Chapter 4
Print ISBN: 978-93-5547-238-0, eBook ISBN: 978-93-5547-243-4

ECG Classification for Heart Arrhythmia Using Deep Machine Learning

Shalin Savalia[1*] and Vahid Emamian[2]

DOI: 10.9734/bpi/nvst/v9/14430D

ABSTRACT

Healthcare professionals commonly use Electrocardiogram (ECG) as a low-cost diagnostic tool for monitoring heart electrical signals. Arrhythmia, which is an abnormal heart signal, can be dangerous and cause death. The arrhythmia can be categorized in various types including tachycardia, bradycardia, supraventricular arrhythmias, and ventricular. The automated monitoring of arrhythmia and classification with ECG is very helpful for doctors. In this research we use deep machine learning for automated arrhythmia classification with the focus on the recent trends in arrhythmia classification. Using St. Mary's University Deep Learning Platform, we conducted heavy and complex simulations to measure the performance of the various arrhythmia classification and detection models. Finally, we present the accuracy of the proposed deep learning algorithms, which surpasses the performance of the existing algorithms in precision and sensitivity.

Keywords: Electrocardiogram (ECG); arrhythmia classification; deep machine learning.

1. INTRODUCTION

ECG is often used to evaluate the electric activity of the heart by simply placing number of electrodes on various parts of the skin and has been broadly used for identifying heart diseases due to its simplicity and non-invasive nature. By examining the electrical activity of each heartbeat, i.e., the mixture of action instinct waveforms produced by different cardiac tissues found in the heart, it is possible to identify some of its heart abnormalities. Features like P waves, T waves, QRS complex can be extracted from ECG and studying and classifying them is crucial in diagnosis of various heart disease [1]. An ECG signal with a its features is shown in Fig. 2. Detecting and classifying various kinds of arrhythmias is possible by studying such feature, including abnormal heart rate or abnormal features of the signal. Abnormal heartbeats known as arrhythmias has its unique pattern, therefore it is possible to classify and detect the type [2]. Two main categories are considered for arrhythmias. The first category consists of arrhythmias shaped by an abnormal heartbeat, a.k.a as morphological arrhythmia. The other category is arrhythmias formed by a set of abnormal heartbeats, known as rhythmic arrhythmias. Abnormal heartbeats form alterations in the wave frequency or morphology, and all these alterations may be detected by the ECG tests. Our focus in this research is on recognizing heart diseases by using ECG feature extraction and deep machine learning. This is feasible by classifying regular and irregular ECGs by using deep neural networks techniques and subsequently extract the features of the ECG signals. We use state-logic machine algorithm which is able to identify heart diseases, such as bradycardia, tachycardia, and first and second-degree Atrioventricular (AV) block [3]. There are other arrhythmias types that are described in this research, mainly ventricular tachycardia, atrial fibrillation, malignant ventricular, atrial flutter, and ventricular bigeminy. We are able to detect these with deep machine learning algorithms [4].

Researchers normally try to distinguish regular and irregular heartbeats automatically. These researchers have typically followed three steps: signal processing, feature extraction, and

[1]Department of Electrical Engineering, St. Mary's University, 1 Camino Santa Maria, San Antonio, TX 78228, USA.
[2]School of Science, Engineering and Technology, St. Mary's University, San Antonio, TX 78228, USA.
**Corresponding author: E-mail: shalin2395@gmail.com;*

classification. Signal processing of raw ECG signals will extract the heartbeats from the continuous ECG waveform into individual heartbeats [5]. Next, feature extractions transform the variable-length time-domain heartbeat waveforms into fixed-length feature vectors that encode the heartbeat's features [6]. Number of features have been extracted from ECG signals to identify the heartbeats, such as morphological features, Hermite coefficients, wavelet transform features, heartbeat interval features, and sparse decomposition. On the other hand, for classification, various deep machine learning algorithms have been used here, including artificial neural networks (ANNs), deep neural networks (DNNs), support vector machines (SVMs), convolutional neural networks (CNNs), and multi-layer perceptron (MCP). When there are a large number of datasets available, machine learning techniques are a good to consider and often exceed human agreement rates [7].

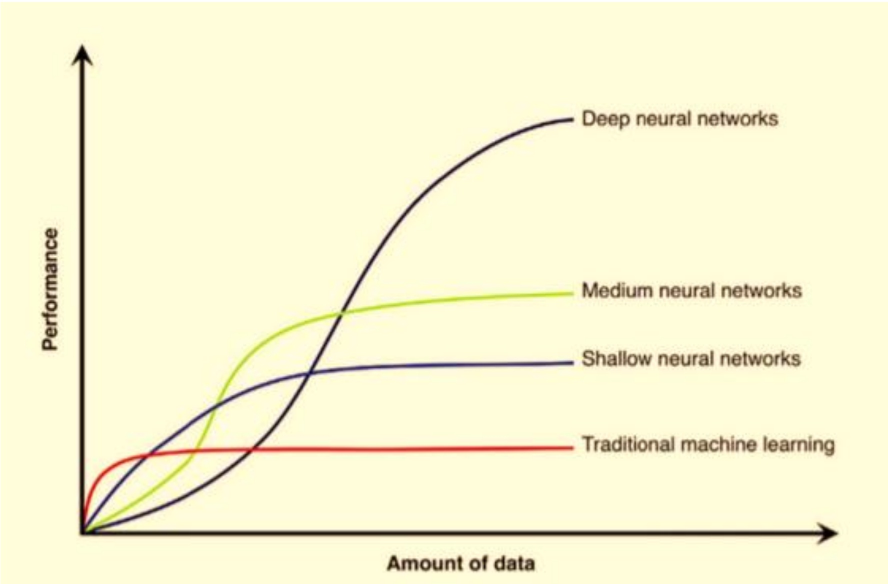
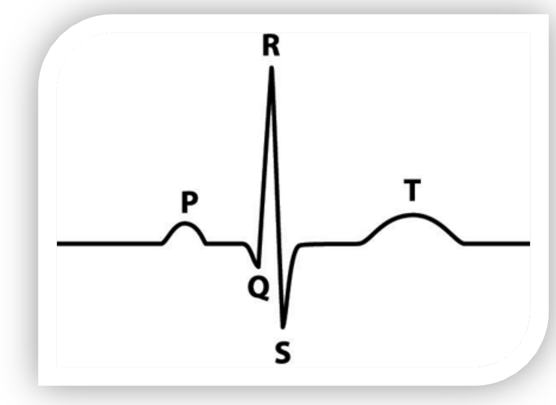


