Classification of Atrial Fibrillation and Congestive Heart Failure using Convolutional Neural Network with Electrocardiogram

Yunendah Nur Fu'adah
Kumoh National Institute of Technology
Ki Moo Lim (kmlim@kumoh.ac.kr)
Kumoh National Institute of Technology

Research Article

Keywords: CNN algorithm, provided precision values, recall, f1-score, AF, CHF, NSR

Posted Date: May 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-521179/v1

Abstract

Delayed diagnosis of atrial fibrillation (AF) and congestive heart failure (CHF) can lead to death. Early diagnosis of these cardiac conditions is possible by manually analyzing electrocardiogram (ECG) signals. However, manual diagnosis is complex, owing to the various characteristics of ECG signals. Several studies have reported promising results using the automatic classification of ECG signals. The performance accuracy needs to be improved considering that an accurate classification system of AF and CHF has the potential to save a patient's life. An optimal ECG signal classification system for AF and CHF has been proposed in this study using a one-dimensional convolutional neural network (1-D CNN) to improve the performance. A total of 150 datasets of ECG signals were modeled using the1-D CNN. The proposed 1-D CNN algorithm, provided precision values, recall, f1-score, accuracy of 100%, and successfully classified raw data of ECG signals into three conditions, which are normal sinus rhythm (NSR), AF, and CHF. The results showed that the proposed method outperformed the previous methods. This approach can be considered as an adjunct for medical personnel to diagnose AF, CHF, and NSR.

Introduction

Atrial fibrillation (AF) and congestive heart failure (CHF) are the most common cardiac diseases and are potentially life-threatening [1]. In 2017, AF affected 37.574 million people worldwide, and its frequency has risen by 33% in the last 20 years [2]. The number of AF cases is expected to increase by more than 60% by 2050 [2]. Similarly, the global prevalence of heart failure in the year 2017 was 64.34 million people [3]. The lifetime risk of developing CHF is over 20% after the age of 40 years [3]. Undiagnosed AF and CHF can endanger the patient's life [4]; therefore, prompt diagnosis is important.

Electrocardiogram (ECG) is a non-invasive method that records the electrical activity of the heart which is commonly measured and analyzed by researchers [5]. ECG signals can be used to detect abnormalities in the heart, such as cardiac arrhythmias [6] or heart failure. Therefore, ECG is important to represent objectively the occurrence of AF and CHF. Several studies have proposed algorithms to classify various cardiac diseases including AF and CHF based on ECG signals.

Rizal et al. analyzed ECG signals based on parameters of the Hjorth descriptor method such as activity, mobility, and complexity [7,8]. A study conducted in 2015 using various classifier algorithms such as k-mean clustering, k-nearest neighbor, and multi-layer perceptron reported an accuracy of 88.67%, 99.3%, and 99.3% respectively to classify AF, CHF, and NSR. The same researchers reported an accuracy of 94% in a study conducted in 2017 using the higher order complexity and k-nearest neighbor classifier [8].

Furthermore, Thaweesak et al. used the primary datasets of AF, CHF, and NSR which consisted of 90 recordings. Their findings indicated that the Hjorth descriptor is capable of class separation among cardiac arrhythmia patient groups and reported accuracy rates of 84.89%, 88.22%, and 76 % using the least-squares, maximum likelihood, and support vector machine respectively [9].

The most challenging aspect of ECG analysis is feature extraction. The aforementioned studies [7–9], extracted the features of ECG signals using the Hjorth descriptor method. However, owing to noise interference, some features did not perfectly represent the characteristics of ECG signals. Therefore, the classification systems used in the previous studies remain prone to false detection that decreases the accuracy performance.

Recently, convolutional neural networks (CNNs) have been identified as potential approaches for ECG classification. CNNs automatically extract features from the raw ECG signal data. Several studies have applied CNNs to classify cardiac diseases and have shown promising results. Ping et al. reported the highest f1-scores value of 89.55% using a CNN and long short-term memory to classify AF and normal conditions [10]. Sidrah et al. reported a classification accuracy of 86.5% using a CNN and long short- term memory to classify AF and normal conditions [11]. Similarly, Georgios et al. reported sensitivity of 97.87% and specificity of 99.29% using a CNN and long short-term memory to classify AF and normal conditions [12]. Moreover, Nurmaini et al. reported a classification accuracy of 99.17% using a CNN to classify three conditions, including normal, AF, and non-AF conditions [13].

