



## A Novel Approach for ECG Classification Using Probability Continuous Wavelet Transform and Alexnet-Deep Neural Network

Mays D. AL-Bayati<sup>1\*</sup>Duraid Y. Mohammed<sup>1</sup>Mohammad Sarfraz<sup>2</sup><sup>1</sup>*Computer Science Department, Aliraqia University, Baghdad, Iraq*<sup>2</sup>*Electrical Engineering Department, Aligarh Muslim University, Aligarh, India*\* Corresponding author's Email: [1maysalbayati6@gmail.com](mailto:1maysalbayati6@gmail.com)


---

**Abstract:** Arrhythmia identification using electrocardiograms (ECGs) is a critical component of biomedical signal processing and pattern recognition. The classification of ECG arrhythmias is presented in this article using features extracted from the Extreme Value distributor of the probability density function (PDF), also training model built using the Alexnet architecture of a conventional neural network (CNN). A continuous wavelet transform (CWT) was used to convert the features space to a two- dimensional image. Additionally, a comparison of the two approaches was conducted. In the first case, we used CWT to convert ECG data to two-dimensional images. In the second case, features were extracted from PDF files and converted to the image domain using a CWT. With a learning rate of 0.001 and batch size of 250, the accuracy of tests was 98.67 %. The results indicated that the proposed method is more accurate than the conventional method when it comes to feature extraction.

**Keywords:** Cardiovascular diseases (CVDs), Diagnosis, ECG arrhythmias classification, Alexnet CNN.

---

### 1. Introduction

In recent decades, many methods have been used to diagnose cases of electrocardiography (ECG), which allows the observation and recording of electrical activities during heart operations using leads located around the heart. Doctors can identify persons with cases of cardiac arrhythmia by observing the ECG records. In addition, the diagnosis of heart activities based on a feature from a signal helps the doctors significantly [1]. Lately, deep neural networks get high interest in biomedical signal processing implementation and applications for analysis. The concept of deep learning is merges ideas from neuroscience and environmental, statistics, arithmetic, and physics, to make computers learn of given data to be smart enough to take an action (s). One of deep learning algorithm is convolutional neural network (CNN) and can applying an ECG data as a 1D vector to a CNN without preprocessing or feature selection. Also, the input of CNN designed to process an image data, for this case, our proposed

work prepared data as 2D to each high accuracy. In [2], two scenarios were designed for comparing different dataset (set1=7 classes, set2=11) using DNN model, the accuracy belong to the first set (7 classes) is 92.24 % and accuracy belonging to the second set (11 classes) is 96.13 %, comparing to our work, the method proposed in our work during a feature normalization is based on probability density function (PDF) for and then apply continuous wavelet transform (CWT) to convert feature from 1D to image spectrogram domain, the accuracy is 98.67. In [3], method block-based neural networks have been applied and optimized, working for ECG heartbeat cases pattern classification, and compared to our proposed work, a CNN configured with different sets of batch size and learning rate. 1-D CNN method has been applied for classification ECG classes in [4], compared to our work, 1D feature has been converted to 2D image spectrogram based on CWT. In [5], a new SoftMax layer has been added over the hidden layer, which is deep neural network method, where the labelling for classes of ECG in the

test record during the iteration and applying this data to modifying the DNN weights, and the other 9-layers of CNN has been justified to classify 1-D ECG arrhythmias in [6], by setting number of layers within CNN, a training processing take more time and this compared to our work by set the size of CNN layer to 3 which has less time of processing during training in [7], they evaluate ECG classification by combining both RNN and CNN where the number of classes of ECG is 5, they proposed ECG classification by combined both RNN and CNN where the number of classes of ECG is 5. This work focus on one cardiovascular disease, arrhythmia, where the heart signs diseases were complex and varied. In [8] RBF in NN parameters had optimized based on cuckoo search algorithm after converting 1D ECG signal to 2D spectrogram based on CWT, and the dataset of 12100 recorded beats, the implemented algorithms accomplished of distinguishing of 2 classes which are normal and abnormal beats where the accuracy is 98.32 and compared to our work the number of classes is (3) and accuracy is 98.67. In [9] proposed noise removal by applying a wavelet algorithm followed by transforming 1D to 2D by using CWT, where the accuracy reached 97.38 % and 99.02 % respectively for 1D and 2D, in this work, many stages of processing consuming resources such as processor and memory, in our work, PDF function has basic math operations and to consuming high processing and space in memory. In [10], Two approaches were used in this study. The first approach is based on ECG data in the time-series domain and the spectrogram (2D) image, and the number of layers set within CNN is 15 layers where this consuming processor and memory resources during training. In [11] a new technique for detecting seizures in epileptic based on ECG signal. The features extract from the ECG signal are spectral analysis means of absolute deviation of fast fourier transform coefficients and spectral entropy. The accuracy of this work is 94.2 %, this work focus on epilepsy case and compared to our work; three cases of ECG taken and the result is 98.67 % in [12] noise removal in ECG signal using independent component analysis (ICA) with different capabilities comparisons had been studied in this paper. The metric used is signal to noise ratio (SNR), and this work has only focus on comparisons methods of noise removal of ECG signal, where our work has classification of ECG cases.

