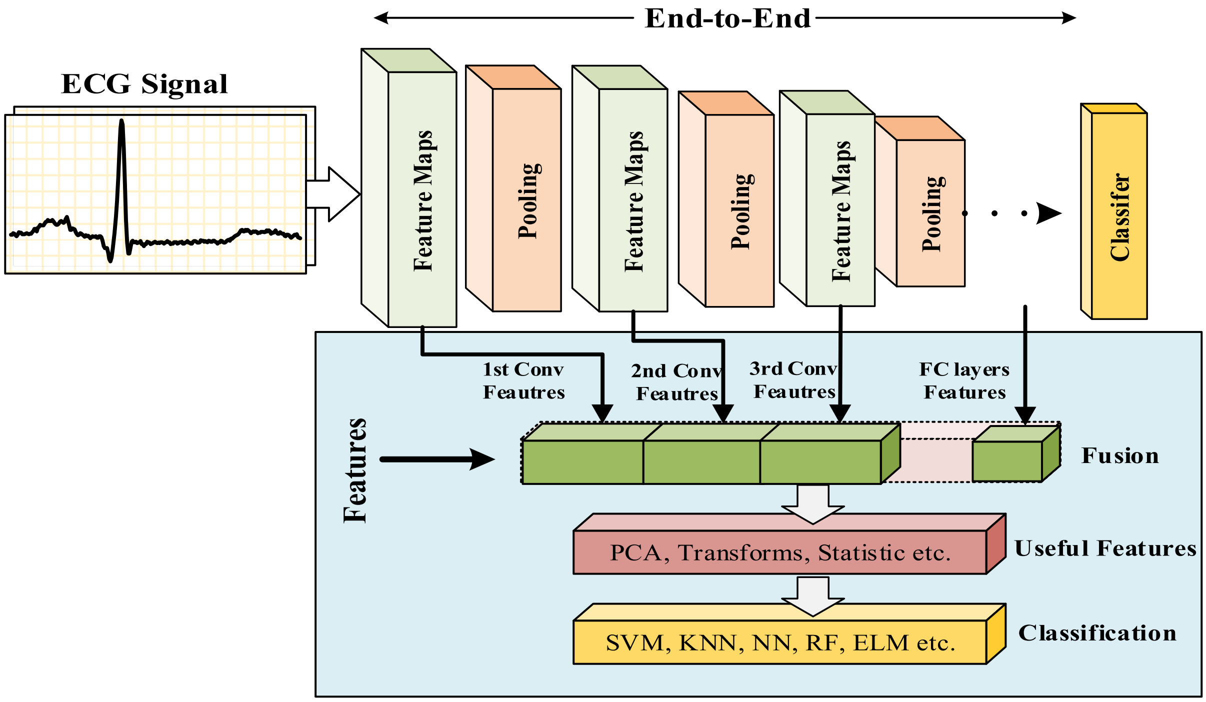


Fig. 6. An illustration of employing CNN models to extract features from ECG input signals.

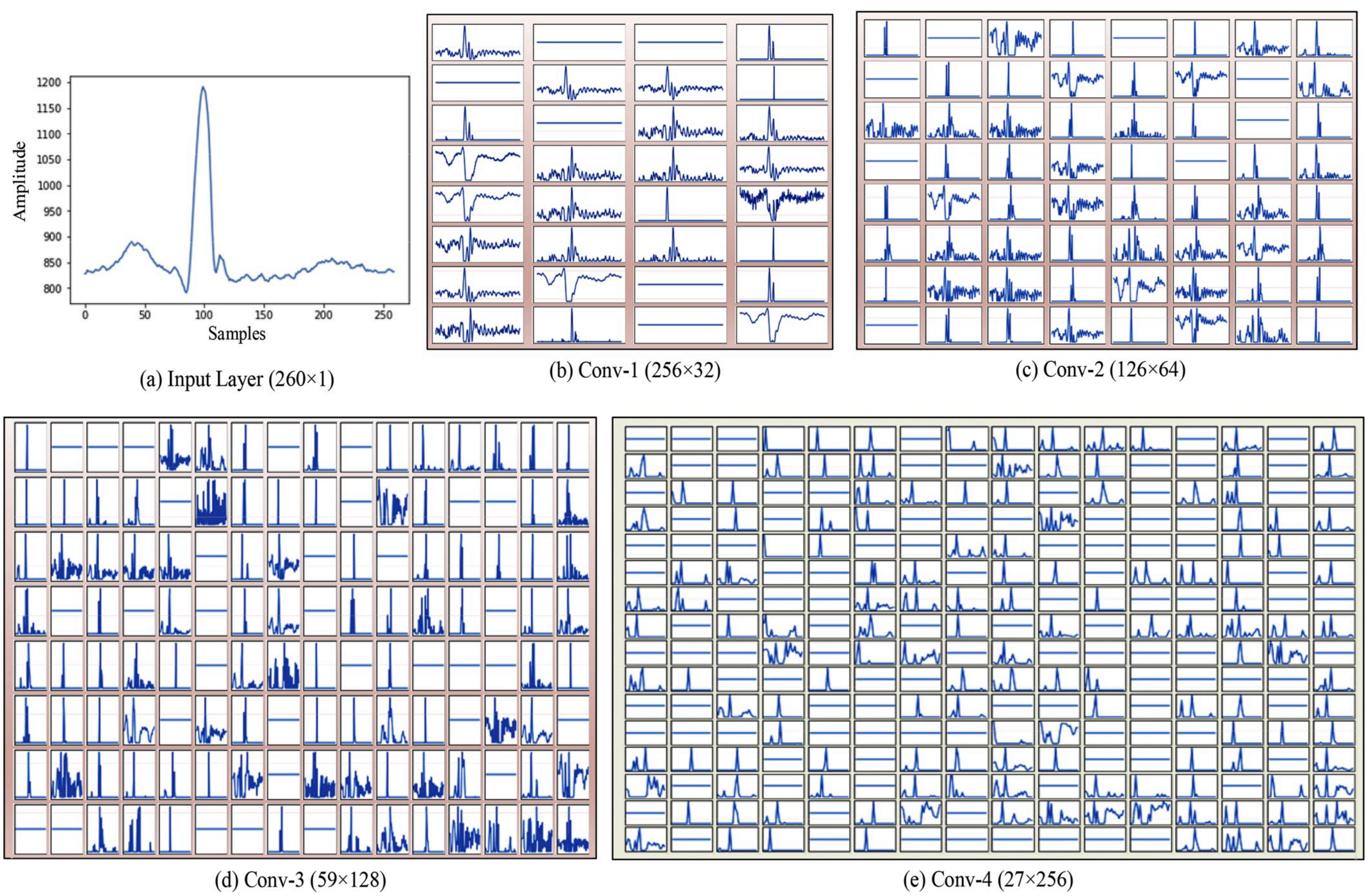


Fig. 7. Feature maps obtained from convolution layers of the CNN-4 model. a) raw input signal, b) feature maps of the first convolution layer, c) features of the second convolution layer, d) features of the third convolution layer, and e) features of the final convolution layer.