As well as several studies developed automated systems for detecting AF, Wang et al. reported a classification accuracy of 99.85% using a CNN and long short-term memory to classify CHF and normal conditions [14]. Ning et al. reported a classification accuracy of 99.3% using a recursive neural network to classify CHF and normal conditions [15]. Meanwhile, Porumb et al. reported accuracy of 100 % using a CNN to classify CHF and normal conditions [16].

The aforementioned studies [10–16] developed the automatic system detection for AF or CHF conditions separately. Furthermore, Padmavati et al. proposed a classification system to identify four classes of cardiac disease including atrial fibrillation (AF), myocardial infarction (MI), congestive heart failure (CHF), and normal conditions [17]. The study reported a classification accuracy of 80.1% to classify CHF and normal, 85.9% of accuracy to classify MI and CHF, and 65.4% of accuracy to classify MI and normal conditions. However, the accuracy performance of their proposed model decreased up to 63.5% and 31.2% while extended to classify three classes (CHF, MI, and normal) and four classes (AF, CHF, MI, and normal) respectively.

Performance results of certain previous studies on the classification of ECG signals revealed that conventional machine learning methods were sensitive to noise; therefore, data cleaning was required. In contrast to the conventional approaches that required separate dataset preprocessing, feature extraction, and classification processes, CNNs can directly extract features from raw input. As a result, in the presence of sufficient training samples, the features extracted by a CNN model would be more detailed than those extracted manually.

According to the advantages of CNNs, some limitations which previously existed in conventional machine learning can be overcome with CNNs. Therefore, a one-dimensional (1-D) CNN was proposed to design an optimal ECG classification system that can improve the performance accuracy of the previous methods. The aim of this study was to investigate the performance of the 1-D CNN to classify raw ECG signals of AF, CHF, and NSR.

Results

A total of 38 test datasets consisting 12 datasets of AF, 14 datasets of CHF and 12 datasets of NSR were used to evaluate the system performance. After conducting several simulations, the Adam optimizer with a learning rate of 0.001 was selected as the optimal hyperparameter that provided the highest accuracy and a minimum loss value compared with other optimization algorithms (Table 1).

Table 1
The comparison of optimization algorithms performance.

Optimizer	Learning Rate	Accuracy (%)	Loss
Adam	0.001	100	0.0127
Nadam	0.001	98	0.4278
RMSprop	0.001	98	0.0946
SGD	0.001	42	0.3947

According to the results, optimization algorithms selections were the influential terms that affected the performance of the system. Adam optimizer algorithm obtained 100% accuracy with loss value of 0.0127. The great accuracy with minimum loss indicated low errors on a few data. Nadam and RMSprop optimizer algorithms obtained accuracy 98 % with loss value of 0.4278 and 0.0946 respectively. Based on accuracy performance, Nadam, and RMSprop were potential optimizer algorithms for the proposed model. However, the greater value of loss obtained by Nadam optimizer indicated there were huge errors on a few data. Meanwhile, SGD optimizer obtained 42% accuracy with loss value of 0.3947. The low accuracy with huge loss indicated huge errors on a lot of data.

As shown in Fig. 1, 100% accuracy was achieved after 500 epochs with a batch size of 32. According to the model accuracy performance, there was no over-fitting which indicated the proposed model ability to generalize well the test datasets. Table 2 shows the confusion matrix of the 1-D CNN for detailed information regarding performance of the algorithm. Based on results of the confusion matrix, the proposed model successfully classified AF, CHF, and NSR according to their class and provided accuracy, precision, recall, and f1-score values of 100%.

Table 2
The confusion matrix of the one-dimensional convolutional neural network.

Condition	AF	CHF	NSR	Precision (%)	Recall (%)	F1-Score (%)	Acc (%)
AF	12	0	0	100	100	100	100
CHF	0	14	0	100	100	100	
NSR	0	0	12	100	100	100	

AF, atrial fibrillation; CHF, congestive heart failure; NSR, normal sinus rhythm; 1-D CNN, one-dimensional convolutional neural network.

Discussion

Researchers have considered two approaches of ECG classification, namely the conventional classification and the 1-D CNN (Table 3). The conventional classification requires separate preprocessing, feature extraction and classification processes [7–9]. However, the 1-D CNN integrates the feature extraction and classification processes and is more efficient than other machine learning algorithms.

