

# Application of Convolutional Neural Network and Deep Learning for Detection of Cardiac Arrhythmia Heart Disease

*Nahom Ghebremeskel<sup>1\*</sup>, Vahid Emamian<sup>2</sup>*

<sup>1</sup>*Department of Electrical Engineering, St. Mary's University, 1 Camino Santa Maria, San Antonio, TX 78228, USA*

<sup>2</sup>*School of Science, Engineering and Technology, St. Mary's University, San Antonio, TX 78228, USA*

*\*Corresponding Author*

*E-Mail Id: nahom.edu2017@gmail.com*

## ABSTRACT

*The goal of this paper is apply convolutional neural networks to Electrocardiogram signals to detect cardiac arrhythmia, which is a form of heart diseases. The new high-tech Electrocardiogram sensors and other medical devices have significantly improved the quantity and quality of Electrocardiogram recordings in high volume. Furthermore, the availability of high computing GPUs have made it easy to process large amount of data in a short amount of time. We have developed a method for Electrocardiogram arrhythmia classification which converts Electrocardiogram signals to two dimensional images to be processed with convolutional neural networks, which is a form of deep machine learning. Deep learning has been proven to be an effective means for complex data analysis with minimal pre- and post-processing requirement. It is the primary tool in this research. We use the proposed convolutional neural networks architecture for classifying cardiac arrhythmia into three distinct categories: normal sinus rhythm, paced rhythm, and other rhythm. The Electrocardiogram signal is converted into a two-dimensional grayscale image and used as an input data for the convolutional neural networks classifier. We use various deep learning techniques such as batch normalization, data augmentation, and averaging-based feature aggregation across time. We use several image crop techniques for data augmentation and K fold cross validation for overcoming over-fitting. The proposed classifier can reach a classification accuracy of over 95% on the data we acquired from PhysioNet/CinC Challenge 2017.*

**Keywords:** *Deep machine learning, electrocardiogram (ECG), arrhythmia, convolutional neural network (CNN), data augmentation*

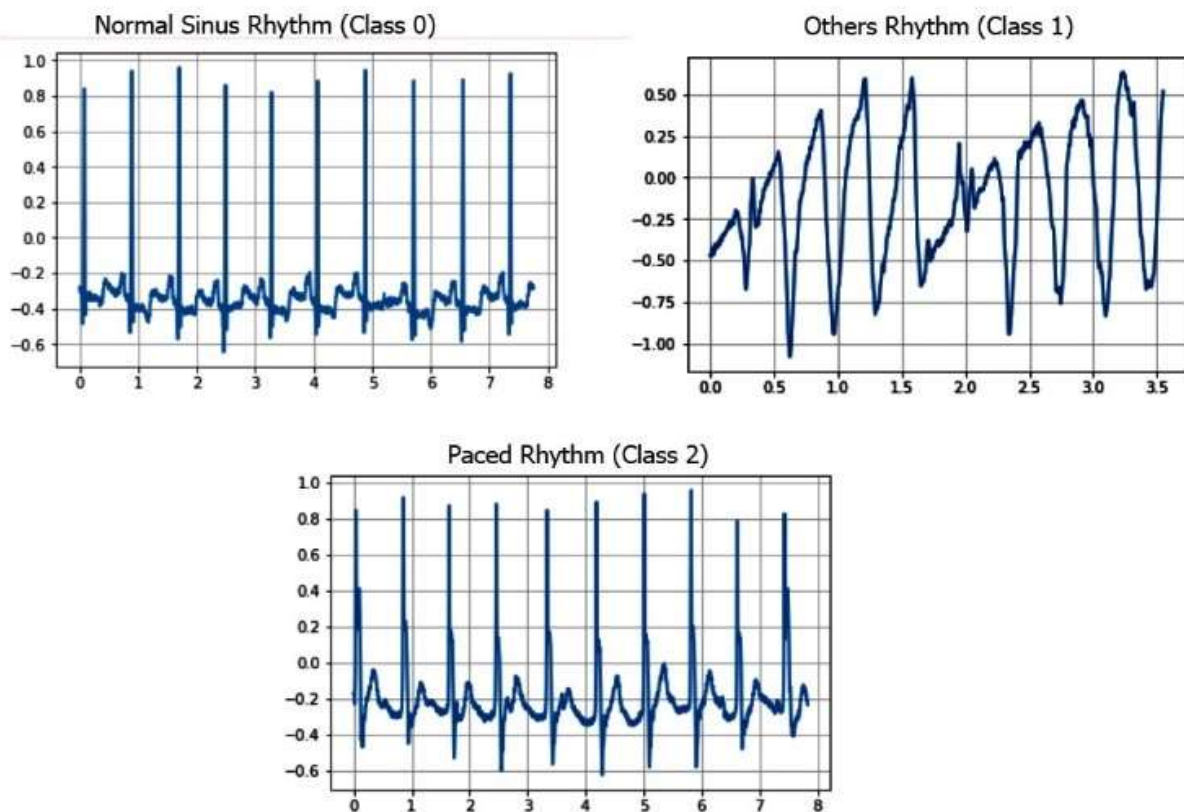
## INTRODUCTION

Arrhythmia is a characteristic type of Cardiovascular Diseases (CVDs) that leads to any irregular change from normal heart rhythms. There are numerous types of arrhythmia including atrial fibrillation, premature contraction, ventricular fibrillation, and tachycardia. Although a single arrhythmia heartbeat may not have an important effect on life, continuous arrhythmia beats can consequence in deadly conditions. For example, the beats of prolonged premature ventricular