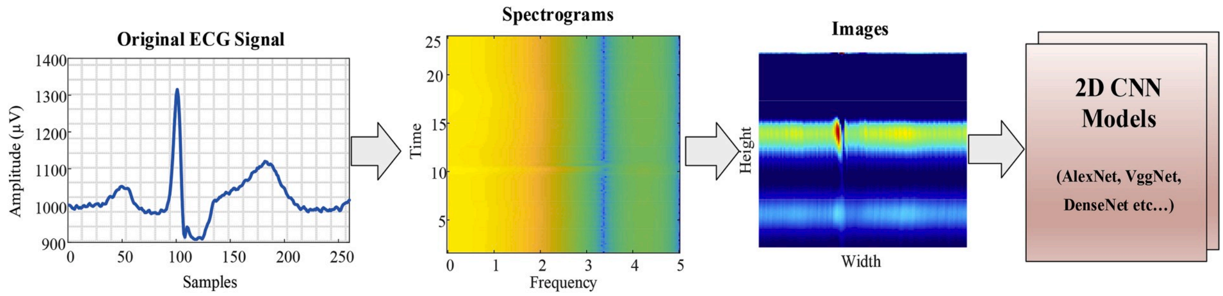


Fig. 8. A block diagram illustrating the acquisition of representations of 1D ECG signals for processing with the 2D CNN model.

architectures have been created by adding an internal memory to feedforward neural networks [55,56]. Although RNNs are successful in short-term memory operations, they have failed to learn long-term dependencies. The most important reason for this is the vanishing gradient problem. LSTM networks have been introduced by Hochreiter and Schmidhuber [57] to solve the problem of vanishing gradient, one of the major difficulties in performing long-term memory. Thanks to the gates (input, forget and output) in the LSTM, the model can be taught using backpropagation through time to avoid gradient problems. In Fig. 9, an illustration of the LSTM structure is given.

As with many sequential problems, LSTM networks have been used effectively in the classification of arrhythmia signals. Faust et al. [58] have proposed an LSTM network that utilizes heart rate (HR) signals as input for the recognition of AF and normal signals. Gao et al. [59] used an LSTM model with focal loss to classify imbalanced arrhythmia data. Yildirim et al. [29] proposed a wavelet transform-based layer to improve the performance of LSTM networks. With this layer, the wavelet coefficients are used as additional features of the signal.

Yildirim et al. [40] employed a coded features-based LSTM approach to classify arrhythmia data. In their study, they first converted the raw signals of 260 samples into 32-dimensional encoded features with an 18-layer convolutional autoencoder. They achieved 99% accuracy by feeding these encoded features into the deep LSTM network. Due to the CAE structure, LSTM networks have significantly reduced the computation time to classify arrhythmia data. However, the CAE structure they used to obtain the encoded features is both complex and time-consuming for coding. Similarly, an interesting study utilizing LSTM networks as a feature extractor is described by Hou et al. [60]. They employed an LSTM-based autoencoder model for arrhythmia recognition, and input the high-level features to the SVM classifier.

In order to compare the performance of LSTM networks in the classification of arrhythmia data, we have prepared different LSTM models that will work on our ECG dataset. These models are commonly used in LSTM based classification problems. Vanilla LSTM (see Fig. 10 (a)) with a single LSTM layer, and stacked LSTM models containing multiple LSTM layers (see Fig. 10 (b)) were created to contain 32 memory units. In addition, the bidirectional LSTM (BLSTM) model (see Fig. 10 (c)), which takes forward and backward sequences as input, was implemented. It is also possible to obtain hidden state outputs for each input time step in the LSTM networks. In the Keras environment, this adjustment is made with the return sequences parameter in LSTM layers. In our study, we used "False" and "True" states to observe the effect of this parameter. The hyper-parameters (learning rate, batch-size, etc.) of these models are the same as in previous models.

From the results (Fig. 11 and Table 4), obtaining the hidden state for each input time step (return_sequences is true) and the stacked use of LSTM layers both increase performance accuracy. However, the number of LSTM layers added to the model doubled the computation time. Furthermore, the performance of the BLSTM model for the dataset we used underperformed the stacked LSTM structure in terms of both time cost and detection accuracy.

In order to reduce the cost of computing and improve the performance of LSTM networks, hybrid techniques have been developed. In particular, CNN and LSTM networks are widely preferred for this purpose [10,61,62]. The main objective of CNN-LSTM networks is to create powerful models of input data by combining both representative learning and sequence learning (see Fig. 12). Oh et al. [10] proposed a CNN-LSTM model for the classification of variable length arrhythmia data. Andersen et al. [61] have proposed a CNN-LSTM model that uses RR interval segments as input for the classification of AF and normal

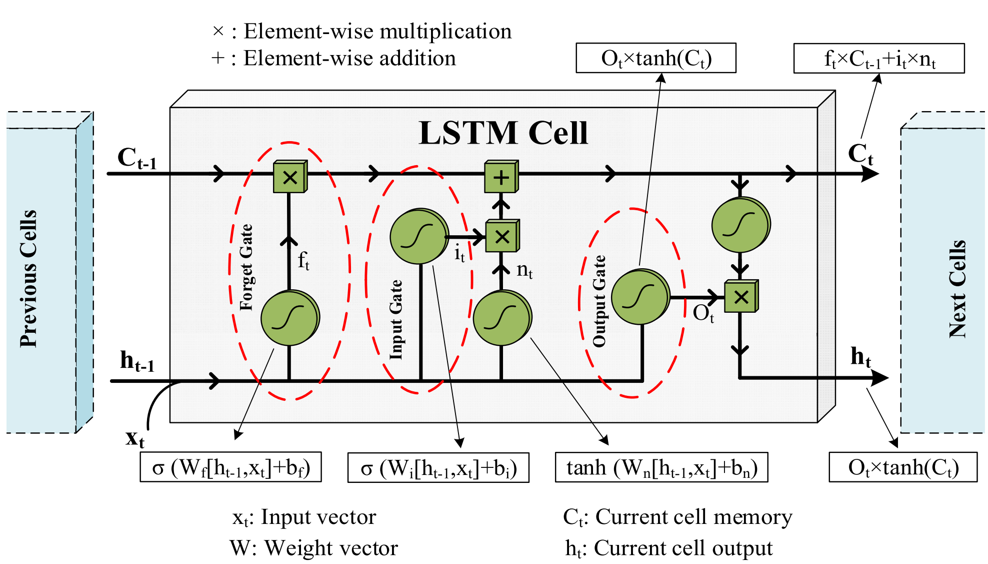


Fig. 9. A block representation of the operating structure of the LSTM cell.