Fig. 1. In traditional Machine learning methods, most of the features need to be extract by a domain expert to reduce the ease of the data and make patterns more noticeable to learning algorithms to be successful. The major improvement of Deep Learning algorithms is that they aim to learn high-level features from data in an incremental manner. This reduces the need of domain expertise and hard-core feature extractionHence DNN have far better performance when we compare it with traditional methods

CNN was used for automated detection of coronary artery disease, and it was found that CNN remains robust despite shifting and scaling invariance which makes it a better choice [8]. In this research, we propose robust methods for heart disease diagnosis using CNN and multilayer perceptron (MLP). We also use CNN to distinguish normal and abnormal heart sound recordings with an accuracy of 82% [8]. The deep machine learning method for single-image super-resolution (SR) was also tested using a CNN algorithm with better performance compared to the state-of-the-art method [9]. In the 2017 PhysioBank competition, Fernando et al. [10] proposed a method with an accuracy of 83% on PhysioBank data, which applies CNN to identify four different arrhythmias from short segments of ECG recordings. In the same competition, Ghiasi et al. [11] proposed algorithms to detect atrial fibrillation using a feature-based algorithm and CNN with 80% accuracy on training datasets.

Key distinction between simple neural network and deep neural network (DNN) method is the problem-solving approach. DNN has shown to solve the problem end to end, whereas simple neural network methods need problem statements to break down to multiple parts to be solved in the beginning and then the results will be merge at final stage. Normally, when there are more than three layers of neurons, including input and output, the method is referred to as "deep learning" [12]. Fig. 3 showed differences between simple NN and deep NN. Usually, a deep neural network algorithm takes longer time to train compares to simple neural network due to large number of parameters it contains.

However, using the state of art HPC deep learning platform at St. Mary's University, we were able to run deep learning algorithms very fast. The main advantage of DNN is that it can detect more complex features compares to simple neural network because of the number of hidden layers. This function of DNN makes it able to handle large amount of data which contains a large number of features. Deep learning neural networks often end in an output layer: a logistic, softmax, or classifier that assigns a chance to a particular outcome [12].



(**a**)

P wave	Depolarization of the atria
Q wave	Activation of the anterioseptal region of the ventricular myocardium
R wave	Depolarization of the ventricular myocardium
S wave	Activation of the posteriobasal portion of the ventricles
T wave	Rapid ventricular repolarization

(**b**)

Fig. 2. Ideal electrocardiogram (ECG) signal with key features indicated; (a) P wave, QRS complex, and T wave which play important roles in diagnosis abnormality of heart signal; (b) Features of an ECG signal; how and which part of heart is used to generates each feature [13]

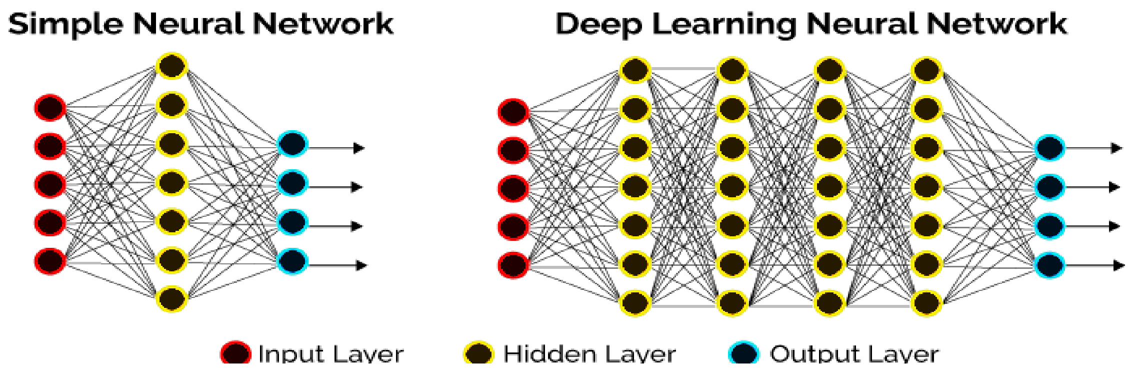


Fig. 3. Comparison between simple neural network (NN) and deep NN; simple neural networks contain only one hidden layer as well as the input and output layers, while deep learning neural networks contain more than one hidden layer. In this case, there are four hidden layers between the input and output layers [12]

For our proposed algorithms, we used two PhysioBank datasets, i.e. normal sinus rhythm database (NSR-DB), and MIT/BIH arrhythmia database to classify regular and irregular ECG signals, using the multilayer-perceptron technique. Another technique uses a four-layer of CNN to detect several arrhythmias in arbitrary length ECG dataset features. The dataset used in this study contains various heart diseases, such as normal sinus, arrhythmia, second degree AV block, first degree AV block, atrial flutter, malignant ventricular, atrial fibrillation, ventricular tachycardia, and ventricular bigeminy. The data were downloaded from kaggle.com. The NN algorithms were trained using Google TensorFlow library, which is a free and open-source software library for machine and deep learning. The algorithm can be utilized across a range of tasks but has a specific focus on training for deep neural networks. Once both algorithms had been trained on the downloaded dataset, they were trained using another dataset with separate characteristics from the training dataset.

2. APPROACH

2.1 Problem Formulation

The proposed algorithm for detection and classification of ECG arrhythmias is a sequence-to-sequence task which takes an input, i.e. ECG signal, S = [s1, …, sk] and gives labels as an output in the form of r = [r1, …, rn], where r_i can take any of m labels. Here, for a multilayer perceptron algorithm, we used m = 2, and for the CNN algorithm, we used m = 9. The individual output label refers to a segment of the input and output labels cover the whole sequence [14].

For a example in the training set, we enhance the cross-entropy function below

$$\mathrm{L}(S, \mathrm{r}) = \frac{1}{n}\sum_{i=1}^{n} \log p\ (R = r_i \ \mid S) \tag{1}$$

where p is the probability the network assigns to the i_{th} output, taking on the value r_i.