The proposed work focuses on increasing the performance during the classification of the ECG signals. The data has been obtained by the MIT-BIH arrhythmia database based on using PDF as a new feature of ECG data and which this function of statistics operation produced accurate in prediction of

ECG signal (variable and continuous data), and this method is applied for all records in the dataset. The feature of PDF is a 1D vector, and for preparation as input to CNN, the data may be a 2D vector. In this case, the CWT is applied as transformer data from 1-D to 2-D, and the result is illustrated as a spectrogram image. The last phase of work is postprocessing, where the training and validation procedure is applied, and one of the deep learning techniques is used called CNN with pre-trained architecture named Alexnet. Finally, the confusion matrix is used to measure the performance of the conventional and proposed methods.

## 2. ECG dataset

The MIT-BIH database of ECG recorded signals from PhysioNet has depended on many approaches involving analyzing ECG signals [13]. The structure of the database contains two files, the first file represents the data of ECG, and the second one contains the list of labels. The total number of samples is 192 for three classes; the first class ARR has 96 samples, CHF 30 samples, and NSR 35 samples. Each signal sample has a length of data 65536. In this work, 30 records are taken from each class with full length, which is 65536. ARR is an irregular sequence of heartbeats, may beat is fast, or slow, CHF is occurring if the heart muscle not pumping blood normally, and NSR is a rhythm that defines the rhythm of the healthy heart of a human.

## 3. Feature extraction based on PDF

In this work, the PDF method is applied during the preprocessing of the ECG signal. The PDF is considered as a statistical method that describes the relative possibility for the vector to take on a given value. The PDF is typically associated with a continuous invariant vector to fall under a specific region is found by the integral of this variable's density over the region [14].

## 4. Extreme value distribution function

The extreme value function comprises parameters, model description, and sample data for a distribution probability of extreme value. The function of the extreme value distribution is used to model the smallest or largest value amongst a part over-sized number of random values that make measurements or observations that are similarly distributed. The distribution of extreme values is suited for modeling the minimum values from a distribution where tails descends exponentially fast, working as the normal distribution. Besides, this function considers negative

values of the original values despite the maximum distribution value, such as regular or exponential distribution functions [15]. The equation below describes a probability density function for the extreme value distribution with the scale parameter  $\sigma$  and location parameter  $\mu$ .

$$f(x|\mu,\sigma) = \sigma^{-1} \exp\left(\frac{x-\mu}{\sigma}\right) \exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right) \quad (1)$$

Where  $x$  is data,  $\mu$  is the mean of  $x$ , and  $\sigma$  is the standard deviation of  $x$ .

### 5. 1-D vector to 2-D using CWT

Continues wavelet transform (CWT) is used as diagnostic wavelets. CWT take action arithmetically on all scales [21]. In case of high domain of data the scale of the low frequency noises show that the strong correlation. In case of the low domain of data of the scale of the high frequency noises show that the strong correlation.

Complex-valued time/frequency contained filters with threshold support on negative values appear in the frequencies domain and provide the basis for a strong analysis of analog signals [16].

$$W_{\Psi}(t, s) = \int_{-\infty}^{\infty} \frac{1}{s^n} \Psi^* \left( \frac{t-t'}{s} \right) x(t') dt' \quad (2)$$

The CWT set of bandpass operations with the scaling values as normalization  $n=1$ , or as a set of projections with  $n = 1/2$  and has indexed by the scale set  $s$ .