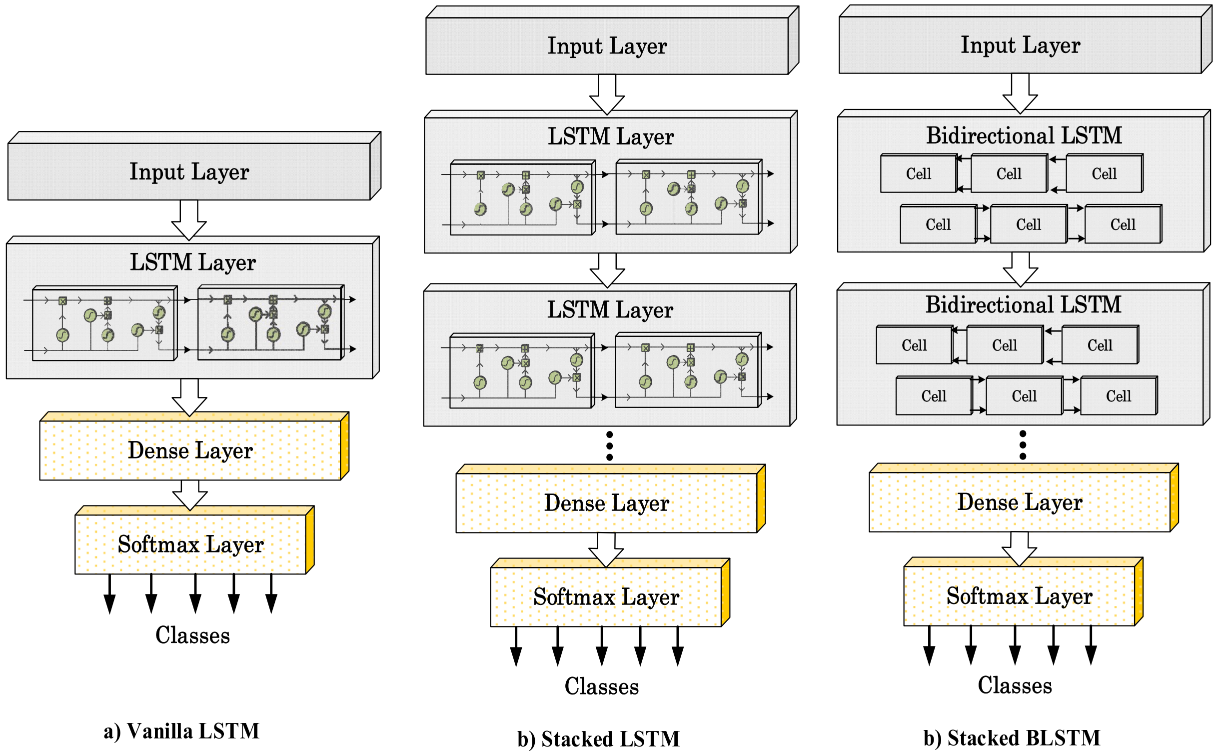


Fig. 10. Three different LSTM models constructed for experimental studies. a) Vanilla LSTM, b) Stacked LSTM and c) Stacked bidirectional LSTM.

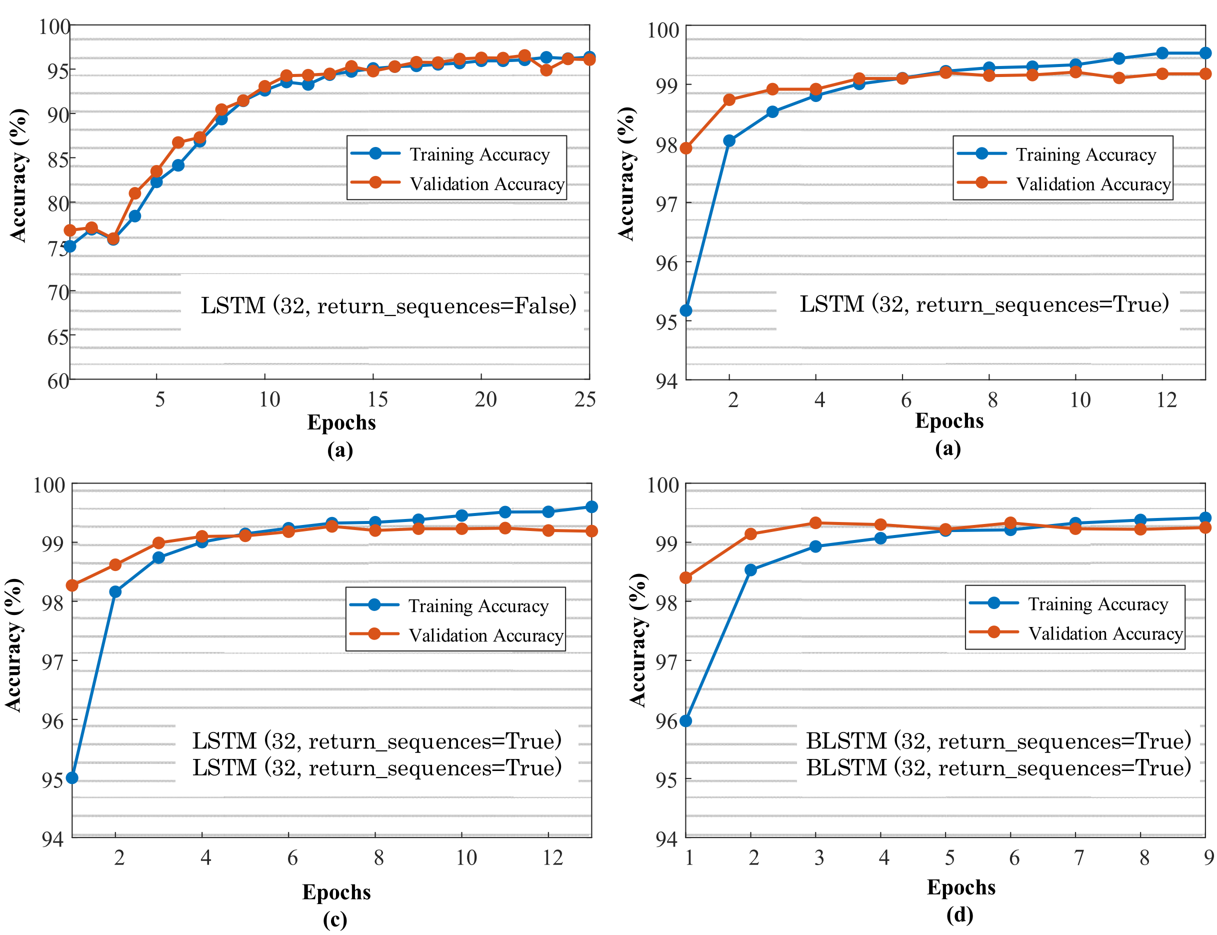


Fig. 11. Accuracy graphs of LSTM models during training. a) Vanilla LSTM, b) Vanilla LSTM with return sequences, c) Stacked LSTM, d) Bidirectional LSTM.