2.2 Convolutional Neural Network (CNN)

CNNs here had significant advantage over other neural networks methods by their superior performance with input signals and images. They are feed-forward ANN inspired by biological processes and are intended to identify patterns directly from data or images, by integrating both feature extraction and classification [14]. A CNN involves three major layers: convolutional layer, pooling layer, as well as fully connected layer. The convolutional layer is the most significant block of a CNN, and it is where majority of the computation occurs. This layer normally requires the following components; i) input data, ii) a filter, and iii) a feature map [15]. Convolutional layers are followed by an activation layer that is non-linear. This helps capturing more complex parameters of the input signal possible. Pooling layers utilized to subsample the last layer by mixing small rectangular subsets of values. This layer, also known as down sampling layer, performs dimension reduction, reducing the number of parameters in the input 18]. Like the convolutional layer, the pooling layer sweeps a filter across the whole input, however, difference is that this filter does not have weights. There are two types of pooling: max pooling and average pooling. Max and average pooling are applied by replacing the input values with the maximum or the average values, respectively [16]. On the other end, a significant amount of data is lost in the pooling layer, it also has several benefits to the CNN as this layer helps to lower complication, increase efficiency, and limit risk of overfitting. Fully connected layer (FCL) does the job of classification based on the features extracted from the past layers and their different filters [16]. While convolutional and pooling layers utilize the ReLU functions, FCL typically leverage a Softmax activation function that classifies inputs appropriately, generating a probability from 0 to 1.

2.3 Multilayer Perceptron (MLP)

Deep neural networks made of multiple layers and multilayer perceptron (MLP) indicates that it is made of more than one perceptron. A single layer perceptron can solve linearly separable problems but when one or more layers are added in single layer perceptron, it is known as MLP [16]. MLP network is recognized as feed-forward neural network that consist of one or several hidden layers and is usually used for classification of input patterns, pattern recognition, prediction based on the input data and approximation [16]. MLP typically made up of an input layer which receive the data, an output layer which makes a decision or prediction about the input data, and in between those two, any number of hidden layers that are the computational portion of the MLP [16,17].

MLPs are often used to supervise learning. They train on a set of input-output data and learn to model the correlation or dependencies between input and output data sets. Training requires modification of the parameters, i.e. the weights and biases of the model to minimize the error [17]. Backpropagation is also applied to make those weights and bias adjustments relative to the error. Here the error can be calculated in various ways, for instant, by using root mean squared error.

2.4 Model Architecture

CNN and MLP with several hidden layers used for sequence-to-sequence learning algorithms. The CNN is one of the central branches of deep, feed-forward machine learning neural networks that can handle large data. CNN, like any other neural networks, has input, output, and a several of hidden layers. The hidden layers mainly made of convolutional layers, pooling layers, normalization layers, fully connected layers, and finally softmax layers. Our proposed CNN algorithm has a convolutional layer with softmax function that provides the output for the trained network. It uses the ReLU, a.k.a rectifier linear unit and activation tool in all convolution layers. The max pooling layer functions independently for each column and row of the input and resizes them spatially [18]. We used the max pooling layer with stride size of 2 by 2 in the function because it gave improved accuracy than original 3 by 3 pooling layer. Use of a 3 by 3 stride layer yields higher info loss. The pooling layer in the CNN lowers the overfitting problem by reducing the input size by half of the actual input. Flowchart diagram of both algorithms are depicted in Fig. 4. Both models take features of an ECG signal as the input of the network and predict the output as labels of the data. Initially, ECG datasets are pre-processed, and for that purpose, the first network reads the datasets and identifies their features and labels. For the MLP, the labels will be arrhythmia and regular sinus, while in the CNN, the labels are normal sinus, arrhythmia, first degree AV block, second degree AV block, atrial fibrillation, atrial flutter, malignant ventricular, ventricular bigeminy, and ventricular tachycardia [15]. Fig. 5 shows our proposed architecture of the CNN in the algorithm where the first and last convolutional layers are not the same as the middle three convolutional layers.

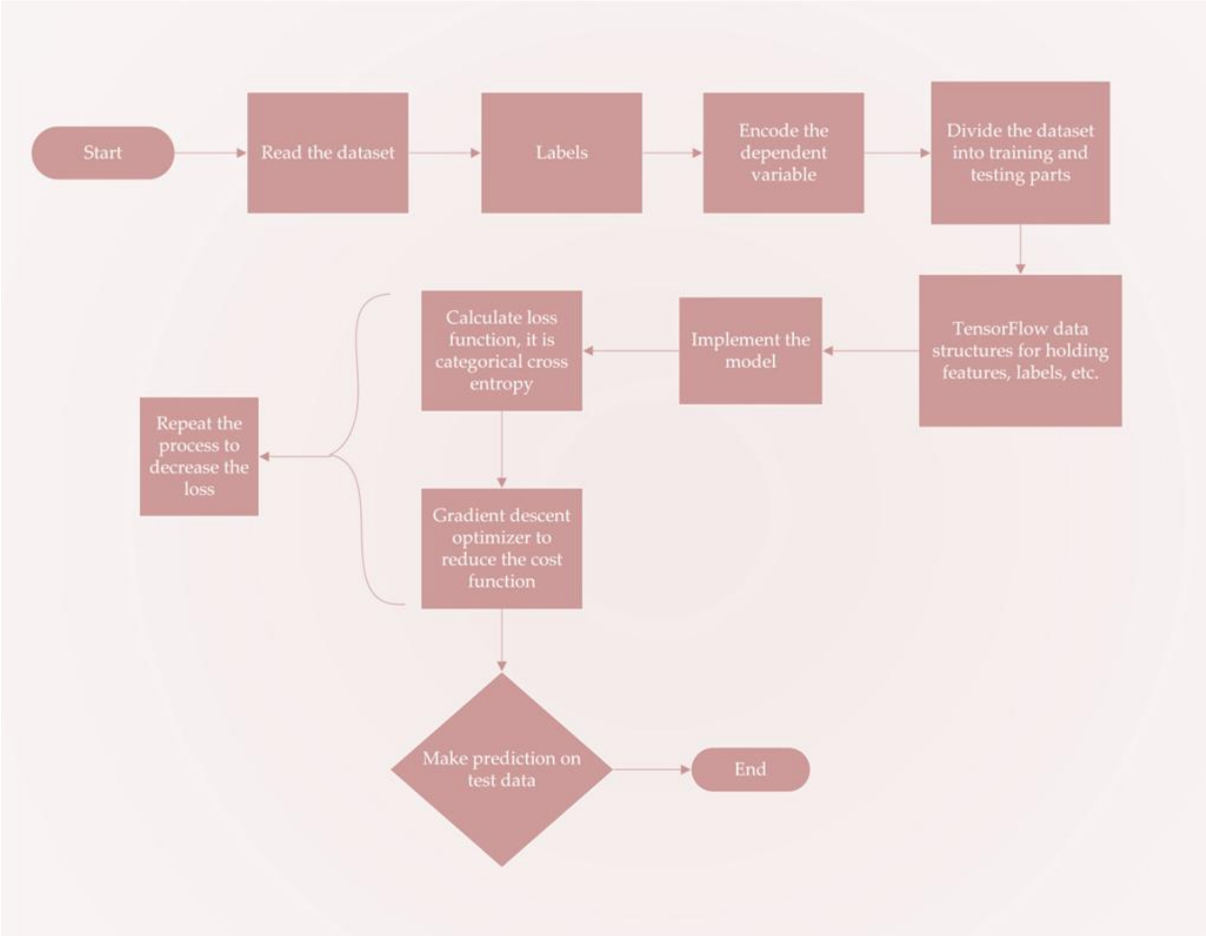


Fig. 4. System process flowchart of Multilayer Perceptron (MLP) and Convolution NN
To define features and labels in the dataset, two TensorFlow variables were defined. One hot encoder was used to encode the dataset