One wavelet type is morse wavelets that effectively join all wavelet types, in addition, the analytic filter, and depict the hard exponentials themselves. These morse wavelets are defined in the frequency domain as

$$W_{\beta,\gamma}(\omega) = \int_{-\infty}^{\infty} \Psi_{\beta,\gamma}(t) e^{-i\omega t} dt = \frac{U(\omega) \alpha_{\beta,\gamma} \omega^{\beta} e^{-\omega\gamma}}{U(\omega) \alpha_{\beta,\gamma} \omega^{\beta} e^{-\omega\gamma}} \quad (3)$$

where  $\alpha, \beta, \gamma$  is a normalization value,  $U(\omega)$  Is the step value function, and  $\beta$  and  $\gamma$  are two values used to control the waveform in the wavelet. Convolutional neural network (CNN)

Many ECG data analysis methodologies have been used for research and development to improve performance. Convolutional neural network (CNN) deep learning is one of these strategies. This includes deep learning algorithms such as those based on convolutional neural networks (CNNs). Signal analysis has gained popularity in the field of signal

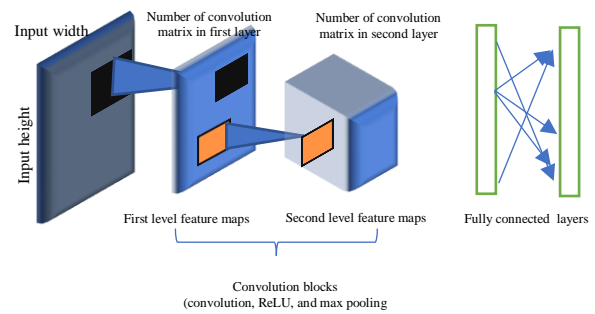


Figure. 1 A typical CNN architecture consisting of an input layer, convolution layer, a max-pooling layer, and a fully connected layer [17]

processing, particularly in the field of electrocardiography. The advantage of CNN is that the input is multidimensional, which means that data in two dimensions may be directly entered into the network. The CNN presented a more comprehensive picture of numerous supervised CNN applications and advised more research into ECG classification techniques [17, 18]. Fig. 1 illustrates a typical CNN architecture.

#### a. Input layer

The input layer is used to input the images dataset two-dimensional vectors.

As explained previously, CNN is composed of multiple layers. The input layer of a CNN should contain image data in the form of a multidimensional matrix. Convolution layer.

The convolutional layer is one of the layers that differentiate the known neural network from the convolutional layer (NN). Each convolutional layer's input is the output of the preceding layer, which is convoluted using several convolution kernels. Convolution kernels are used repeatedly in each sensory field of the entire data set, and the convolution result establishes a feature of the input image. Convolution filters are applied across the width and length of the image by performing multiplication and summation operations between the filter and the input at any point. The convolution kernels are the contents of the convolution layer, consisting of the weight matrix  $w$  and the bias  $b$ . The layer's mathematical expression is as follows:

$$X_{jl} = f \left( \sum_{i \in M_j^{l-1}} x_i^{l-1} k_{ij}^l + b_j^l \right) \quad (4)$$

Where  $l$  represents the layer,  $k$  is the convolution kernel,  $b$  is the bias,  $x_i^l$  is current data in current layer and  $M_j$  represents the feature map.

#### b. Pooling layer

Another main point of CNN is the pooling layer.

The theoretical basis of the pooling layer is that images contain the property of “stationarity”, which means that features that are useful in a certain region are also probably useful for other ones. Rectangular blocks are taken from the convolutional layer and sub-sampled by the pooling to produce a single output. The strength of image variations like distortion, noise, and rotation is enhanced by the pooling. It also reduces the output dimensions and retains the noteworthy features. Pooling has two ways, the average pooling, and the max pooling. The maximum pool formula is [19]:

$$y_{j(m,n)}^{l+1} = \max_{0 \leq r, k} \{x_{(m.s+r, n.s+k)}^l\} \quad (5)$$

where  $m \geq 0, n \geq 0, s \geq 0$ , and  $y^{l+1}$  is the value of the neuron unit (m, n) that in the  $j^{\text{th}}$  output feature map  $y^{l+1}$  at the  $l+1$  layer,  $(m.s+r, n.s+k)$  is neuron unit in the  $i$ th input map  $x^l$  at the  $l+1$  layer, whose corresponding value is  $x^l$ ,  $y^{l+1}$  is obtained by computing the biggest value over  $(m.s+r, n.s+k)$  an  $s \times s$  non-overlapping local region in the input map  $x^l$ .

#### c. ReLU layer (rectified linear unit)

‘ReLU’ is a non-linear operation, and it involves units that employ the rectifier. An element-wise operation is implemented per pixel, and zero reconstitutes all negative values in the feature map. If positive, it maintains its value without any change. It is presumed that there is a neuron input given as  $x$  in an attempt to comprehend how the ReLU functions, and from that, the rectifier is represented as:

$$f(x) = \max(0, x) \quad (6)$$

One gets a better CNN along with stable performance through using those optimization techniques. A lot of auxiliary data can be introduced to pre-train the CNN to improve the generalization ability of the system [20].