contractions (PVCs) seldom turn into ventricular tachycardia (VT) or ventricular fibrillation (VF) beats, which can immediately lead to heart failure [1]. Thus, it is crucial to frequently observe heart rhythms to control and avoid CVDs. An electrocardiogram (ECG) is a medical tool that displays the rhythm and condition of the heart. Therefore, the involuntary result of improper heart rhythms from ECG signals is an important task in the field of cardiology.

Different approaches have been recently researched for automatic identification of ECG arrhythmia based on signal feature extraction, such as support vector machines (SVM) [2,3], discrete wavelet transformation (DWT) [4,5], feed forward neural networks (FFN) [6], learning vector quantization (LVQ) [7,8], back propagation neural networks (BPNN) [9], and regression neural networks (RNN) [10]. When a large amount of data is available, deep learning models are a good approach and often surpass identification

by humans [11]. CNN was used for automated detection of coronary artery disease and it remains robust despite shifting and scaling invariance, which makes it advantageous [12]. In our research, we propose deep neural network architecture for classifying electrocardiogram (ECG) recordings from a single-channel handheld ECG device into three distinct categories: normal sinus rhythm (N), paced rhythms (A), or other rhythm (O).



**Fig. 1:** Three Classes of the Data Set such as Normal Sinus Rhythm (class 0), Paced Rhythm (class 2), Other Rhythms (class 1)

For the classification of arbitrary-length ECG recordings, we evaluate them using the AF (atrial fibrillation) classification data set provided by the PhysioNet/CinC Challenge 2017. AF happens in 1-2% of the population due to an increase in age and is associated with significant mortality rate and disease. Unfortunately, current AF classification methods are unsuccessful

at solving the potential of automated AF classification to have poor generalization capabilities experienced by training and/or evaluation on small and/or carefully selected data sets. Our architecture uses an averaging-based feature aggregation with 24-layer convolutional neural network (CNN). CNNs can extract features invariant to local spectral and spatial/temporal

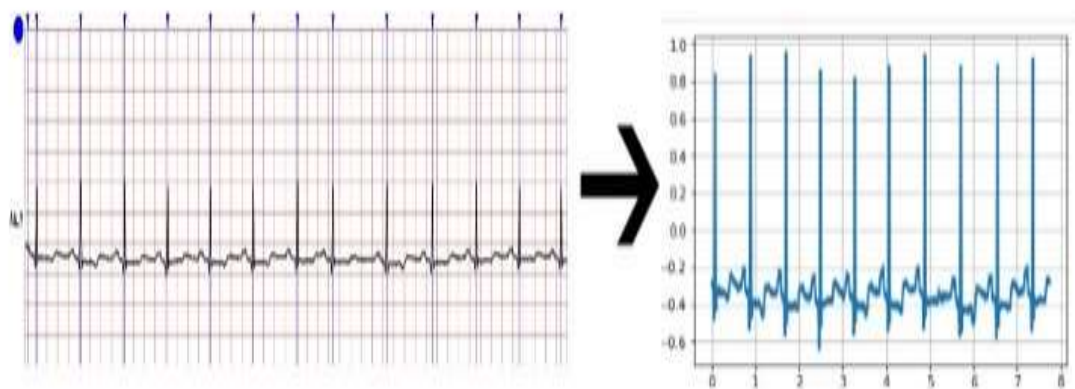
variations, and have led to many breakthrough results, most prominently in computer vision [13].

## METHODOLOGY

In order to classify the input ECG signal into three classes of interest, the recordings are first cut, and the data is nominated based on the labels. After nominating, each data is transformed into an image of grayscale 200 x 200. After that, the ECG images are taken into a CNN for training and testing, a 24-layered deep CNN. The output of those layers is used to extract features. At the end, averaging-based feature aggregation across time is used for classifying the features. Our research consists of the following steps: data processing, feature extraction using block of convolutional layers, and aggregation of features across time by averaging.

## ECG Data Pre-Processing

In this paper, we used the MIT-BIH arrhythmia database [30] for the CNN model training and testing. The MIT-BIH Arrhythmia Database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979 [23]. The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10-mV range [23]. Since the CNN model uses 2D images as input data, we convert the ECG signal into ECG images in the ECG data pre-processing step. The next step is the CNN classifier step in which we use the ECG image to get classification of three ECG types. Overall procedures are shown in Figure 1.