The next step is to encode the dependent variable—the dataset labels—for the deep network. As the data is classified as containing different arrhythmia names, it is mandatory to encode the data because the labels are not numerical, and it is not possible to read them directly by the algorithm [19]. There are two statistical ways for encoding data. First is integer encoding and second is one-hot encoding. Integer encoding normally assigns an integer value to each unique category. For example "red", "green", and "blue" are 1, 2, and 3 respectively [19]. For categorical variables where no ordinal correlation recognized, integer encoding is not enough. In one-hot encoding, the integer encoded variable is discarded and a new binary variables are added for each unique value of integer. In the "color" variable, there are three classes and consequently three binary variables are required. A "1" value is inserted in the binary variable for the color and "0" values are placed for the other colors. In our proposed machine learning algorithm, we used one-hot encoding to distinct integer encoding. This was then followed by dividing the data into training, testing, and validation sets [19].

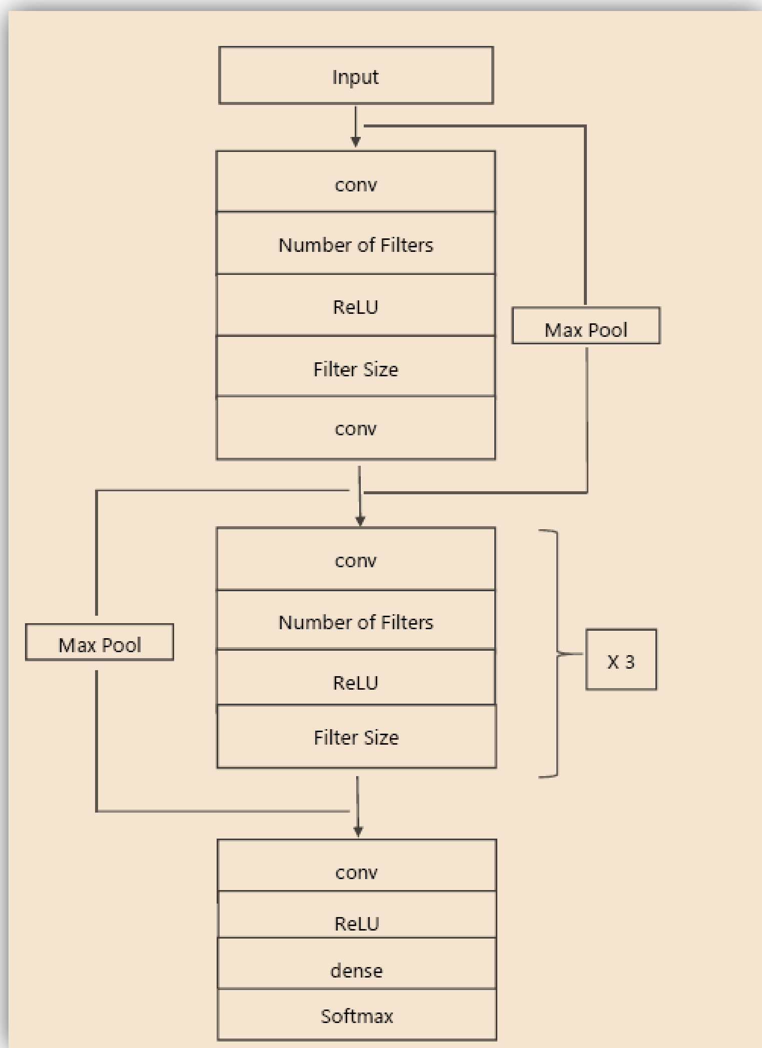


Fig. 5. Proposed Algorithm for CNN

TensorFlow data structures were defined for holding features and labels, which includes defining weights, hidden layers, biases, activation tools, placeholders for inputs, filters, filter size, and desired output. Another Tensor defined to store trained output model. This was then followed by training of the proposed model with the training dataset. After training the network, it will compute how different the trained model's output is from the actual output. After reaching the minimum value, the trained model will yield testing accuracy by training it using the test data [20]. The cross-entropy function is also used to reduce the error to a minimum level.

2.5 ECG Data

We download ECG datasets from PhysioBank.com and kaggle.com for training and testing the proposed CNN and MLP algorithms. The MLP dataset size was 208x61, where total ECG signals were 208 and the total number of features and labels were 61. The first 60 columns contained features, and the last column contains the diseases label for each individual data. The CNN dataset however had dimensions of 26x543x60, following the same configuration as the MLP dataset, but this dataset consisted of 9 labels. Both algorithms consisted of 80% of the total data for training and 20% for testing. The training dataset was separated into 70% for actual training and 30% for validation. Each ECG vector in the dataset was 10 second long and contained only one rhythm class. A demonstration of the distribution for the ECG signals used for training, testing, and validation process is shown in Fig. 6.

2.6 Training of Data

We used batch sizes of 50 for the training stage with the standard back propagation for stochastic learning. To update the weights we used the following formula [21]:

$$w_l = \left(1 - \frac{n_\lambda}{ts}\right) w_{l-1} - \frac{n}{x}\frac{\partial c}{\partial w} \tag{2}$$

Where

$\boldsymbol{w}$ = weights
l = layer number
$\boldsymbol{n}$ = learning rate
c = cost function
$\boldsymbol{ts}$ = total number of training samples
$\boldsymbol{x}$ = batch size
$\boldsymbol{\lambda}$ = regulation parameter

Additionally, the biases in the formula were updated through,

$$b_l = b_{l-1} - \frac{n}{x}\frac{\partial c}{\partial w} \tag{3}$$

In the proposed algorithms, the deep neural learning rate were set to 0.002 for MLP and to 0.003 for CNN.