#### d. Fully connected layer

The term (FCL), fully connected layer means that each filter in the preceding layer is linked to every filter in the following one. The results from the convolution layers, ReLU, and max-pooling represent the input image’s high features. The purpose of using the FCL is to use these features to identify the input image into separate categories depending on the training dataset. FCL is considered as the final pooling layer that feeds the features to a classifier that utilizes the activation function of softmax.

#### e. Softmax layer

The softmax classification with powerful non-

linear classification capability is being applied as the ECG classes and characteristics are much more complex, and there is no unified template. Softmax is implemented via a neural layer in the network just before the output layer. The layer of softmax must contain the same neuron number as that of the output layer. Softmax is broadening this idea.

The standard (unit) softmax function  $\sigma: \mathbb{R}^k \rightarrow \mathbb{R}^k$  is defined by the formula

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}} \text{ for } i = 1 \dots K \text{ and } z = (z_1 \dots z_k) \in \mathbb{R}^k \quad (7)$$

#### f. Classification layer

The output layer of the convolution neural network is the classifier layer.

The final layer of the convolutional neural network, the layer containing the answer(s).

### 6. Performance evaluation of classifiers

Accuracy, precision, and recall are statistical measures of the performance of a classification test. All possible outcomes of such a test can be represented by a confusion matrix. Precision, sensitivity (recall), and accuracy are given by the following equations where TP is true positive, TN is true negative, FP is false positive, and FN is false negative:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (8)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (9)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+TF} \quad (10)$$

The true positive (TP) is the count of beats classified correctly for the specified class, true negative (TN) is the count of overall which does not belong to specified class and classified correctly. Similarly, false positive (FP) is the count of beats which are classified belonging to the specified class incorrectly and false negative (FN) is the count of beats classified incorrectly as not a member of the specified class.

### 7. Results and discussion

This work presents robust ECG classification with different viewing and various signal amplitudes and based time. For the determination of representing the performance of the proposed work, experimental results are described below.

1. Description of the preprocessing method applied on ECG data as normalization based on using distribution named Extreme Value of PDF.

probability density extreme value distribution function applied on 1D dataset of ECG is contains (3-classes) and number of records of this dataset is