Table 4
The performance values of Vanilla LSTM, Vanilla LSTM with return sequences, Stacked LSTM, and Bidirectional LSTM models.

Models	Training Time per epochs	Overall Sensitivity (%)	Overall Specificity (%)	Overall Precision (%)	Overall F-Score (%)	Overall Accuracy (%)
Vanilla LSTM (return_sequences = True)	245 s	86.92	97.94	90.97	88.41	95.71
Vanilla LSTM (return_sequences = False)	247 s	96.44	99.48	97.35	96.87	98.98
Stacked LSTM	504 s	96.41	99.52	97.91	97.12	99.12
Bidirectional LSTM	1071 s	96.64	99.55	97.32	96.96	99.00

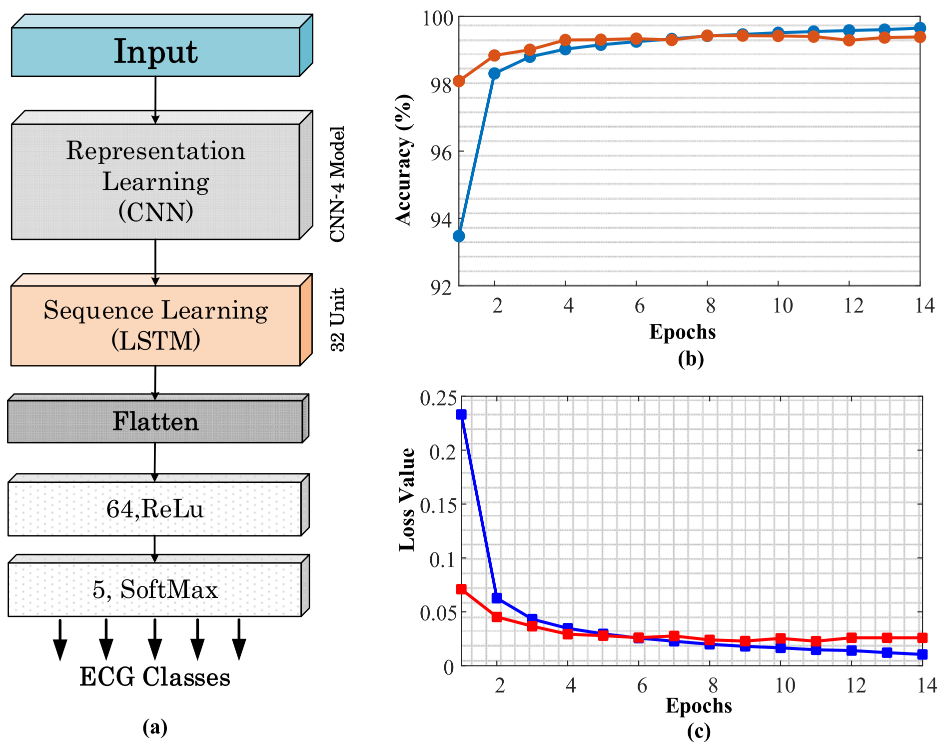


Fig. 12. Representation and sequence learning approach. a) Block representation of the model b) Performance graphs of the model.

ECG signals. They stated that the use of RR segments instead of raw ECG to model inputs reduces the network computational cost. Warric et al. [62] proposed a CNN-LSTM model using the raw ECG input for the AF detection problem.

In addition to LSTM models, different RNN models are employed for arrhythmia recognition. Wang [63] used a CNN-modified Elman neural network (MENN) hybrid for classification of AF signals. Guo et al. [64] utilized CNN and gated recurrent units (GRU) for inter-patient SVEB and VEB arrhythmia detection.

As a final experiment in this study, we designed a CNN-LSTM model for both arrhythmia data and sequence learning. In this model, we have added an LSTM network with 32 memory units to the CNN-4. The structure and performance graphs of this model are presented in Fig. 12.

When the values of the performance criteria for the CNN-LSTM model were examined (see Table 5), a 99.26% overall accuracy yielded a better result than other models. In addition, each epoch time of the CNN-LSTM model requires 294 s to complete, which is similar to that of the vanilla LSTM model. In general, although high performance (sensitivity, precision, and F1-Score) is obtained for all classes except for APB class. The recognition performance is better for classes such as NSR where the number of beats are more. Fewer number of data in other classes and the presence of inappropriate signals in the dataset may lead to misclassifications. Models tend to achieve better performance with more data in each class.

4. Discussion

The researchers conducted their studies on the classes suggested by AAMI standards. Few researchers have frequently used deep learning models on two rhythm classes such as supraventricular ectopic beat

Table 5
The performance values of the CNN-LSTM model on test data.

Classes	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1 Score (%)	Overall Accuracy (%)
APB	99.61	89.51	99.86	94.46	91.92	99.26
LBBB	99.85	99.61	99.86	98.48	99.04	
NSR	99.43	99.69	98.60	99.55	99.62	
RBBB	99.94	99.56	99.96	99.56	99.56	
PVC	99.69	97.37	99.87	98.32	97.84	

(SVEB or S) and ventricular ectopic (VEB or V) [5,6,33–36,38,52,64]. In SVEB and VEB detection process, different records of the same dataset were used to distinguish the events from non-SVEB and non-VEB. Similarly, studies on the detection of atrial fibrillation (AF) is increasing [42,47,58,61–63]. Besides, deep models were frequently used for the classification of beats (Normal, LBBB, RBBB, APC, PVC, Paced, etc.) and arrhythmia classes [1,3,7,10,29,59,66]. Interpatient signals were used to train these models. However, few studies have focused on the patient-specific performance of the models because the beats of the patients showed different characteristics [6,33,36,38].

Table 6
Some state-of-the-art studies using deep learning techniques on ECG signals.