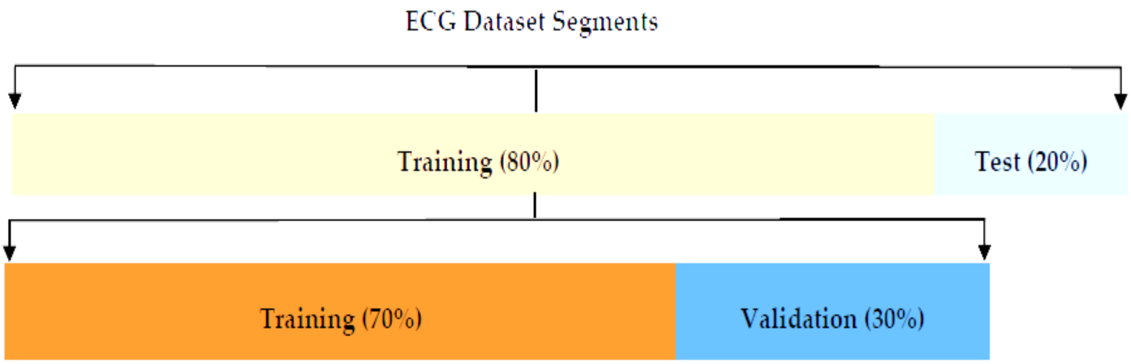


Fig. 6. ECG segments distribution in training and testing data set

2.7 Testing of Data

After running each training epoch, the algorithms are expected to perform testing on the CNN and MLP algorithms to give test accuracy. The CNN and MLP algorithms have 500 and 1000 epochs respectively. We applied 30% of the total training data, that is 80% of the original dataset, to validation part after completion of every epoch for better accuracy. Testing and validation dataset consisted of 20% of the total data used for testing, as described in previous sections [22].

3. RESULTS

We concluded that CNN networks have a remarkable ability to extract all differentiable features invariant to local temporal and spectral variations. This has resulted in significant breakthroughs in higher accuracy results. The proposed CNN algorithm contained the following stages; i) data preprocessing; processing of ECG signals (so that automatic algorithm can understand different diseases), ii) stacking of max pooling layers and convolution layers to extract the known features, iii) activation of the softmax function and layering of a fully connected layer to predict the disease [13]. Table 1 shows the parameters of the CNN layers, the size off filter, and output layer neuron size. To distinguish between regular sinus rhythm and irregular rhythm, we used MLP. To accomplish this, we used four hidden layers each consisting of 60 neurons. The ReLU was then used to activate the first as well as last hidden layers, while two middle hidden layers employed a sigmoidal activation function. At the output layer, these were followed by the linear activation function. Additionally, a gradient descent optimizer was employed to lower the error between the actual output and the trained network output. When the parameters cannot be calculated analytically or by linear algebra, we realized that it was advantageous to implement a gradient descent optimizer. Fig. 7 depicts the accuracy and mean square error graphs for the MLP algorithm.

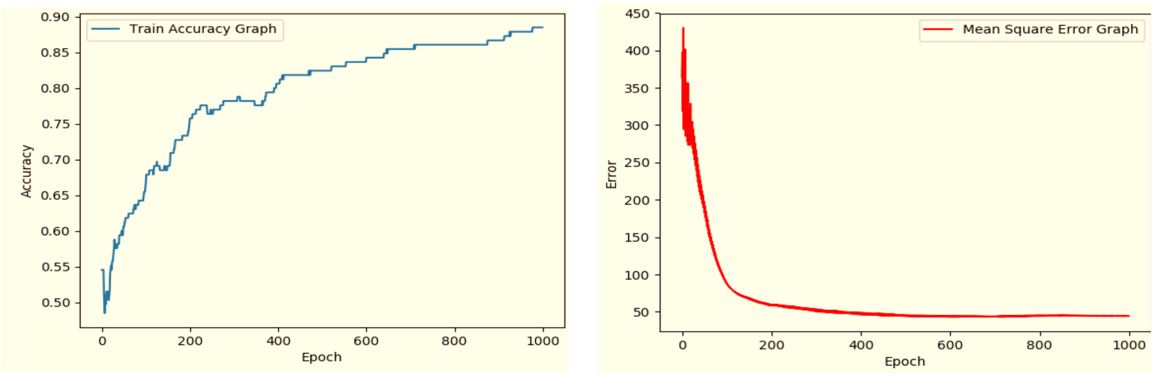
After training the network with 1000 epochs, it yielded an accuracy of about 89% for the dataset from PhysioBank.net. Fig. 8 depicts the visual confusion matrix for the training dataset. The confusion graph is a demonstration of true label vs predicted label, where 0 stands for irregular ECG signal and 1 represents regular sinus rhythm. The dataset used here consisted of 208 ECG data recordings, 97 of which are irregular (arrhythmia) and 111 represent a regular sinus rhythm. The 80% of the data applied for training, consisting of 165 ECG signals, 72 describes arrhythmia and 93 represent regular sinus rhythm. From this training dataset, 63 arrythmia and 81 regular sinus signals were rightfully classified by the algorithm, demonstrating a significant improvement in the accuracy of the MLP algorithm.

Table 1. Parameters for the proposed CNN algorithm

Layers	Type	Size of Neurons (Output Layer)	Filter Size of Each Layer
0–1	Convolution	(None, 1, 60, 1)	32
1–2	Max Pooling	(None, 1, 30, 1)	2
2–3	Convolution	(None, 1, 30, 1)	32
3–4	Max Pooling	(None, 1, 15, 1)	2
4–5	Convolution	(None, 1, 15, 1)	32
5–6	Max Pooling	(None, 1, 8, 1)	2
6–7	Convolution	(None, 1, 8, 1)	32
5–6	Fully connected layer	2048	-