Figs. 2-7 shows the original ECG data and

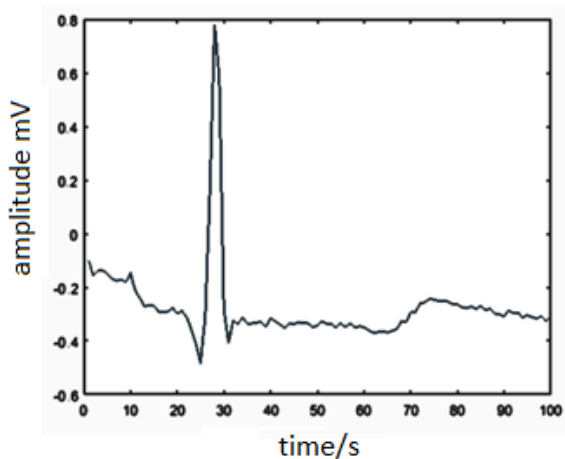


Figure. 2 Original sample of ARR

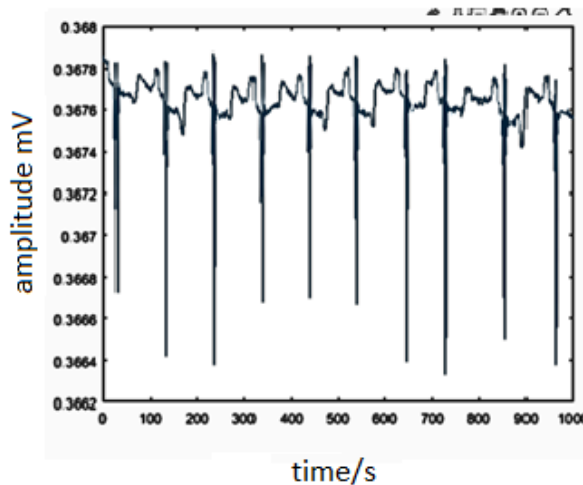


Figure. 3 Extreme value PDF of ARR sample

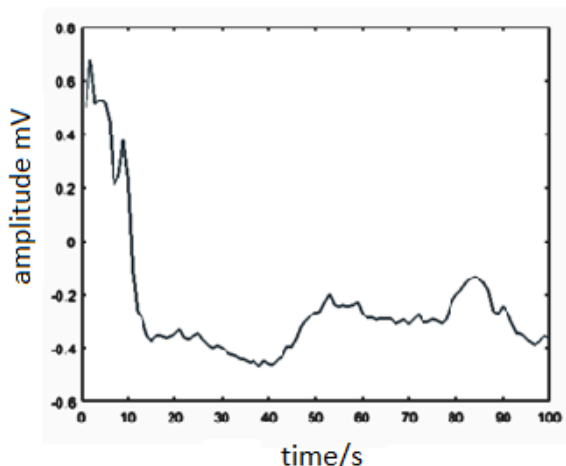


Figure. 4 Original sample of CHF

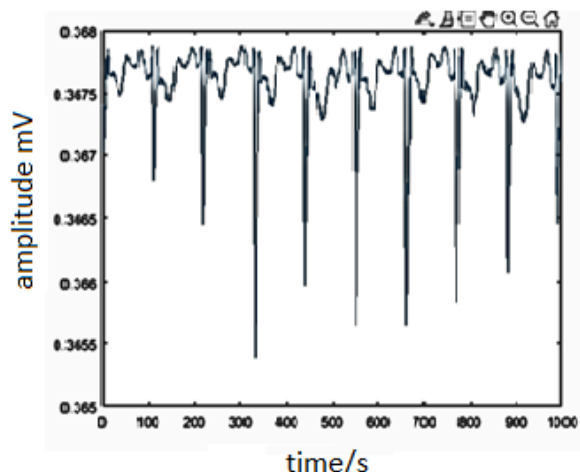


Figure. 5 Extreme value PDF of CHF sample

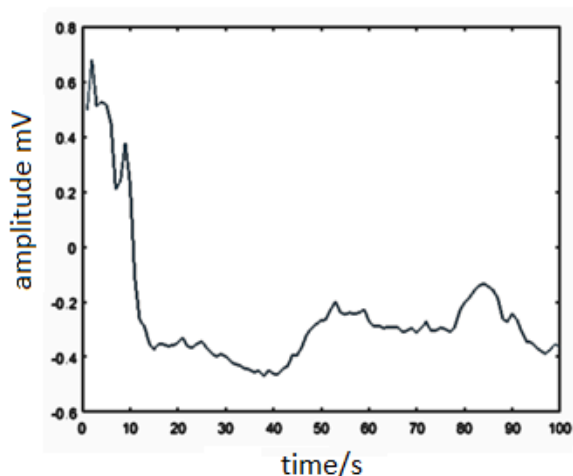


Figure. 6 Original sample of NSR

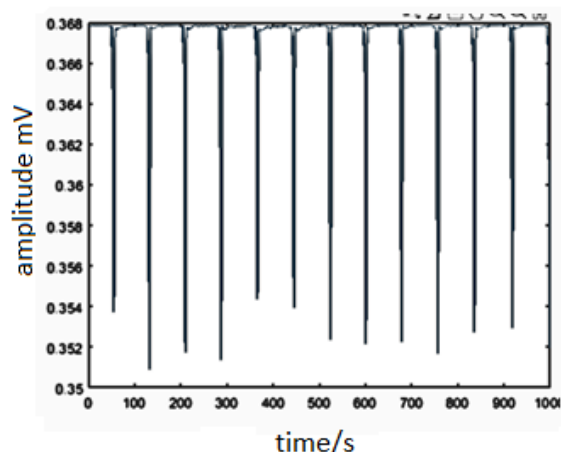


Figure. 7 Extreme value PDF of NSR sample

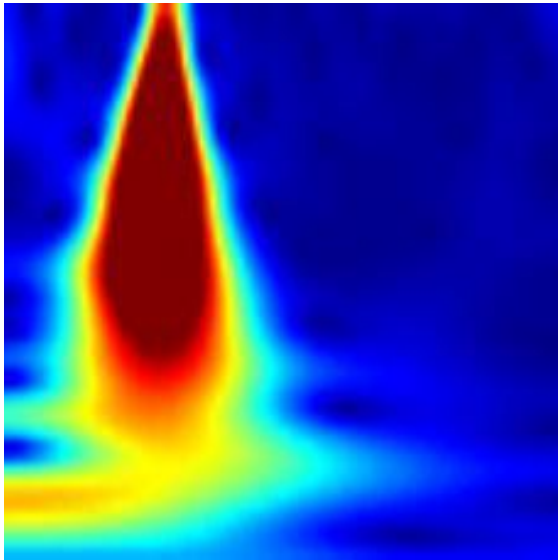


Figure. 8 CWT (ECG signal) sample ARR

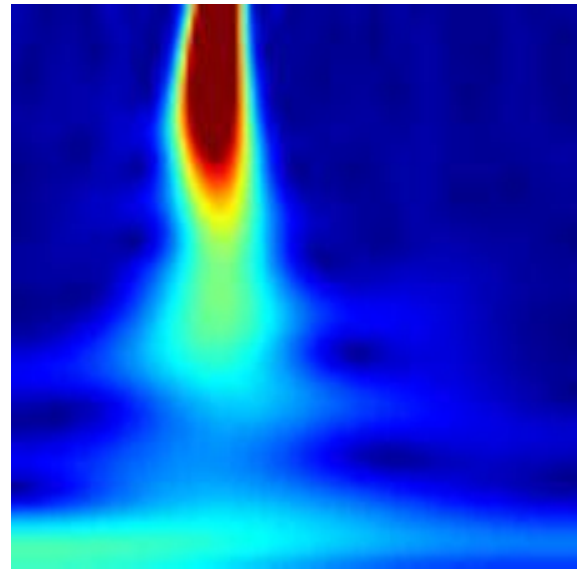


Figure. 9 CWT-PDF of ARR sample

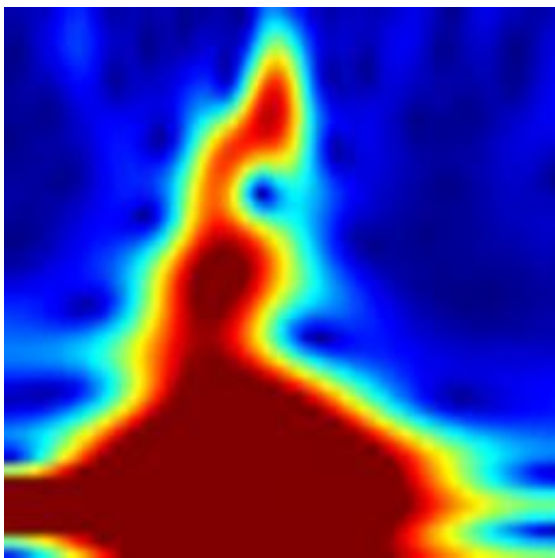


Figure. 10 CWT (ECG signal) sample CHF

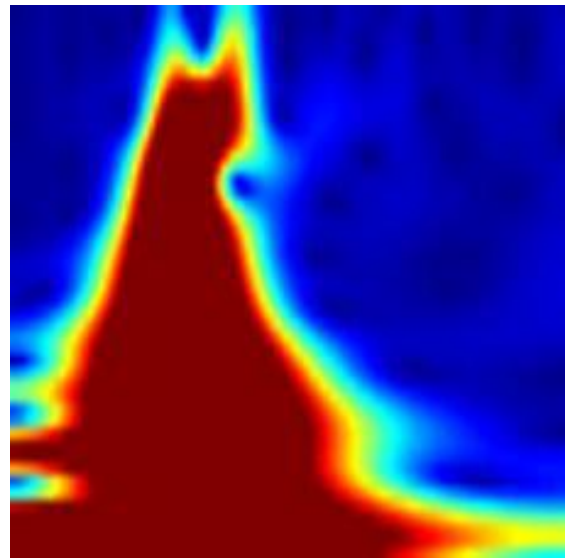


Figure. 11 CWT-PDF of CHF sample

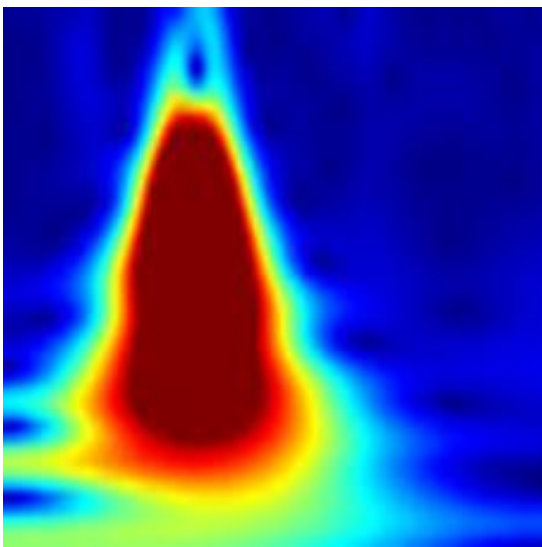


Figure. 12 CWT (ECG signal) sample NSR

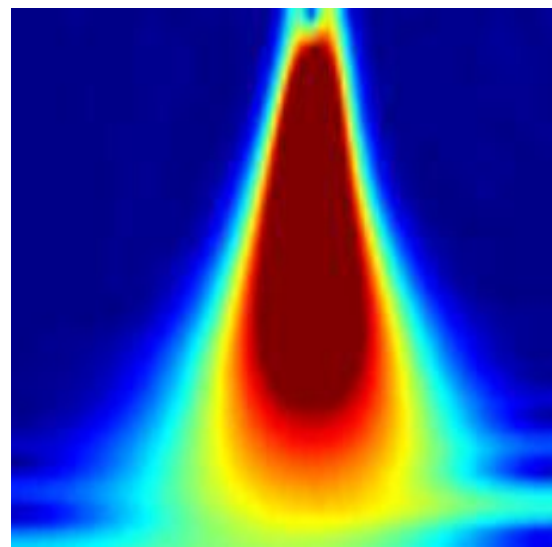


Figure. 13 CWT-PDF of NSR sample

(162), and each record has length (65536). Each class of this dataset labeled as (ARR, CHF, and NSR).