Study	Database	Number of Classes	Total Data	DL Technique	Results
Kiranyaz et al., 2015 [38]	MIT-BIH Arrhythmia Database	5	83,648 beats	1-D CNN	VEB: Acc = 99%, Sen = 93.9%, Spec = 98.9% SVEB: Acc = 97.6%, Sen = 60.3%, Spec = 99.2%
Rahhal et al., 2016 [33]	MIT-BIH Arrhythmia Database INCART SVDB	4	48 records 75 records 78 records (30 min)	DNN-SDAE	VEB: Acc = 99.9%, Sen = 99.3%, Spec = 99.9% SVEB: Acc = 99.9%, Sen = 95.9%, Spec = 100%
Luo et al., 2017 [36]	MIT-BIH Arrhythmia Database	4	49,373 beats	DNN-SDA	Patient-specific scenario: VEB: Acc = 99.1%, Sen = 93.3%, Spec = 99.5% SVEB: Acc = 98.8%, Sen = 71.4%, Spec = 99.8% Inter-patient scenario: VEB: Acc = 95.5%, Sen = 60.4%, Spec = 97.9% SVEB: Acc = 96.2%, Sen = 15.4%, Spec = 99.3%
Rajpurkar et al., 2017 [39]	Zio Patch	14 rhythm	64,121 records	34-layer CNN	PPV = 0.809, Recall = 0.827, F1 = 0.809
Acharya et al., 2017 [8]	MIT-BIH Arrhytmia Database + Synthetic data	5	109,449 beats	9-layer CNN	Set A: Acc = 93.47%, Sen = 96.01%, Spec = 91.64% Set B: Acc = 94.03%, Sen = 96.71%, Spec = 91.54%
Warrick et al., 2017 [62]	PhysioNet Challenge 2017	4	8528 records	LSTM	10-folds CV: F1 = 83.10% Entry: F1 = 84%
Xu et al., 2018 [31]	MIT-BIH Arrhythmia Database	5	50,977 beats	DNN	Exp.1: Acc = 93.1% Exp. 2: Acc = 94.7% Exp. 3: Acc = 99.9% Exp. 4: Acc = 99.7%
Hanbay 2018 [34]	MIT-BIH Arrhytmia Database	4	–	DNN	VEB: Acc = 99.9%, Sen = 99.8%, Spec = 100% SVEB: Acc = 99.9%, Sen = 99.2%, Spec = 100%
Sannino et al., 2018 [32]	MIT-BIH Arrhythmia Database	2	4576 beats	7-layer DNN	Acc = 99.68%, Sen = 99.48%, Spec = 99.83%
Xia et al., 2018 [35]	MIT-BIH Arrhythmia Database Wearable Device Data Base	4	100,700 beats 160,420 beats	DNN	VEB: Acc = 99.8%, Sen = 99.4%, Spec = 99.9% SVEB: Acc = 99.8%, Sen = 98.5%, Spec = 99.9%
Li et al., 2018 [6]	MIT-BIH Arrhythmia Database	5	42,244 beats	GCNN, TDCNN	VEB: Acc = 98.8%, Sen = 95.5%, Spec = 99.1% SVEB: Acc = 98.3%, Sen = 68.7%, Spec = 99.8%
Yildirim 2018 [7]	MIT-BIH Arrhythmia Database	5	7376 beats	7-layer DBLSTM-WS	DULSTM-WS2: Acc = 99.25% DBLSTM-WS3: Acc = 99.39%
Oh et al., 2018 [10]	MIT-BIH Arrhythmia Database	5	16,499 beats with variable length	CNN-LSTM	Acc = 98.1%, Sen = 97.5%, Spec = 98.7%
Yildirim et al., 2018 [40]	MIT-BIH Arrhythmia Database	13 15 17	833 fragments (10s) 976 fragments (10s) 1000 segment (10s)	1D-CNN	Acc = 95.20% Acc = 92.51% Acc = 91.33%
Zhai et al., 2018 [52]	MIT-BIH Arrhythmia Database	5	44 records	2D-CNN	VEB: Acc = 99.1%, Sen = 96.4%, Spec = 99.5% SVEB: Acc = 97.3%, Sen = 85.3%, Spec = 98.0%
Faust et al., 2018 [58]	MIT-BIH Atrial Fibrillation Database	2	100 beat window 99 beats overlap	LSTM	CV: Acc = 98.51%, Sen = 98.32%, Spec = 98.67% Blind fold validation: Acc = 99.77%, Sen = 99.87%, Spec = 99.61%
Hannun et al., 2019 [4]	Zio Monitor	12	91,232 records	DNN	ROC = 0.97, F1 = 0.837
Jiang et al., 2019 [5]	MIT-BIH Arrhythmia Database European ST-T Database MIT-BIH ST Change Database	4	Intra-patient: 230,775 beats Inter-patient: 232,357 beats	DAE & 1D CNN	Intra-patient: Acc = 98.4% Inter-patient: VEB: Acc = 98.8%, Sen = 91.0%, spec = 99.3% SVEB: Acc = 97.3%, Sen = 64.4%, Spec = 98.6%
Huang et al., 2019 [1]	MIT-BIH Arrhythmia Database	5	2520 segments (10 s)	2-D Deep CNN	Acc = 99.0%
Yildirim et al., 2019 [29]	MIT-BIH Arrhythmia Database	5	100,022 beats	16-layer deep CAE & 5-layer LSTM Network	Raw ECG: Acc = 99.23% Coded Features: Acc = 99.11%
Oh et al., 2019 [66]	MITDB	5	94,667 beats	Modified U-Net	Acc = 97.32%
Gao et al., 2019 [59]	MIT-BIH Arrhythmia Database	8	93,371 beats	4-layer LSTM, Focal Loss	Acc = 99.26%, Recall = 99.26%, Spec = 99.14%
Fujita et al., 2019 [45]	MIT-BIH Malignant Ventricular Arrhythmia Database MIT-BIH Atrial Fibrillation Database MIT-BIH Arrhytmia Database	4	25,284 beats	CNN	Normal: Acc = 98.45%, Sen = 99.87%, Spec = 99.27% Arr: Acc = 98.45%, Sen = 99.27%, Spec = 99.87%
Fan et al., 2019 [42]	MIT-BIH Arrhythmia Database CINC Challenge Dataset	3	8249 records	CNN	Balanced set: F1 = 84%, Acc = 85% Imbalanced set: F1 = 85%, Acc = 87%
Guo et al., 2019 [64]	MIT-BIH Arrhythmia Database MIT-BIH Supraventricular Arrhytmia Database	5	289,666 beats	CNN	VEB: Acc = 93.71%, Sen = 91.25%, Spec = 94.77% SVEB: Acc = 93.61%, Sen = 62.70%, Spec = 96.40%
Yao et al., 2020 [3]	1st China Physiological Signal Challenge	8	9831 records (60 s)	ATI-CNN	PPV = 82.6%, Recall = 80.1%, F1 = 81.2%

VEB: V class versus [N, S, and F]; **SVEB:** S class versus [N, V and F]; **Acc:** accuracy; **Sen:** sensitivity; **Spec:** specificity; **CV:** cross-validation; **PPV:** precision; **ROC:** receiver operating characteristic curve, **Arr:** Arrhythmia, **GCNN:** Generic CNN, **TDCNN:** Tuned Dedicated CNN, **ATI-CNN:** Attention-based time-incremental CNN, **DBLSTM-WS:** Deep bidirectional LSTM- Wavelet sequence.