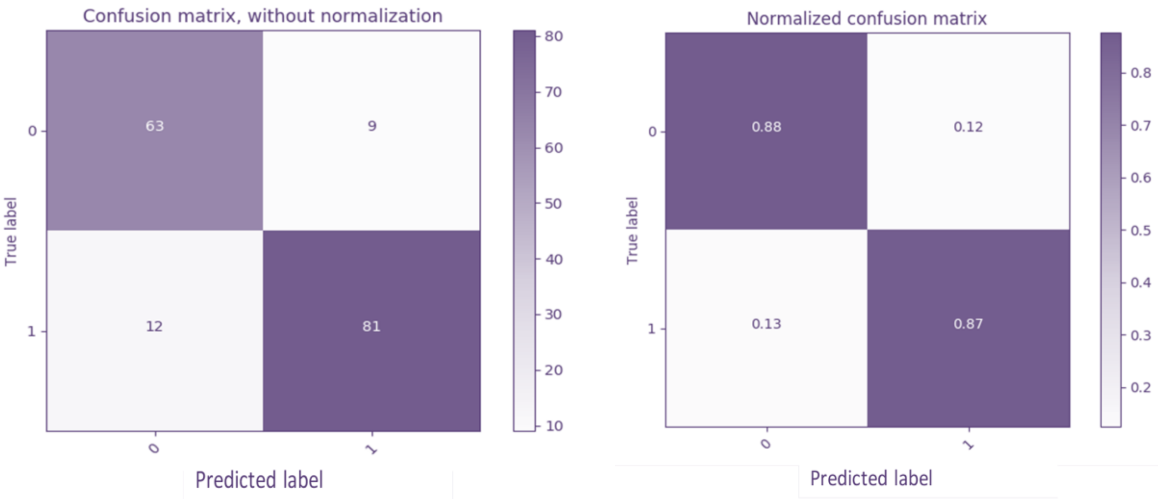
We then used the deep machine learning technique to create CNN to identify different cardiovascular diseases. Here, we used the ReLU non-linear activation tool to activate the CNN alongside the gradient descent optimizer to minimize the error. This method proved to work great when the parameters cannot be obtained analytically [23]. The architecture of the CNN for each convolution layer had 32 filters and each filter had a size of 5 by 5. Fig. 9 shows the accuracy and MSE error. On average, the accuracy improves with every epoch and after about 500 epochs, reaches the max of 83.5%. The error reduces continuously with each epoch and gradually reaches a minimum.

We defined two variables as features and labels for the datasets. The proposed algorithm reshaped dimensions of features 1 by 4 because the convolution layer here only accepts 4-dimension vectors [24-27]. Upon completion of the simulation using the deep learning platform HPC, the first three convolution layers were defined. The output of first layer was fed into the max pooling layer to reduce the dimension of the vector to make the network faster and also to avoid overfitting. The same approach was followed for the second and third convolution layers as well. The result of the third pooling layer was then fed into the fully connected layer, followed by the softmax layer so that the algorithm predicts the diseases [28]. Same as before, we showed the classification results of the algorithm by using a confusion matrix. Each cell consisted of the raw number of exemplars classified for the matching combination of expected and actual outputs. Fig. 10 below shows a visual representation of the confusion matrix for the proposed CNN algorithm. We noticed that many arrhythmias were confused with first-degree AV Block and ventricular bigeminy, but overall the network gives a good prediction accuracy for the other diseases. We anticipate part of this is due to the ambiguous location of the exact onset and offset of the arrhythmia in the ECG vectors [29-31].



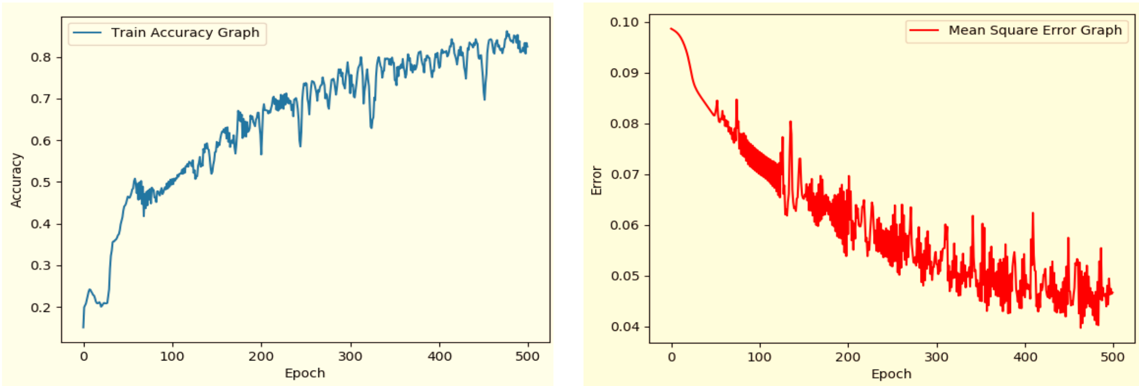
(**a**) (**b**)

Fig. 7. Accuracy and mean square error for the MLP algorithm; (a) accuracy increases as the number of epochs increases; (b) error reduces with every epoch reaches the lowest after 1000 epochs



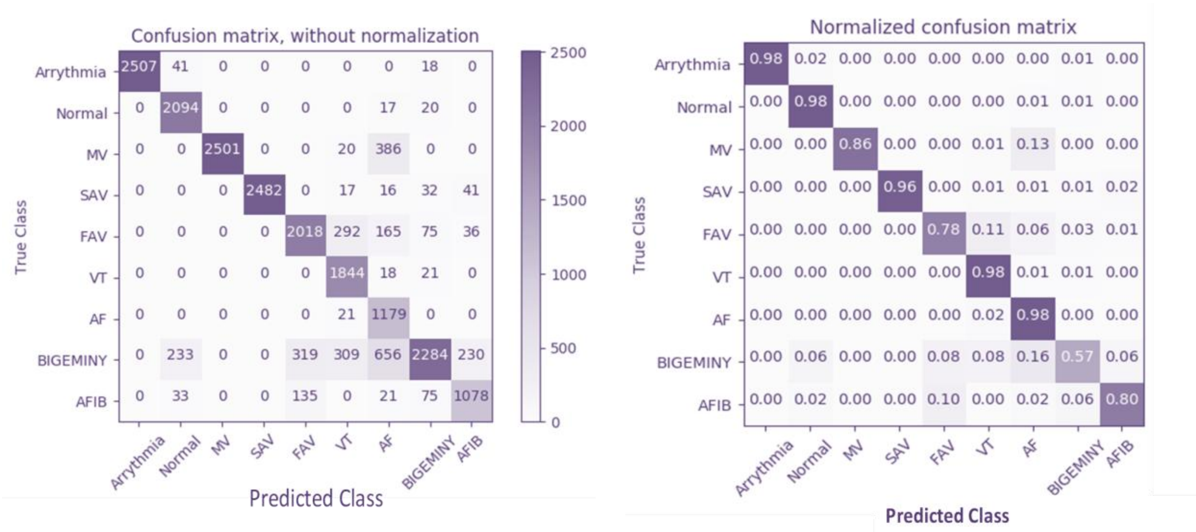
(**a**) (**b**)

Fig. 8. Confusion matrix (CM)



(**a**) (**b**)

Fig. 9. Accuracy and error of the proposed CNN



(**a**) (**b**)

Fig. 10. CM of the CNN algorithm

4. CONCLUSIONS

In this article we used deep machine learning for automated arrhythmia classification with the focus on the recent trends in arrhythmia classification. Using St. Mary's University Deep Learning Platform, we conducted heavy and complex simulations to measure the performance of the various arrhythmia classification and detection models [32]. The proposed algorithms in this research were tested on ECG signals obtained from Physio.net and keggar.com and succeeded in detecting abnormal states in each signal with significant accuracy using MLP and CNN models. Our results showed that the proposed algorithms can make accurate diagnoses of various heart diseases with 89% and 83% accuracy for the proposed MLP and CNN algorithms respectively [32].