2. The formed input data to the first layer of CNN can be two-dimension vectors. In this step, the processed ECG data is a 1-D vector and the proposed method to transform this vector to a 2-D vector is CWT filter bank. Finally, the spectrogram image result has been resized to (227, 227), which this size suited as size depended on AlexNet (CNN).

The same samples shown during preprocessing have been shown in Figs. 8-13 in the CWT domain.

3. The result of training and classification has been using AlexNet architecture type of CNN for both cases (convention and proposed methods) as shown in Fig. 14.

4. An evaluation in this paper based on test data over the training model, the number of tested data of three classes is (750) samples. During the evaluation phase, a performance measurement selected depends on the confusion matrix (accuracy). Table 1 compares three state-of-the-art with our proposed method and shows that the proposed method has the highest accuracy. For instance, in [4], proposed a method during feature extraction (average, mean, standard

Table 1. Comparison of three approaches with proposed work

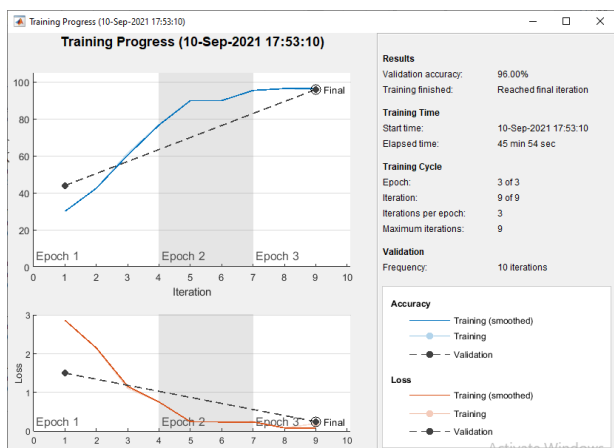
Work	Accuracy %
[4]	96.95
[7]	95.90
[10]time-series	97.10
<b>Proposed work</b>	<b>98.00</b>

Table 2. Evaluation compared with fixed batch size and different learning rate

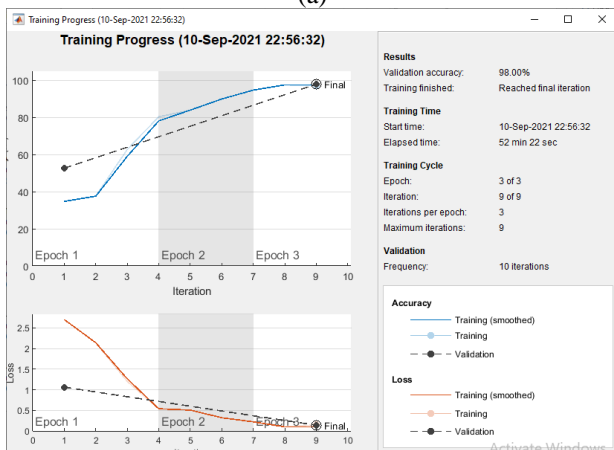
Batch size	Learning rate	Accuracy	Precision	Recall
<b>250</b>	<b>0.001</b>	<b>98.67</b>	<b>98.67</b>	<b>98.67</b>
250	0.0025	98.67	98.67	98.67
250	0.005	98.00	98.00	98.00

Table 3. Evaluation compared with fixed learning rate and different batch size

Batch size	Learning rate	Accuracy	Precision	Recall
<b>250</b>	<b>0.001</b>	<b>98.67</b>	<b>98.67</b>	<b>98.6</b>
200	0.001	98.00	98.00	98.00
150	0.001	98.67	98.67	98.67



(a)



(b)

Figure. 14: (a) Result of CNN training conventional and (b) Result of CNN training proposed method

deviation, energy, and entropy) of ECG based on applied discrete wavelet transform (DWT) and accuracy achieved is 96.95. In [7], 1D data normalized between 0 and 1 of ECG signal proposed and accomplished accuracy in 95.90, and in [10], CWT method used for transformation 1D to 2D spectrogram image and applied to CNN and accomplished accuracy reached 97.10.