*Fig. 2: MIT-BIH Arrhythmia Transformed to ECG Image.*

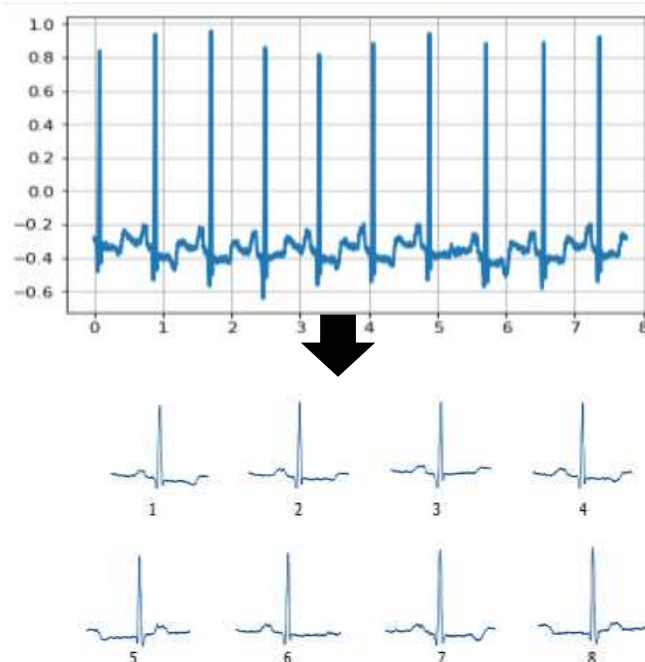
## ECG image

We converted ECG signals into ECG images because a two-dimensional CNN requires an image as input data. We then plotted each ECG beat as an individual 200 x 200 grayscale image. In the MIT-BIH arrhythmia database, every ECG beat is divided based on Q-wave peak time. More specifically, the type of arrhythmia is considered at the Q-wave peak time of each ECG beat. Thus, we defined a single ECG beat image by positioning the Q-

wave peak signal while eliminating the first and the last 10 ECG signals from the Q-wave peak signals. Based on the time information, a single ECG beat range can be defined with the following:

$$T(Q_{\text{peak}}(n-1) + 10) \leq T(n) \leq T(Q_{\text{peak}}(n+1) - 10)$$

For example, for a signal with 10 beats, 8 ECG beat segments would be converted to images.



**Fig. 3:** *Plotting Each ECG Beat as an Individual 200×200 Scale Image.*

We converted ECG signals into ECG images by plotting each ECG beat. We used the Biosppy module of Python for detecting the R – peak in the ECG signals. After the R-peaks were found, we took the present R-peak and the last R-peak, took half of the distance between the two, and included those signals in the present beat. Using this technique, we segmented R-peaks to a beat. We did this step for the next beat. We used Matplotlib and OpenCV to convert these segmented signals into grayscale images. Figure 3 shows the segmented signals.

### Feature Extraction

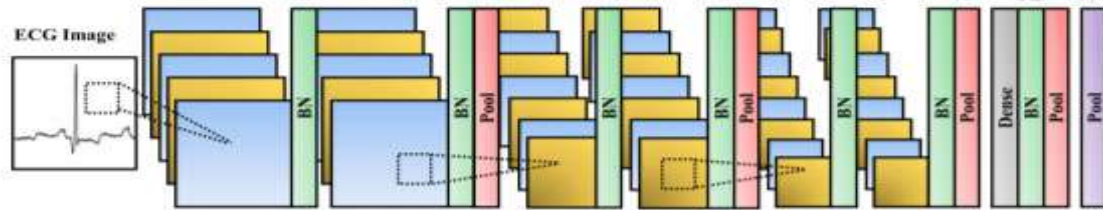
Convolutional neural networks were first developed by Fukushima in 1980 and were improved in later years [15]. It is a form of DNN which involves one or more convolutional layers followed by one or more fully connected layers as in a standard multilayer neural network [15]. The main advantages of CNNs are that they are easier to train and have fewer parameters than fully connected networks with the same number of hidden layers [15]. CNNs are self-learned and self-

organized networks which remove necessities of supervision. Nowadays, image classification, object recognition, and handwriting recognition are important concentrations of CNN. In addition, they play an important role in the medical field for automated disease diagnosis [16]. CNN does not need prerequisites such as pre-processing of datasets and separate feature extraction techniques, but some machine learning algorithms do. This makes CNN advantageous and reduces liability during training and picking the best feature extraction procedure for the automatic detection of arrhythmias [15,16]. We used a kernel size of  $3 \times 3$  for all the convolutional layers, then we proceeded to Batch-normalization and ReLU activation. After the spectrogram conversion, the convolutional layers were arranged into 6 Convolutional Blocks in which each block had four layers. The number of filters was initially set to 32 for the first three convolutional layers but increased by 32 in the last layer of each convolutional block and this last layer also applied stride 2 while all other layers kept a stride of 1[13]. We reduced the size of the output image



after each block by using stride 2 for the last layer in each block. We used an ECG image with 200 X 200 grayscale image. This resulted in a 200 x 200 x 1 input

dimension of the network. The Convolutional neural network at the output of the last Block provided for the feature aggregation.