Few state-of-the-art ECG classification studies are given in Table 6. Deep learning approaches used in these studies are: deep neural networks (DNN) [31–36], denoising autoencoders (DAE) [5,33–36], convolutional autoencoders (CAE) [29], CNN [1,5,6,8,38–40,42,45,64,65], LSTM [29,39,58] and CNN-LSTM [3,10,62]. Researchers often design new models with different layer sizes, or they try to improve the input to the model. For example, Guo et al. [64] have made ECG classification by adding dense connections to the standard CNN structure, and thus allowing the use of all former layer outputs. Hannun et al. [4] and Rajpurkar et al. [39] used the CNN structure by adding shortcut connections as it was used in the residual structure. Oh et al. [66] employed a modified U-net model, which is often used for image segmentation studies. In addition, researchers had the opportunity to employ common models used for image processing by simply converting the one-dimensional signal data into two-dimensional data [1,52]. Various preprocessing techniques are used on ECG records before feeding the models. Few common preprocessing operations are removing noise [8, 59], removing baseline-wondering [32,35,36,64], normalization [3,8, 10,59] segmentation [8,10,31,32,36,52,66] and feature extraction [31, 32,35].

Most of the studies have been carried out using the public databases. One of the most important problems encountered in these datasets is the data imbalance problem. Researchers have proposed various approaches to deal with this situation. Acharya et al. [8] preferred to produce synthetic data to overcome imbalance data. These synthetic data were obtained by changing the standard deviation, and Z-score mean calculated from the original ECG signals. Jiang et al. [5] also produced synthetic data, but they oversampled the minority class data. They trained their model with unmodified balanced dataset, and then they fed the model with the unbalanced data to perform classification with fine-tuning process. Similarly, Lu et al. [49] obtained a balanced dataset by increasing the number of minority classes with the random over sampler method. For the imbalanced data problem, Guo et al. [59] reduced the contribution of normal ECG samples during training phase by using the focal loss. Another significant difficulty in public datasets is that these datasets usually contain records of a small number of subjects for few classes. Datasets with more number of subjects contribute better and stable results. For example, Hannun et al. [4] and Rajpurkar et al. [39] used the dataset containing long duration records for large number of patients. They reported an efficient model better than the cardiologist's performance using a well trained model.

In this study, we have analyzed literature reports that use deep learning on arrhythmia ECG data. Some important observations obtained as a result of these examinations are as follows:

- It is an important advantage to classify raw ECG signals with deep learning based systems without using any manual feature extraction. However, some studies have shown that the use of certain temporal features (i.e., RR interval) along with raw signals improves model performance [32,33,58].
- Imbalance of ECG datasets is an important problem. Because there is much data in some classes as compared to other classes, it can give misleading information concerning model performance. Some researchers have focused on this problem and proposed solutions [5,8, 49,59].
- Much recent research in this area has focused on CNN modeling. In our experimental studies, both representation and sequential features of ECG signals improve classification performance. Hence, efficient hybrid models can provide more distinctive features from ECG signals.
- The most important problem evident for CNN modeling is the design of a suitable structure for various datasets. Development of the layer parameters and hyperparameters are an important optimization problem in the formulation of a deep network. For this purpose, models similar to the effective models prepared on big data in image processing should be created for ECG analysis. Therefore, effective results can likely be obtained in this field with a transfer learning approach.
- The methods of converting ECG signals to 2-D images by some researchers for the use of models trained on two-dimensional images for one-dimensional ECG datasets are of interest [1,44,46,49,52,53]. The investigations in this field can be used effectively in arrhythmia classification by utilizing deep models trained on large image datasets.
- Another interesting application in this field would be to employ distinctive models for shallow classification by incorporating deep models as the feature extractors. With this approach, the advantages of shallow classification can be utilized.
- Two different approaches are used, namely signal-wise and subject-wise, during performance evaluation. The main problem of the signal-wise approach is that the performance of the model is high due to the fact that the signals belonging to the same subject can be included in both the training and test sets. Therefore, subject-wise evaluations can give more accurate results about the generalization ability of the model. However, the subject number should be sufficient for this approach.
- In a real life scenario, ECG signals have a noisy structure. In the approaches used in the literature, noise elimination is performed with various pre-processing techniques. Since these pre-processing steps add additional computational cost, more robust models are needed.
- Deep learning models perform well when run on databases containing large amounts of quality information. Consequently, conducting research on recently established large ECG datasets [4,39, 65] may lead to more effective models.

For future studies, research should be expanded on the correct and efficient clinical applications of models created with deep learning techniques. For this purpose, research should be carried out in critical areas such as the integration of models into cloud and mobile systems. In addition, the development of models that work with integrated low power consumption wearable technologies is an important research area. Another important issue that will be needed in the use of these technologies is data security. Research on the protection of personal data stored and transmitted in cloud systems is critical. It is obvious that new approaches that will emerge in parallel with the advances in the field of deep learning can help in the advancement of the field. Furthermore, handcrafted feature extraction and research progress with shallow classifiers are necessary for progress in this area. With the increase of public databases and the increase of the data of specific classes in this direction, it will be a vital source of motivation for deep learning approaches to produce more successful results in the future. Finally, what features are taken into account during the diagnostic process, due to the black-box nature of deep learning methods, is an important question mark. For this reason, research on parameters that the models should consider for input data will play a significant role in developing more reliable methods.

Some of the recent ECG classification studies are given in Table 6 using the deep learning technique. When these studies are examined, it is evident that CNN models are preferred over other methods. Besides the difficulties in the design and parameter adjustment of CNN models, the high computational cost is the most crucial disadvantage of these networks. They also require a big dataset for proper training, which is the another drawback. Furthermore, hybrid models such as CNN-LSTM tend to produce successful results. An important problem in the use of LSTM models is the high resource utilization. This technique requires more time and cost compared to other methods. The most important disadvantage in using the deep learning methods are the requirement of costly hardwares such as graphics processing unit (GPU), layer and parameter optimizations are difficult when developing multi-layer models. Effective techniques such as transfer learning, residual connections and data augmentation will help to overcome these problems over time.

5. Conclusions

This study comprised a comprehensive review and evaluation of deep learning techniques for arrhythmia classification. Peer-reviewed journal articles that utilized deep learning for arrhythmia detection were examined and discussed. An experimental study was presented to provide information concerning techniques that make deep learning effective for arrhythmia detection. In order to examine the performance of the proposed approaches, we constructed deep learning models for categorization of a five-class arrhythmia ECG dataset. We presented results for various deep learning models for arrhythmia detection, and suggested solutions to some important problems in the field.