Table 3
Performance comparison with previous studies.

References	No of Classes	Feature Extraction	Classifier	Accuracy (%)
Rizal et al. (2015)	3 classes (AF, CHF, NSR)	Hjorth descriptor	K Mean Clustering, K-NN, and MLP.	88.67, 99.3, and 99.3
Rizal et al. (2017)	3 classes (AF, CHF, NSR)	Higher order complexity	K-NN	94
Thaweesak et al. (2018)	3 classes (AF, CHF, NSR)	Hjorth descriptor	LS, ML, SVM	84.89, 88.22, and 76.94
Nurmaini et al. (2020)	3 classes (AF, non- AF, normal)		CNN	99.17
Porumb et al. (2020)	2 classes (CHF, normal)		CNN	100
Padmavathi et al. (2020)	2 classes (CHF, normal), 2 classes (CHF, MI), 2 classes (MI, normal), 3 classes (CHF, MI, normal), and 4 classes (AF, CHF, MI, normal)	EMD	CNN	80.1, 85.9, 65.4, 63.5, and 31.2
Our Method	3 classes (AF, CHF, NSR)		CNN	100

AF, atrial fibrillation; CHF, congestive heart failure; NSR, normal sinus rhythm; MI, myocardial infarction; CNN, convolutional neural network; K-NN, k-nearest neighbor; MLP, multi-layer perceptron; LS, least-squares; ML, maximum likelihood; SVM, support vector machine; EMD, empirical mode decomposition.

In this study, we proposed a new configuration of the 1-D CNN to classify AF, CHF, and NSR conditions. While AF, CHF, and NSR classification using machine learning in the previous studies [7–9] that relied on the extraction of Hjorth descriptor features to make classifiers capable of classifying AF, CHF, and NSR conditions, we significantly advanced the method by using raw ECG signals as input of 1-D CNN. As a result, we improved the classification accuracy performance from the aforementioned studies that used the same datasets [7,8] (Table 3). We claimed this advantage due to the ability of 1-D CNN to extract and learn the pattern of ECG signal rather than relying on specific features that might not completely represent the characteristics of ECG signal.

While determining the configuration of the 1-D CNN model, the number of filters and depth of the model must be considered. The feature maps as well as the model complexity are influenced by these parameters. If the model is too simple, it will not be able to extract the unique features. On the other hand, if the model is too deep, it will increase the model complexity as well as slow the training process.

Therefore, the proposed model in this research proposed the configuration of 1-D CNN model as shown in Fig. 1. Furthermore, several hyper parameters including optimization algorithms and learning rate were carefully selected after conducting extensive simulations to avoid over-fitting and increase robustness and generalization capability.

Several studies that used CNNs showed promising results in classifying ECG signals as shown in Table 3. The previous studies developed the system to detect the occurrence of AF or CHF separately [10–16]. Meanwhile, Padmavati et al. proposed the system to classify several conditions of ECG signal. Their proposed model performed quite well to classify two conditions. However, the accuracy performance significantly decreased to classify three conditions and four conditions [17]. It showed that the performance was still affected by false detection that decreased the accuracy performance.

The main goal of this study was to verify the ability of the 1-D CNN to learn the characteristics of each condition (AF, CHF, and NSR) based on raw ECG signal data without requiring separate data cleaning, preprocessing, feature extraction, and classification processes. Furthermore, the classification performance of the 1-D CNN was not affected by false detection. Therefore, the values of accuracy, recall, precision and f1-score were equal to 100% in classifying the 38 test datasets of AF, CHF, and NSR. In addition to classification performance, a large dataset is required for training to be implemented for clinical use.

Methods

Through this study, we have proposed an optimal ECG signal classification system using a 1-D CNN that directly processes the raw ECG signal data and classifies them into three conditions, namely, AF, CHF, and NSR. The configuration of the proposed 1-D CNN model is shown in Fig.2.

Dataset

ECG signal data were collected from the MIT-BIH database that can be accessed from PhysioNet [18], the datasets have previously been used in a number of studies [7,8]. The data consisted of three cardiac classes, namely, AF, CHF, and NSR. Each class had 50 datasets with a sampling rate of 250 Hz. A total of 150 ECG signal datasets were obtained, which were divided as training (n=112) and testing (n=38) data. Each ECG signal consisted of 245 samples as inputs for the convolutional layer in the CNN.