**Fig. 5:** Convolutional Neural Network of our Proposed Network.

### Activation function

The role of an activation function is to define the output value of kernel weights in the model. In modern CNN models, nonlinear activation is widely used, including rectified linear units (ReLU), leakage rectified linear units (LReLU) [17], and exponential linear units (ELU) [18]. While ReLU is the most widely used activation function in CNN, a small negative value is generated by LReLU and ELU because the ReLU translates whole negative values to zero. This results in the dropping of participation of some nodes in learning. We used ELU after the experiment as the performance for ECG arrhythmia classification was better than LReLU. ReLU, LReLU, and ELU are shown in the following [16]:

$$\text{ReLU}(x) = \max(0, x)$$

$$\text{LReLU}(x) = \max(0, x) + \alpha \min(0, x)$$

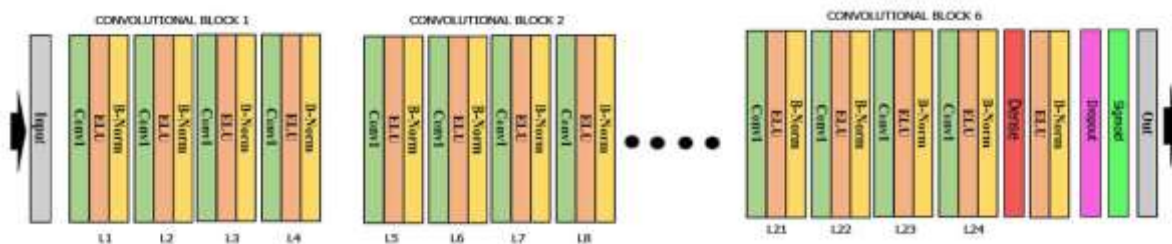
$$\text{ELU}(x) = \begin{cases} x & \text{if } x \geq 0 \\ \gamma(\exp(x) - 1) & \text{if } x < 0 \end{cases}$$

### Aggregation of features across

While feature selection removes characteristics from the input file, feature aggregation combines input features into a smaller set of features called aggregated features. Variable length outputs are produced when the Convolutional Blocks process the variable length input of ECG signals in full length. These variable length outputs need to be gathered across time before they are fed to a standard classifier, which typically needs the dimension of the input to be unchanging. Averaging can be used to attain temporal aggregation in our CNN architecture.

### DATA SET

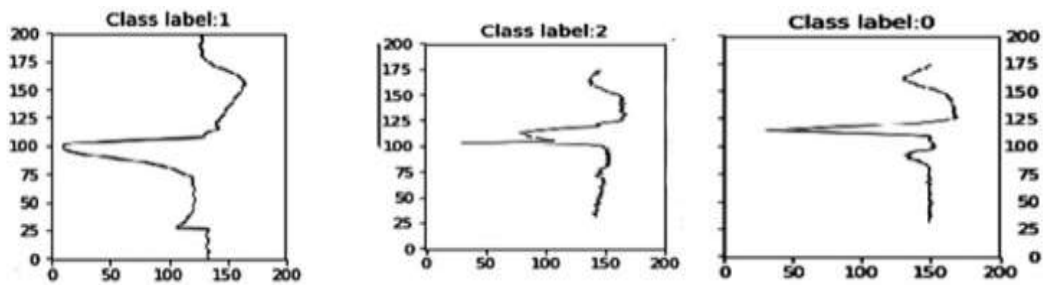
The ECG arrhythmia recordings were retrieved from the MIT-BIH arrhythmia database. The database holds 8528 single lead ECG recordings of length varying from 9 to 61. The ECG recording is sampled at 360 samples per second. The MIT-BIH database contains approximately 110,000 ECG beats with 15 different types of arrhythmia including normal.



**Fig. 6:** Architecture of Proposed CNN Model.

The aim of this paper is to validate the performance of the proposed CNN. From the MIT-BIH database, each record was labelled as normal beat (NOR), AF rhythm, other rhythm, and noise record. For our network architectures we used the cross-entropy loss (reweighted as to account for the class frequencies) as a training objective and employed the Adam

optimizer with the default parameters recommended in [19]. The batch size was set to 64. We used 7177 Normal beat ECG Images (class 0), 8917 Paced rhythm ECG images (class 2) and 472 Other rhythm ECG images (class 1). In total, we used 16566 images as a data set before using data augmentation and K fold cross validation.



**Fig. 7:** Spectrogram of a Sample Data Instance Belonging to Each Class.

### Data augmentation

The poor generalization performance of a model is a result of overfitting, which occurs due to training on too few examples. Infinite training data can eradicate overfitting as every possible instance can be considered. Obtaining new training data is not easy in most machine learning applications, especially in image classification tasks, thereby limiting us to the training set at hand. We can, however, generate more training data through data augmentation, which enhances the training data by randomly transforming the existing data by generating new examples. Therefore, overfitting is reduced through the artificial boosting of the size of the training set.