Declaration of competing interest

The authors declare no conflicts of interest.

References

[1] J. Huang, B. Chen, B. Yao, W. He, ECG arrhythmia classification using STFT-based spectrogram and convolutional neural network, IEEE Access 7 (2019) 92871–92880.
[2] S.N. Yu, K.T. Chou, Integration of independent component analysis and neural networks for ECG beat classification, Expert Syst. Appl. 34 (4) (2008) 2841–2846.
[3] Q. Yao, R. Wang, X. Fan, J. Liu, Y. Li, Multi-class Arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network, Inf. Fusion 53 (2020) 174–182.
[4] A.Y. Hannun, P. Rajpurkar, M. Haghpanahi, G.H. Tison, C. Bourn, M.P. Turakhia, A.Y. Ng, Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, Nat. Med. 25 (1) (2019) 65.
[5] J. Jiang, H. Zhang, D. Pi, C. Dai, A novel multi-module neural network system for imbalanced heartbeats classification, Expert Syst. Appl. X 1 (2019), 100003.
[6] Y. Li, Y. Pang, J. Wang, X. Li, Patient-specific ECG classification by deeper CNN from generic to dedicated, Neurocomputing 314 (2018) 336–346.
[7] Ö. Yildirim, A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification, Comput. Biol. Med. 96 (2018) 189–202.
[8] U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, M. Adam, A. Gertych, R. San Tan, A deep convolutional neural network model to classify heartbeats, Comput. Biol. Med. 89 (2017) 389–396.
[9] S.N. Yu, K.T. Chou, Integration of independent component analysis and neural networks for ECG beat classification, Expert Syst. Appl. 34 (4) (2008) 2841–2846.
[10] S.L. Oh, E.Y. Ng, R. San Tan, U.R. Acharya, Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats, Comput. Biol. Med. 102 (2018) 278–287.
[11] R.J. Martis, U.R. Acharya, L.C. Min, ECG beat classification using PCA, LDA, ICA and discrete wavelet transform, Biomed. Signal Process Contr. 8 (5) (2013) 437–448.
[12] Y.H. Hu, W.J. Tompkins, J.L. Urrusti, V.X. Afonso, Applications of artificial neural networks for ECG signal detection and classification, J. Electrocardiol. 26 (1993) 66–73.
[13] N. Izeboudjen, A. Farah, A new neural network system for arrhythmia's classification, NC 98 (1998) 23–25.
[14] J.S. Wang, W.C. Chiang, Y.L. Hsu, Y.T.C. Yang, ECG arrhythmia classification using a probabilistic neural network with a feature reduction method, Neurocomputing 116 (2013) 38–45.
[15] C.V. Banupriya, S. Karpagavalli, Electrocardiogram beat classification using probabilistic neural network, Int. J. Comput. Appl.(IJCA) 1 (7) (2014) 31–37.
[16] N. Acır, Classification of ECG beats by using a fast least square support vector machines with a dynamic programming feature selection algorithm, Neural Comput. Appl. 14 (4) (2005) 299–309.
[17] M. Moavenian, H. Khorrami, A qualitative comparison of artificial neural networks and support vector machines in ECG arrhythmias classification, Expert Syst. Appl. 37 (4) (2010) 3088–3093.
[18] W.M. Zuo, W.G. Lu, K.Q. Wang, H. Zhang, Diagnosis of cardiac arrhythmia using kernel difference weighted KNN classifier, in: 2008 Computers in Cardiology, IEEE, 2008, September, pp. 253–256.
[19] A. Uyar, F. Gurgen, Arrhythmia classification using serial fusion of support vector machines and logistic regression, in: 2007 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IEEE, 2007, September, pp. 560–565.
[20] M.H. Song, J. Lee, S.P. Cho, K.J. Lee, S.K. Yoo, Support Vector Machine Based Arrhythmia Classification Using Reduced Features, 2005.
[21] M.K. Sarkaleh, A. Shahbahrami, Classification of ECG arrhythmias using discrete wavelet transform and neural networks, Int. J. Comput. Sci. Eng. Appl. 2 (1) (2012) 1.
[22] H. Khorrami, M. Moavenian, A comparative study of DWT, CWT and DCT transformations in ECG arrhythmias classification, Expert Syst. Appl. 37 (8) (2010) 5751–5757.
[23] K.I. Minami, H. Nakajima, T. Toyoshima, Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network, IEEE Trans. Biomed. Eng. 46 (2) (1999) 179–185.
[24] L. Khadra, A.S. Al-Fahoum, S. Binajjaj, A quantitative analysis approach for cardiac arrhythmia classification using higher order spectral techniques, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 52 (11) (2005) 1840–1845.
[25] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[26] C. François, Deep Learning with Python, 2017.
[27] O. Faust, Y. Hagiwara, T.J. Hong, O.S. Lih, U.R. Acharya, Deep learning for healthcare applications based on physiological signals: a review, Comput. Methods Progr. Biomed. 161 (2018) 1–13.
[28] A.L. Goldberger, , et al.PhysioToolkit PhysioBank, PhysioNet, Components of a. New research res. Resource for, Complex Physiologic. Physiol. Signals 101 (23) (2000) e215–e220.
[29] O. Yildirim, U.B. Baloglu, R.S. Tan, E.J. Ciaccio, U.R. Acharya, A new approach for arrhythmia classification using deep coded features and LSTM networks, Comput. Methods Progr. Biomed. 176 (2019) 121–133.
[30] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT press, 2016.
[31] S.S. Xu, M.W. Mak, C.C. Cheung, Towards end-to-end ECG classification with raw signal extraction and deep neural networks, IEEE J. Biomed. Health Inf. 23 (4) (2019), https://doi.org/10.1109/JBHI.2018.2871510.
[32] G. Sannino, G. De Pietro, A deep learning approach for ECG-based heartbeat classification for arrhythmia detection, Future Generat. Comput. Syst. 86 (2018) 446–455.
[33] M.M. Al Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, R.R. Yager, Deep learning approach for active classification of electrocardiogram signals, Inf. Sci. 345 (2016) 340–354.
[34] K. Hanbay, Deep neural network based approach for ECG classification using hybrid differential features and active learning, IET Signal Process. 13 (2) (2018) 165–175.
[35] Y. Xia, H. Zhang, L. Xu, Z. Gao, H. Zhang, H. Liu, S. Li, An automatic cardiac arrhythmia classification system with wearable electrocardiogram, IEEE Access 6 (2018) 16529–16538.
[36] K. Luo, J. Li, Z. Wang, A. Cuschieri, Patient-specific deep architectural model for ECG classification, J. Healthc. Eng. (2017), 2017.
[37] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[38] S. Kiranyaz, T. Ince, M. Gabbouj, Real-time patient-specific ECG classification by 1-D convolutional neural networks, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 63 (3) (2015) 664–675.
[39] P. Rajpurkar, A.Y. Hannun, M. Haghpanahi, C. Bourn, A.Y. Ng, Cardiologist-level Arrhythmia Detection with Convolutional Neural Networks, 2017 arXiv preprint arXiv:1707.01836.
[40] Ö. Yıldırım, P. Pławiak, R.S. Tan, U.R. Acharya, Arrhythmia detection using deep convolutional neural network with long duration ECG signals, Comput. Biol. Med. 102 (2018) 411–420.
[41] U.R. Acharya, H. Fujita, O.S. Lih, Y. Hagiwara, J.H. Tan, M. Adam, Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network, Inf. Sci. 405 (2017) 81–90.
[42] Fan, X., Hu, Z., Wang, R., Yin, L., Li, Y., & Cai, Y. A novel hybrid network of fusing rhythmic and morphological features for atrial fibrillation detection on mobile ECG signals. Neural Comput. Appl., 1-13.
[43] X. Fan, Q. Yao, Y. Cai, F. Miao, F. Sun, Y. Li, Multiscaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ECG recordings, IEEE J. Biomed. Health Inf. 22 (6) (2018) 1744–1753.
[44] J. Rubin, S. Parvaneh, A. Rahman, B. Conroy, S. Babaeizadeh, Densely connected convolutional networks for detection of atrial fibrillation from short single-lead ECG recordings, J. Electrocardiol. 51 (6) (2018) S18–S21.
[45] H. Fujita, D. Cimr, Computer aided detection for fibrillations and flutters using deep convolutional neural network, Inf. Sci. 486 (2019) 231–239.
[46] Y. Xia, N. Wulan, K. Wang, H. Zhang, Detecting atrial fibrillation by deep convolutional neural networks, Comput. Biol. Med. 93 (2018) 84–92.
[47] X.C. Cao, B. Yao, B.Q. Chen, Atrial fibrillation detection using an improved multi-scale decomposition enhanced residual convolutional neural network, IEEE Access 7 (2019) 89152–89161.
[48] B. Pourbabaee, M.J. Roshtkhari, K. Khorasani, Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients, IEEE IEEE rans. Syst. Man Cybern. : Systems 48 (12) (2017) 2095–2104.
[49] W. Lu, H. Hou, J. Chu, Feature fusion for imbalanced ECG data analysis, Biomed. Signal Process Contr. 41 (2018) 152–160.
[50] Z. Golrizkhatami, A. Acan, ECG classification using three-level fusion of different feature descriptors, Expert Syst. Appl. 114 (2018) 54–64.
[51] M. Amrani, M. Hammad, F. Jiang, K. Wang, A. Amrani, Very deep feature extraction and fusion for arrhythmias detection, Neural Comput. Appl. 30 (7) (2018) 2047–2057.
[52] X. Zhai, C. Tin, Automated ECG classification using dual heartbeat coupling based on convolutional neural network, IEEE Access 6 (2018) 27465–27472.
[53] K.S. Rajput, S. Wibowo, C. Hao, M. Majmudar, On Arrhythmia Detection by Deep Learning and Multidimensional Representation, 2019 arXiv preprint arXiv: 1904.00138.
[54] O. Yildirim, M. Talo, B. Ay, U.B. Baloglu, G. Aydin, U.R. Acharya, Automated detection of diabetic subject using pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals, Comput. Biol. Med. 113 (2019) 103387.
[55] J.L. Elman, Finding structure in time, Cognit. Sci. 14 (1990) 179–211.