AUTHOR CONTRIBUTIONS

Savalia and Emamian proposed the idea, revised the work and contributed the materials and analysis tools. Savalia interpreted the data, designed and performed the simulations, and wrote the paper. Their effort supported by the DURIP Grant W911NF-19-1-0159 (73980-NS-RIP) to St. Mary's University. The simulations were performed on St. Mary's Deep Learning HPC.

REFERENCES

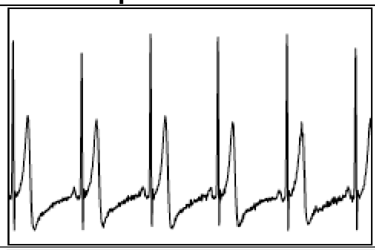
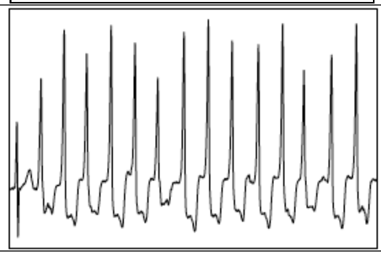
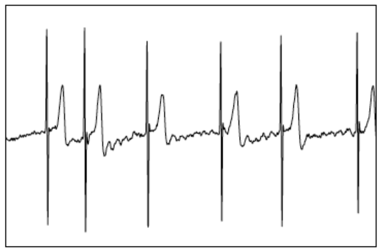
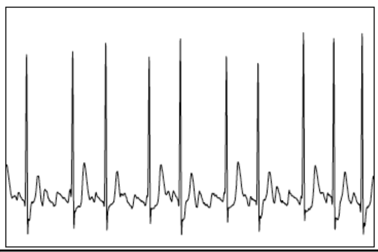
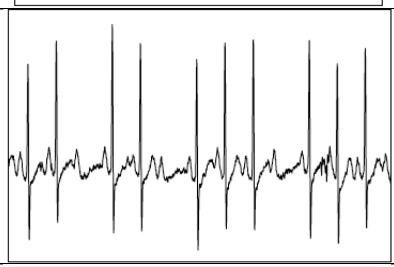
1. Artis SG, Mark RG, Moody GB. Detection of atrial fibrillation using artificial neural networks. In Computers in Cardiology 1991, Proceedings of the 18th Annual Conference on Computers in Cardiology, Venice, Italy, 23–26 September 1991; Alan, A.R., Werner, R., et.al., Eds; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 1991;173–176.
2. Melo SL, Caloba LP, Nadal J. Arrhythmia analysis using artificial neural network and decimated electrocardiographic data. In Computers in Cardiology 2000, Proceeding of the IEEE Conference on Computers in Cardiology, Cambridge, MA, USA, 24–27 September 2000; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2000;73–76.
3. Moody, G.B.; Mark, R.G. The impact of the MIT-BIH arrhythmia database. IEEE Eng. Med. Biol. Mag. 2001;20:45–50.
4. Salam AK, Srilakshmi G. An algorithm for ECG analysis of arrhythmia detection. In Electrical, Computer and Communication Technologies (ICECCT), Proceedings of 2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 5–7 March 2015; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2015;1–6.
5. Perez RR, Marques A, Mohammadi F. The application of supervised learning through feed-forward neural networks for ECG signal classification. In Electrical and Computer Engineering (CCECE), Proceedings of 2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Vancouver, BC, Canada, 15–18 May 2016; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2016;1–4.
6. Palreddy S, Tompkins WJ, Hu YH. Customization of ECG beat classifiers developed using SOM and LVQ. In Engineering in Medicine and Biology Society, 1995, Proceedings of IEEE 17th Annual Conference on Engineering in Medicine and Biology Society, 1995, Montreal, QC, Canada, 20–23 September 1995; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 1995;813–814.
7. Zebardast B, Ghaffari A, Masdari M. A new generalized regression artificial neural networks approach for diagnosing heart disease. Int. J. Innov. Appl. Stud. 2013;4:679.
8. Acharya UR, Fujita H, Lih OS, Adam M, Tan JH, Chua CK. Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network. Knowl.-Based Syst. 2017;132:62–71
   doi: 10.1016/j.knosys.2017.06.003
9. Nilanon T, Yao J, Hao J, Purushotam S, Liu Y. Normal/abnormal heart recordings classification by using convolutional neural network. In Computing in Cardiology Conference (CinC), 2016, Proceedings of the IEEE Conference on Computing in Cardiology Conference (CinC), Vancouver, BC, Canada,11–14 September 2016, Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2016;585–588.
10. Savalia S, Acosta E, Emamian V. Classification of cardiovascular disease using feature extraction and artificial neural networks. Journal of Biosciences and Medicines. 2017;5:64,
11. Jigar AD. Swift single image super resolution using deep convolution neural network. In Communication and Electronics Systems (ICCES), Proceedings of International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 21–22 October 2016; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2016;1–6.
12. Shadi G, Mostafa A, Nasimalsadat M, Kamran K, Ali G. Atrial fibrillation detection using feature based algorithm and deep conventional neural network. In Computing in Cardiology (CinC), 2017, Proceedings of conference on Computing in Cardiology (CinC), Rennes, France, 24–27 September 2017; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA; 2017.
13. Fernando A, Oliver C, Marco AFP, Adam M, de Maarten V. Comparing feature based classifiers and convolutional neural networks to detect arrhythmia from short segments of ECG. In Computing in Cardiology (CinC), Proceedings of conference on Computing in Cardiology (CinC), Rennes, France, 24–27 September 2017; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA; 2017.
14. Chow GV, Marine JE, Fleg JL. Epidemiology of arrhythmias and conduction disorders in older adults. Clin. Geriatr. Med. 2012;28:539–553.