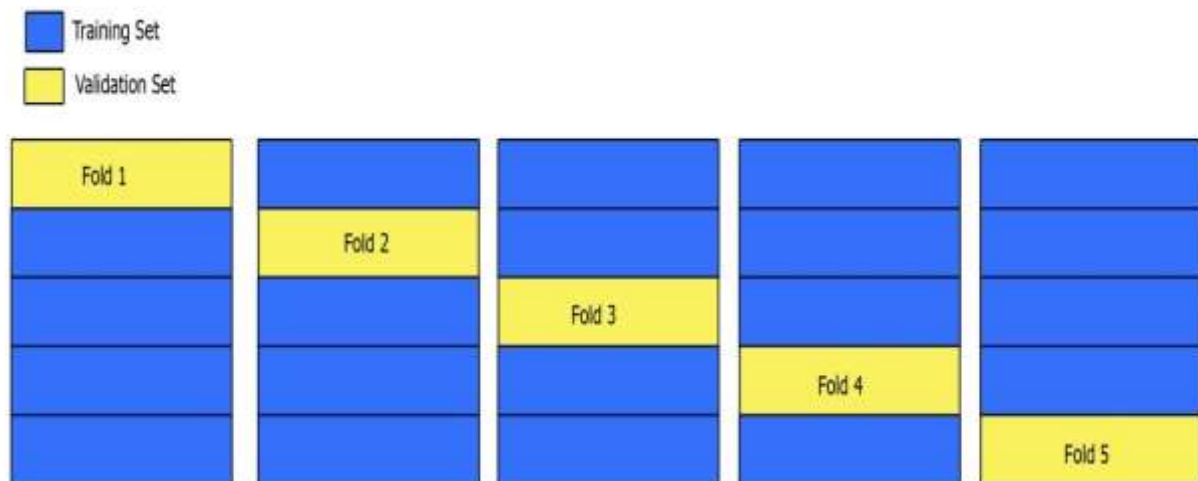
It was demonstrated in [20] that data augmentation can regularize and prevent overfitting in neural networks and improve classification performance in problems with imbalanced class frequencies [21].

In our dataset the third class (Other rhythm ECG images) are very few compared with the other two classes, so we used data augmentation to increase the number of

data sets for this class to 7740 images. Therefore, we augmented Other rhythm ECG images with nine different cropping methods: left top, centre top, right top, centre left, centre, centre right, left bottom, centre bottom, and right bottom. Each cropping method results in the size of an ECG image, that is 128×128 grayscale. These augmented images are then resized to the original size, which is 200×200.

### Training and evaluation

After data augmentation K fold cross validation, the proposed CNN algorithms used 953360 ECG beat images for training and 238340 ECG beat images for validation. Furthermore, 5056 ECG image were used for testing. We trained the CNN end-to-end from scratch without encountering any issues. Training the convolutional layers in the CNN from scratch, on the other hand, did not lead to convergence. We therefore used feature averaging across time and the convolutional layers, which were trained together with a linear classifier for 150 epochs. We also used K fold cross validation to overcome overfitting.



**Fig. 8:** *K – Fold Cross Validation.*

### Testing of data

The algorithm does test on the CNN model to give test accuracy after completion of each training epoch. Our CNN algorithms used 150 epochs for the test data set. After completion of every epoch, we used 20% of the data as validation part to improve accuracy. Twenty percent of the total training data (70% of the original dataset) was used as the validation part and was used to improve accuracy.

### RESULTS

Our research further shows the important role of CNN in extracting all the dissimilar features, which are comparatively invariant to local spectral and temporal variations. This has resulted in higher accuracy performance. The proposed CNN algorithm contains three stages: (1) data pre-processing of input, where ECG signals are processed so that the computer can understand different diseases, (2) stacking of convolution layers to extract the features, and (3) layering of a fully connected layer and activation of the sigmoid function, which will predict the disease.

Table 1 shows the parameters of the CNN layers and their filter size and output size. The proposed CNN algorithm was used to classify between Normal sinus rhythm

(class 0), Paced rhythm (class 2), and Other rhythm (class 1). We used 24 hidden layers. The ReLU function was used to activate each hidden layer and batch normalization was used to normalize the input layer by adjusting and scaling the activations. After the convolutional layers, the resulting outputs were passed to reshape them. At the output of the layer, a linear activation function was then implemented.

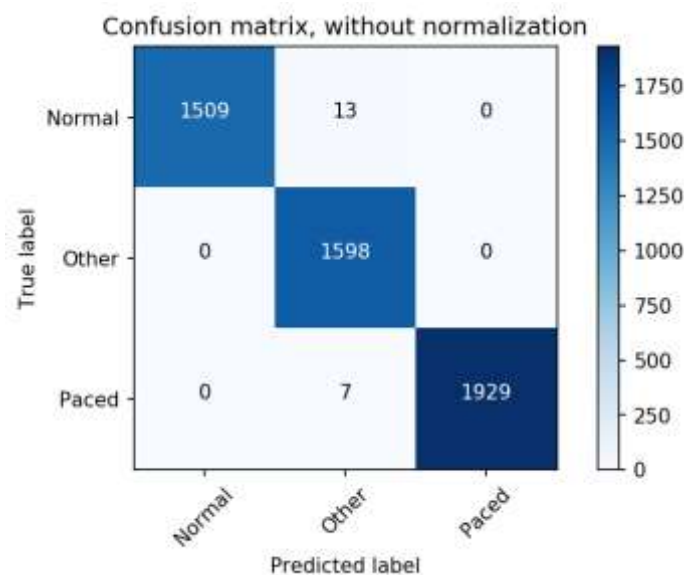
The network was trained with 150 epochs and 50 steps per epoch. It gave an accuracy over 90% for the MITBIH arrhythmia database. Figure 10 shows the confusion matrix for the validation part of the dataset. The confusion graph is a graph which plots the true label versus the predicted label. As shown in the graph, the blue square indicates the high number of correct responses and the white square indicates the low number of incorrect responses. The dataset contains a total of 23834 ECG recordings. 7177 are Normal sinus rhythm (Class 0), 7740 are Paced rhythm ECG (Class 2), and 8917 are Other rhythm (Class 1). After K-fold cross validation, we used 80% of the data for training, which is 953360 ECG signal images and 20% of the data for validation, which is 238340 ECG signal images.