[56] M.I. Jordan, Tech. Rep. No. 8604, in: Serial Order: A Parallel Distributed Processing Approach, University of California, Institute for Cognitive Science, San Diego, 1986, 1986.
[57] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997) 1735–1780.
[58] O. Faust, A. Shenfield, M. Kareem, T.R. San, H. Fujita, U.R. Acharya, Automated detection of atrial fibrillation using long short-term memory network with RR interval signals, Comput. Biol. Med. 102 (2018) 327–335.
[59] J. Gao, H. Zhang, P. Lu, Z. Wang, An effective LSTM recurrent network to detect arrhythmia on imbalanced ECG dataset, J. Healthc Eng. (2019) 1–10, 2019.
[60] B. Hou, J. Yang, P. Wang, R. Yan, LSTM based auto-encoder model for ECG arrhythmias classification, IEEE Trans. Instrum. Meas. (2019), https://doi.org/10.1109/TIM.2019.2910342.
[61] R.S. Andersen, A. Peimankar, S. Puthusserypady, A deep learning approach for real-time detection of atrial fibrillation, Expert Syst. Appl. 115 (2019) 465–473.
[62] P. Warrick, M.N. Homsi, Cardiac arrhythmia detection from ECG combining convolutional and long short-term memory networks, in: 2017 Computing in Cardiology (CinC), 2017, September, pp. 1–4.
[63] J. Wang, A deep learning approach for atrial fibrillation signals classification based on convolutional and modified Elman neural network, Future Generat. Comput. Syst. 102 (2020) 670–679.
[64] L. Guo, G. Sim, B. Matuszewski, Inter-patient ECG classification with convolutional and recurrent neural networks, Biocybern. Biomed. Eng. 39 (3) (2019) 868–879.
[65] S. Tan, G. Androz, A. Chamseddine, P. Fecteau, A. Courville, Y. Bengio, J.P. Cohen, Icentia11K: an Unsupervised Representation Learning Dataset for Arrhythmia Subtype Discovery, 2019 arXiv preprint arXiv:1910.09570.
[66] S.L. Oh, E.Y. Ng, R. San Tan, U.R. Acharya, Automated beat-wise arrhythmia diagnosis using modified U-net on extended electrocardiographic recordings with heterogeneous arrhythmia types, Comput. Biol. Med. 105 (2019) 92–101.
[67] M. Sharma, R.S. Tan, U.R. Acharya, Automated heartbeat classification and detection of arrhythmia using optimal orthogonal wavelet filters, Inf. Med. Unlocked 16 (2019), 100221.
[68] T. Tuncer, S. Dogan, P. Pławiak, U.R. Acharya, Automated arrhythmia detection using novel hexadecimal local pattern and multilevel wavelet transform with ECG signals, Knowl. Base Syst. 186 (2019), 104923.
[69] R.J. Martis, U.R. Acharya, C.M. Lim, K.M. Mandana, A.K. Ray, C. Chakraborty, Application of higher order cumulant features for cardiac health diagnosis using ECG signals, Int. J. Neural Syst. 23 (2013) (04), 1350014.
[70] P. Pławiak, U.R. Acharya, Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals, Neural Comput. Appl. (2019) 1–25.