Feature Extraction using 1-D CNN

The feature extraction layer of the 1-D CNN used in this study consisted of five convolutional layers with a kernel size of five for each layer. Convolutional layers one through five had a number of filters (8, 16, 32, 64, and 128 respectively). Rectified linear unit activation functions were applied to each convolution layer. Similarly, following the convolutional layer, we applied max pooling to each layer and dropout 0.5 at the last feature extraction layer to avoid network complexity as well as over-fitting. Subsequently, the feature maps were extracted from the convolutional and pooling layers as inputs for the classification layers.

Classification using 1-D CNN

The classification layer of the 1-D CNN is a fully connected layer that is responsible for classifying the data. There is a flattening process prior to creating the fully connected layers, which included 256, 128, 64, and 32 dense layers respectively (Fig. 2). Finally, the softmax activation function was applied to classify the signals into three conditions, namely, AF, CHF, and NSR.

To optimize the 1-D CNN algorithm, the hyperparameter performance was investigated including optimizer algorithms (Adam, Nadam, SGD, and RMSprop) and the optimal learning rate (0.1 to 0.0001). The optimizer algorithms, including the learning rate value, were applied while training the network. The optimizer algorithm minimizes the error related to accuracy performance.

System Performance

A confusion matrix was used to obtain the accuracy, precision, recall, and f1-score when evaluating the system performance. The following equations were used to calculate parameters to measure the effectiveness of the system in diagnosing AF, CHF, and NSR conditions.

$$Accuracy = (TP + TN)/(TP + FP + TN + FN) \tag{1}$$

$$Precision = TP/(TP + FP) \tag{2}$$

$$Recall = TP/(TP + FN) \tag{3}$$

$$F1\,score = 2 \cdot \frac{Recall \cdot Precision}{Recall + Precision} \tag{4}$$

In Equation (1), (2), and (3), a true positive (TP) is a result in which the model predicts the positive class correctly. A true negative (TN), on the other hand, is a result in which the model correctly predicts the negative class. A false positive (FP) is a result in which data is negative but incorrectly classified as positive. Meanwhile, a false negative (FN) is a result in which data is positive but incorrectly classified as negative [19].

Declarations

Acknowledgments

This research was partially supported by NRF (National Research Foundation) under basic engineering research project (2016R1D1A1B0101440) and the EDISON (NRF-2011-0020576) Programs, and the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program (IITP-2021-2020-0-01612) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation).

Author contribution statement

This manuscript is the intellectual product of the entire team. YNF wrote the convolutional neural network source code and manuscript, performed data analysis, and interpreted the results. KML designed the study, reviewed, and revised the entire manuscript based on the results. All authors have read and approved the final manuscript.

Additional information

Data availability

Publicly available datasets were analyzed in this study. Datasets can be accessed here: https://www.physionet.org/.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary Information

References

1. M. W. Mak & C. C. Cheung. Towards End-to-End ECG Classification with Raw Signal Extraction and Deep Neural Networks. *IEEE J. Biomed. Heal. Informatics***23**, 1574–1584 (2019).
2. Lippi, G., Sanchis-Gomar, F. & Cervellin, G. Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. *Int J Stroke* (2021).
3. Lippi, G. & Sanchis-Gomar, F. Global epidemiology and future trends of heart failure. *AME Med. J.* (2020).
4. Erdenebayar, U. Automatic Prediction of Atrial Fibrillation Based on Convolutional Neural Network Using a Short-term Normal Electrocardiogram Signal. *J. Korean Med. Sci.* (2019).
5. Taye, G. T., Hwang, H.-J. & Lim, K. M. Application of convolutional neural network for predicting the occurrence of ventricular tachyarrhythmia using heart rate variability features. *Sci. Rep.* (2020).
6. Taye, G. T., Shim, E. B., Hwang, H.-J. & Lim, K. M. Machine Learning Approach to Predict Ventricular Fibrillation Based on QRS Complex Shape. *Front. Physiol.* (2019).
7. Rizal, A. & Hadiyoso, S. ECG Signal Classification using Hjorth Descriptor. *Int. Conf. Autom. Cognitive Sci. Opt. Micro Electro-Mechanical Syst. Inf. Technol.* (2015).
8. Hadiyoso, S. & Rizal, A. Electrocardiogram Signal Classification Using Higher-Order Complexity of Hjorth Descriptor. *Adv. Sci. Lett.***23**, 3972–3974 (2017).