**Table 1: Architecture of Proposed CNN Model.**

Layers	Type	Kemel Size	Stride	#Kernel	Output size
Layer 1	Conv2D	3 x 3	1	32	200 x 200 x 32
Layer 2	Conv2D	3 x 3	1	32	200 x 200 x 32
Layer 3	Conv2D	3 x 3	1	32	200 x 200 x 32
Layer 4	Conv2D	3 x 3	2	64	100 x 100 x 64
Layer 5	Conv2D	3 x 3	1	64	100 x 100 x 64
Layer 6	Conv2D	3 x 3	1	64	100 x 100 x 64
Layer 7	Conv2D	3 x 3	1	64	100 x 100 x 64
Layer 8	Conv2D	3 x 3	2	96	50 x 50 x 96
Layer 9	Conv2D	3 x 3	1	96	50 x 50 x 96
Layer 10	Conv2D	3 x 3	1	96	50 x 50 x 96
Layer 11	Conv2D	3 x 3	1	96	50 x 50 x 96
Layer 12	Conv2D	3 x 3	2	128	25 x 25 x 128
Layer 13	Conv2D	3 x 3	1	128	25 x 25 x 128
Layer 14	Conv2D	3 x 3	1	128	25 x 25 x 128
Layer 15	Conv2D	3 x 3	1	128	25 x 25 x 128
Layer 16	Conv2D	3 x 3	2	160	13 x 13 x 160
Layer 17	Conv2D	3 x 3	1	160	13 x 13 x 160
Layer 18	Conv2D	3 x 3	1	160	13 x 13 x 160
Layer 19	Conv2D	3 x 3	1	160	13 x 13 x 160
Layer 20	Conv2D	3 x 3	2	192	7 x 7 x 192
Layer 21	Conv2D	3 x 3	1	192	7 x 7 x 192
Layer 22	Conv2D	3 x 3	1	192	7 x 7 x 192
Layer 23	Conv2D	3 x 3	1	192	7 x 7 x 192
Layer 24	Conv2D	3 x 3	2	224	4 x 4 x 224
Layer 25	Full			1024	3
Layer 26	Out				3

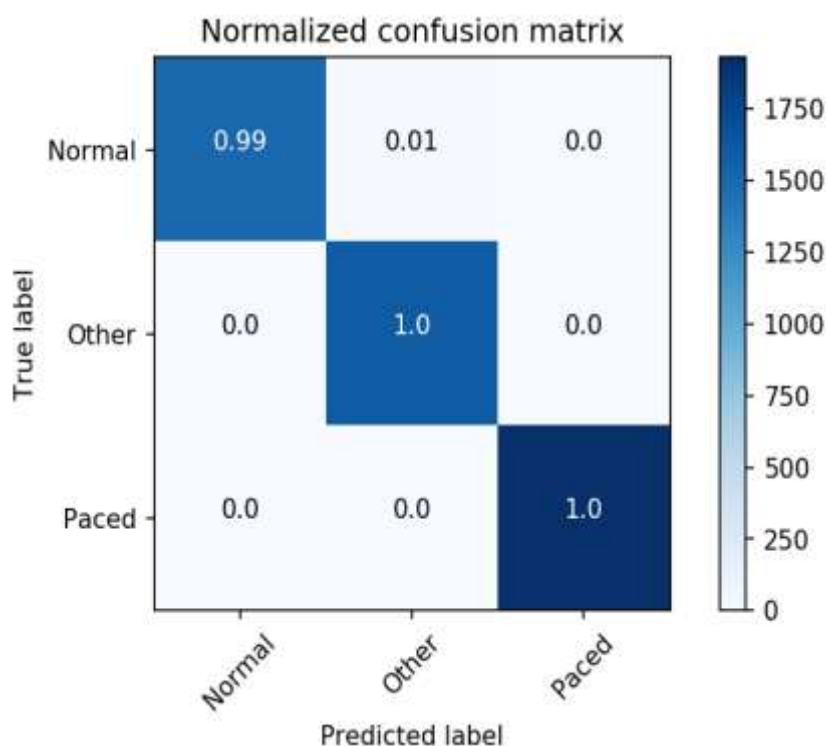
From the Validation data 5056, 1509 Normal sinus rhythm, 1929 Paced rhythm ECG and 1598 other rhythm signals were successfully classified by the algorithm, an improvement in the accuracy of the CNN model. Figure 10 shows a graphic

representation of the confusion matrix for the CNN algorithm. The network provides a reasonable prediction accuracy for the diseases. We expect a reasonable confusion because of unbalanced classes in the data set.