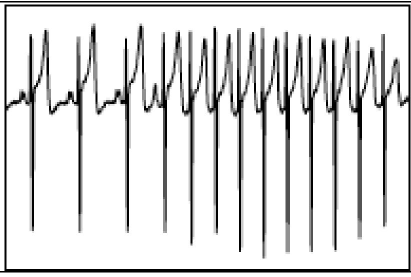
15. Oquab M, Bottou L, Laptev I, Sivic J. Is object localization for free?—Weakly-supervised learning with convolutional neural networks. In Computer Vision and Pattern Recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 12 June 2015; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2016;685–694.
16. Acharya UR, Fujita H, Adam M, Oh SL, Tan JH, Sudarshan VK, Koh JEW. Automated characterization of arrhythmias using nonlinear features from tachycardia ECG beats, In Systems, Man, and Cybernetics (SMC), Proceedings of 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 9–12 October 2016. Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2016.
17. Desai U, Martis RJ, Acharya UR, Nayak CG, Seshikala G, Shettyk R. Diagnosis of multiclass tachycardia beats using recurrence quantification analysis and ensemble classifiers. J. Mech. Med. Biol. 2016, 16, doi: 10.1142/S0219519416400054.
18. Zubair M, Kim J, Yoon CW. An automated ECG beat classification system using convolutional neural networks. In IT Convergence and Security (ICITCS), Proceedings of the 2016 6th International Conference on IT Convergence and Security (ICITCS), Prague, Czech Republic, 26–29 September 2016; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2016;1–5.
19. Krizhevsky A, Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems 25 (NIPS 2012). Proceedings of the Neural Information Processing Systems 2012, Lake Tahoe, NV, USA, 3–8 December 2012; Pereira, F., Burges, C.J.C., et al., Eds.; The MIT Press: Cambridge, MA, USA, 2012; 1097–1105.
20. Kiranyaz S, Ince T, Gabbouj M. Real-time patient-specific ECG classification by 1-D convolutional neural network. IEEE Trans. Biomed. Eng. 2016;63:664–675.
21. Fahim S, Khalil, I. Diagnosis of cardiovascular abnormalities from compressed ECG: A datamining-based approach. IEEE Trans. Inf. Technol. Biomed. 2011;15:33–39.
22. Ciresan DC, Meier U, Gambardella LM, Schumacher J. Convolutional neural network committees for handwritten character classification. In 2011 International Conference on Document Analysis and Recognition (ICDAR 2011). Proceedings of 2011 International Conference on Document Analysis and Recognition (ICDAR), Beijing, China, 18–21 September 2011; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2011;1135–1139.
23. Sainath TN, Mohamed A-R, Kingsbury B, Ramabhadran B. Deep convolutional neural networks for lvcsr. In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Proceedings of 2013 IEEE international conference on Acoustics, speech and signal processing (ICASSP), Vancouver, BC, Canada, 26–31 May 2013; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2011;8614–8618.
24. Golkov V, Dosovitskiy A, Sperl JI, Menzel MI, Czisch M, Sämann P, Brox T, Cremers D. q-Space deep learning: Twelve-fold shorter and model-free diffusion MRI scans. IEEE Trans. Med. Imaging 2016;35:1344–1351.
25. Martis RJ, Acharya UR, Adeli H, Prasad H, Tan JH, Chua KC, Too CL, Yeo SWJ, Tong L. Computer-aided diagnosis of atrial arrhythmia using dimensionality reduction methods on transform domain representation. Biomed. Signal Process. Control 2014;13:295–305.
26. Martis RJ, Acharya UR, Prasad H, Chua KC, Lim CM, Suri JS. Application of higher order statistics for atrial arrhythmia classification. Biomed. Signal Process. Control 2013;8:888–900.
27. Wang Y, Zhu YS, Thakor NV, Xu YH. A short-time multifractal approach for arrhythmia detection based on fuzzy neural network. IEEE Trans. Biomed. Eng. 2001;48:989–995.
28. Yan Z, Zhan Y, Peng Z, Liao S, Shinagawa Y, Zhang S, Metaxas DN, Zhou XS. Multi-instance deep learning: Discover discriminative local anatomies for body part recognition. IEEE Trans. Med. Imaging 2016;35:1332–1343.
29. He, K, Zhang X, Ren S, Sun J. Identity mappings in deep residual networks. In European Conference on Computer Vision, Proceedings of the 14th European Conference on Computer Vision, ECCV 2016, Cham, The Netherlands, 8–16 October 2016; Springer: Berlin, Germany, 2016; pp. 630–645.

30. Savalia S, Emamian V. Cardiac Arrhythmia Classification by Multi-Layer Perceptron and Convolution Neural Networks. Bioengineering 2018;5:35.
31. Elsayad AlM. Classification of ECG arrhythmia using learning vector quantization neural networks, International conference on computer engineering systems, ICCES 2009; 2009. In Computer Engineering & Systems, 2009, Proceedings of 2009 International Conference on Computer Engineering & Systems, Cairo, Egypt, 14–16 December 2009; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2009;139–144.
32. Savalia S, Emamian V. Cardiac Arrhythmia Classification by Multi-Layer Perceptron and Convolution Neural Networks. Bioengineering 2018;5:35.

APPENDIX A

Table A1. A list of all arrhythmia types which the model classifies. For each arrhythmia, we give the label name, a more descriptive name, and an example chosen from the training set. We also give some description of each arrhythmia type [23]

Class	Description	Example
Normal	Normal Sinus Rhythm means normal heart rate, in respect to both heart rate and rhythm. Heart Rate—60 to 100 BPM	
VT	Ventricular Tachycardia is heart rhythm illness instigated by abnormal signals in the lower chambers of the heart. Heart Rate—More than 100 BPM	
AFIB	Atrial Fibrillation is an irregular and fast heart rate than can increase chance of stroke, heart failure. Heart Rate—100 to 175 BPM	
AF	Atrial Flutter is the same as AFIB. But, whereas AFIB causes increased heart rate without a regular pattern, AFL causes increased heart rate in a regular pattern. Heart Rate—100 to 175 BPM	
SAV	Second Degree AV, is a disease of the cardiac conduction system in which the conduction of atrial impulse over the AV node and/or his bundle is delayed or blocked.	
Bigeminy	Ventricular Bigeminy is a heart rhythm problem in which there is a continuous alternation of long and short heart beats.	

FAV	In First Degree AV Block conduction is slowed, there are no missed beats. In first-degree AV block, every atrial impulse is transmitted to the ventricles, resulting in a regular ventricular rate.	

DISCLAIMER
This chapter is an extended version of the article published by the same author(s) in the following journal.
Bioengineering, 5, 35, 2018.