9. Thaweesak, Y. & Pawita, T. Cardiac Arrhythmia Classification Using Hjroth Descriptors. AISC (2018).
10. Ping, Y. & Chen, C. Automatic Detection of Atrial Fibrillation Based on CNN-LSTM and Shortcut Connections. *Healthcare* (2020).
11. Sidrah, L. & Dashtipour, K. Detection of Atrial Fibrillation Using a Machine Learning Approach. *Information* (2020).
12. Georgios, P. & Haris, K. Automated Atrial Fibrillation Detection using a Hybrid CNN-LSTM Network on Imbalanced ECG Datasets. *Biomed. Signal Process. Control* (2020).
13. Nurmaini, S. & Tondas, A. E. Robust detection of atrial fibrillation from short-term electrocardiogram using convolutional neural networks. *Futur. Gener. Comput. Syst.* (2020).
14. Wang, L. & Zhou, W. Deep Ensemble Detection of Congestive Heart Failure Using Short-Term RR Intervals. *IEEE Access* (2019).
15. Ning, W. & Li, S. Automatic Detection of Congestive Heart Failure Based on a Hybrid Deep Learning Algorithm in the Internet of Medical Things. *IEEE,Internet Things J.* (2020).
16. Porumb, M. & Iadanza, E. A convolutional neural network approach to detect congestive heartfailure. *Biomed. Signal Process. Control* (2019).
17. Padhmavathi, C. & V Veenadevi, S. Heart Disease Recognition from ECG Signal Using Deep Learning. *Int. J. Adv. Sci. Technol.* **29**, 2303–2316 (2020).
18. Physionet.org. ECG Database.
19. Novakovic, J. D., Veljovic, A., C, S. S. I. ˇ, C, Z. P. ˇ & Tomovic, M. Evaluation of Classification Models in Machine Learning. *Theory Appl. Math. Comput. Sci.* **1**, 39 – 46 (2017).