(a)



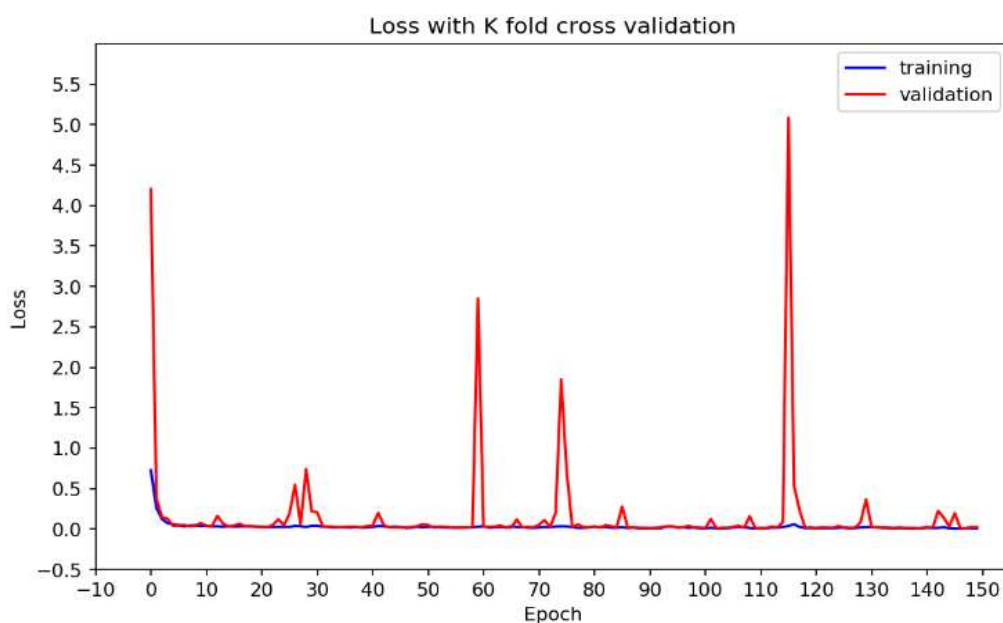


(b)

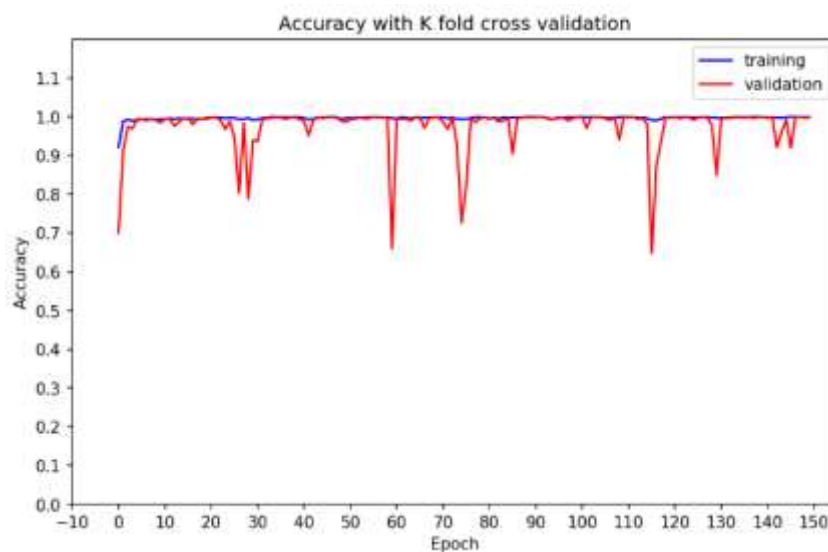
**Fig. 10:** Confusion Matrix (a) with Normalization and (b) Without Normalization of the CNN Algorithm.

Figure 11 shows that the model converges very quickly and presents over 90% accuracy for the validation set. The noticeable peaks in the validation accuracy are most likely due to the unbalanced

classes in the data set. This effect might be reduced by adding weight factors to the loss function, which would penalize those weights that belong to higher-frequency classes [14].



(a)



(b)

**Fig. 11:** (a) Train and Validation Loss Graph (b) Train and Validation Accuracy Graph.

## CONCLUSION

We proposed ECG arrhythmia classification technique using CNN with ECG images as inputs. 200 x 200 grayscale images were converted from a PhysioNet/CinC Challenge 2017 dataset ECG recording. 238340 ECG beat images were attained with three types of ECG beats including Normal sinus rhythm (Class 0), Paced rhythm ECG (Class 2), and Other rhythm (Class 1). An enhanced CNN model was created with significant concepts such as data augmentation, regularization, and  $K$ -fold cross-validation. The proposed algorithms resulted in successful classification of disease states in each signal with significant accuracy, using CNN models (Table 1). As a result, the proposed algorithms can achieve efficient diagnoses of various cardiovascular diseases with the accuracy of over 90%. The results show that detection of arrhythmia with ECG spectrograms and CNN models can be an important method to help the experts analyze cardiovascular diseases using ECG signals. Furthermore, the proposed ECG arrhythmia classification method can be applied to medical robots or scanners that can monitor the ECG signals and help

medical experts identify ECG arrhythmia more precisely and easily.

## REFERENCES

1. World Health Organization (2017). Cardiovascular disease (CVDs). <http://www.who.int/mediacentre/factsheets/fs317/en/>. Accessed 18 Apr 2018.
2. Melo S.L., Caloba L.P., Nadal J. *Arrhythmia analysis using artificial neural network and decimated electrocardiographic data*. In Proceedings of the IEEE Conference on Computers in Cardiology, Cambridge, MA, USA, 24–27 September 2000; IEEE: Piscataway, NJ, USA, 2000. 73–76p.
3. Moody G.B., Mark R.G. *The impact of the MIT-BIH arrhythmia database*. IEEE Eng. Med. Biol. Mag. 2001, 20, 45–50p.
4. Salam A.K., Srilakshmi G. *An algorithm for ECG analysis of arrhythmia detection*. In Proceedings of the IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 5–7 March 2015; IEEE: Piscataway, NJ, USA, 2015. 1–6p.

5. Debbal S.M. *Model of differentiation between normal and abnormal heart sounds in using the discrete wavelet transform*. J. Med. Bioeng. 2014, 3, 5–11p.
6. Perez R.R., Marques A., Mohammadi F. *The application of supervised learning through feed-forward neural networks for ECG signal classification*. In Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Vancouver, BC, Canada, 15–18 May 2016; IEEE: Piscataway, NJ, USA, 2016, 1–4p.
7. Palreddy S., Tompkins W.J., Hu Y.H. *Customization of ECG beat classifiers developed using SOM and LVQ*. In Proceedings of the IEEE 17th Annual Conference on Engineering in Medicine and Biology Society, Montreal, QC, Canada, 20–23 September 1995; IEEE: Piscataway, NJ, USA, 1995, 813–814p.
8. Elsayad A.M. *Classification of ECG arrhythmia using learning vector quantization neural networks*. In Proceedings of the 2009 International Conference on Computer Engineering & Systems, Cairo, Egypt, 14–16 December 2009; IEEE: Piscataway, NJ, USA, 2009, 139–144p.
9. Gautam M.K., Giri V.K. *A neural network approach and wavelet analysis for ECG classification*. In Proceedings of the 2016 IEEE International Conference on Engineering and Technology (ICETECH), Coimbatore, India, 17–18 March 2016; IEEE: Piscataway, NJ, USA, 2016, 1136–1141p.
10. 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, 679p.
11. Acharya U.R., Fujita H., Lih O.S., et al. *Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network*. Knowl. Based Syst. 2017, 132, 62–71p.
12. Nilanon T., Yao J., Hao J. et al. *Normal/abnormal heart recordings classification by using convolutional neural network*. In Proceedings of the IEEE Conference on Computing in Cardiology Conference (CinC), Vancouver, BC, Canada, 11–14 September 2016; IEEE: Piscataway, NJ, USA, 2016, 585–588p.
13. Convolutional Recurrent Neural Networks for Electrocardiogram Classification. Martin Zihlmann, Dmytro Perekrstenko, Michael Tschannen, Dept. IT & EE, ETH Zurich, Switzerland.
14. A convolutional neural network to detect atrial fibrillation from a single-lead ECG. <https://github.com/awerdich/physionet>
15. Shadi G., Mostafa A., Nasimalsadat M., et al. *Atrial fibrillation detection using feature-based algorithm and deep convolutional neural network*. In Proceedings of the Conference on Computing in Cardiology (CinC), Rennes, France, 24–27 September 2017; IEEE: Piscataway, NJ, USA, 2017.
16. Savalia S., Emamian V. *Cardiac Arrhythmia Classification by Multi-Layer Perceptron and Convolution Neural Network*.
17. Maas A.L., Hannun A.Y., Ng A.Y. *Rectifier nonlinearities improve neural network acoustic models*. International Conference on Machine Learning. 2013, 30(1), 3p.
18. Clevert D.A., Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (elus). 2015. arXiv preprint arXiv:1511.07289
19. Kingma D.P., Adam B.J. *A method for stochastic optimization*. In Proc. Int. Conf. on Learn. Representations (ICLR). 2015.

20. Simard P., Steinkraus D., Platt J. *Best practices for convolutional neural networks applied to visual document analysis*. In Proc. Int. Conf. on Document Analysis and Recognition. 2013.
21. Chawla N., Bowyer K., Hall L., Kegelmeyer W. *Smote: Synthetic minority over-sampling technique*. Journal of Artificial Intelligence Research. 2002.
22. Chow G.V., Marine J.E., Fleg J.L. *Epidemiology of arrhythmias and conduction disorders in older adults*. Clin. Geriatr. Med. 2012, 28, 539–553p.
23. MIT-BIH Arrhythmia Database. <https://www.physionet.org/physiobank/database/mitdb/>