Figures

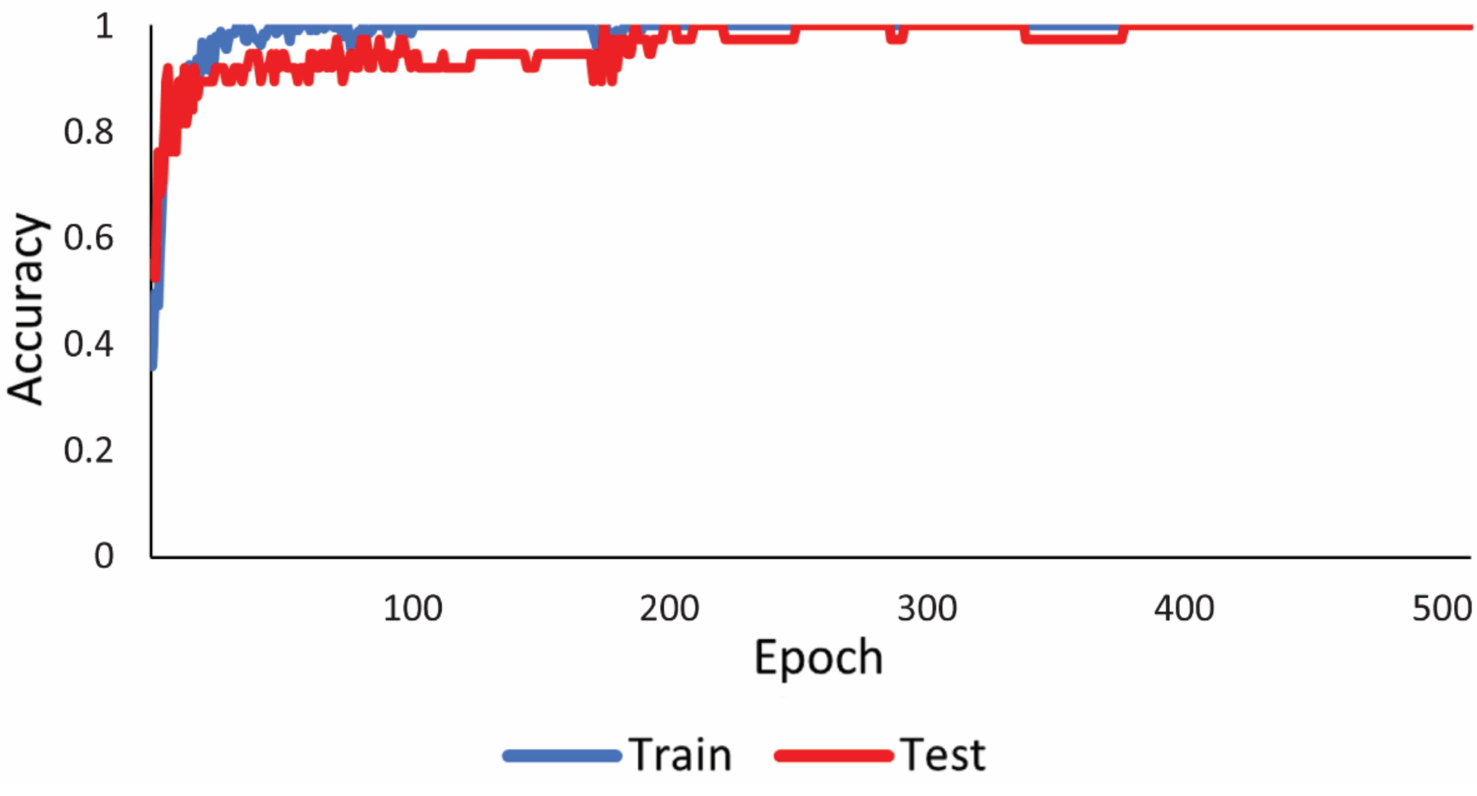

Figure 1

Training and testing accuracy of the proposed model.

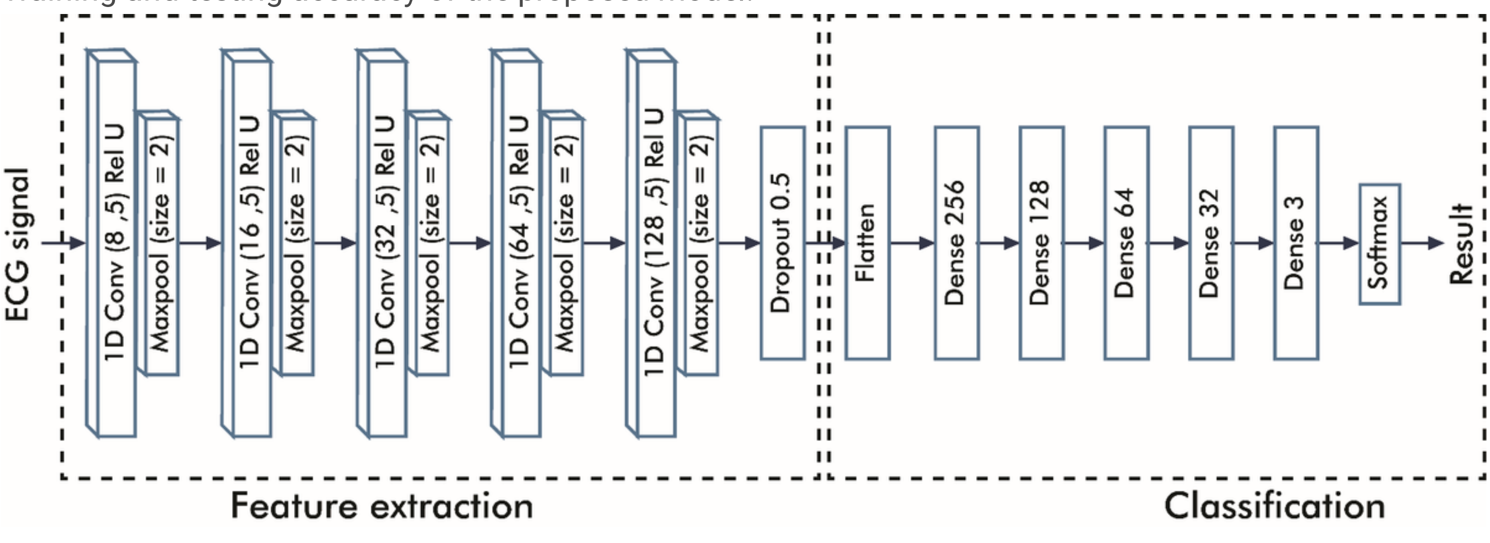

Figure 2

Proposed model of the one-dimensional convolutional neural network for AF and CHF classification.