Another evaluation of the proposed work is represented in the design of the model of 2D-CNN. Two tested during the second evaluation had experimented. The first experiment was tested by a fixed batch size value and set different learning rate values represented in table 2. The second experiment was tested by a fixed value of learning rate and set different values of batch size as a result represented in table 3.

From the experimental works and results during the evaluations phase, we can have concluded that at batch size 250 and learning rate set to 0.001, the model selected of 2D-CNN achieved high accuracy.

### 8. Conclusion

In this paper, the classification method of ECG data is based on deep learning techniques. Belonging to three different classes (ARR, CHF, and NSR) and these classes were obtained from the MIT-BIH arrhythmia database. The ECG signals data were spliced into records, where each record has a length

(65536). In the procedure of the preprocessing of ECG data has been used PDF (Extreme value) as normalization data and minimizing imbalance data method in this work, other method used to transform the ECG signals from the time domain into two-dimensional time-frequency ECG spectrograms by CWT filter bank, where the wavelet type selected in this filter is “AMOR”. One of the pretrained deep CNN architecture used is the “Alexnet” network has a colour image input size of 227-by-227.

An evaluation of performances has been worked during experimental work. This evaluation is based on measuring the performance by confusion matrix where proposed work compared with three approaches by accuracy matrix, and our proposed work achieved accuracy as (98.67 %). Other evaluations compare the different sets of 2D-CNN parameters, such as learning rate and batch size. We’ve found that the best set of CNN parameters during experimental works are (learning rate=0.001 and batch size=250).

## References

- [1] Ç. Sarvan and N. Özkurt, “ECG Beat Arrhythmia Classification by using 1-D CNN in case of Class Imbalance”, In: *Proc of The Medical Technologies Congress (TIPTEKNO), Izmir, Turkey*, pp. 1–4, 2019.
- [2] O. Yildirim, M. Talo, E. Ciaccio, R. San, and U. Rajendra, “Accurate Deep Neural Network Model to Detect Cardiac Arrhythmia on More Than 10000 Individual Subject ECG Records”, *Elsevier, Computer Methods and Programs in Biomedicine*, Vol. 197, pp. 1-12, 2020.
- [3] W. Jiang and G. Kong, “Block-based neural networks for personalized ECG signal classification”, *IEEE Trans. Neural Networks*, pp. 1750-1761, 2007.
- [4] Ç. Sarvan and N. Özkurt, “Multiwavelet feature sets for ECG beat classification”, *IEEE, In Signal Processing and Communications Applications Conference (SIU)*, pp. 1-4, 2017.
- [5] M. AlRahhala, Y. Bazia, H. AlHichria, N. Alajlana, F. Melganib, and R. Yagerc1, “Deep learning approach for active classification of electrocardiogram signals”, *Information Sciences*, Vol. 345, pp. 340-354, 2016.
- [6] U. Rajendra, S. Oh, Y. Hagiwara, J. Tan, M. Adam, A. Gertych, and R. San, “A deep convolutional neural network model to classify heartbeats”, *Computers in Biology and Medicine*, Vol. 89, pp. 389-396, 2017.
- [7] X. Xu, S. Jeong, and J. Li, “Interpretation of Electrocardiogram (ECG) Rhythm by Combined CNN and BiLSTM”, *IEEE Access*, Vol. 8, pp. 125380–125388, 2020.
- [8] A. Harkat, R. Benzid, and L. Saidi, “Features Extraction and Classification of ECG beats using CWT combined to RBF neural network optimized by Cuckoo search Via Levy Flight”, *4th International Conference on, IEEE*, pp. 1–4, 2015.
- [9] A. Ullah, S. Rehman, S. Tu, R. Mehmood, F. Fawad, and M. E. U. Haq, “A Hybrid Deep CNN Model for Abnormal Arrhythmia Detection Based on Cardiac ECG Signal”, *MDPI, Sensors*, Vol. 21, 2021.
- [10] Y. Şen and N. Özkurt, “ECG Arrhythmia Classification By Using Convolutional Neural Network And Spectrogram”, *IEEE, Innovations in Intelligent Systems and Applications Conference (ASYU)*, Vol. 21, pp. 1-6, 2019.
- [11] G. Shamim, U. Khan, M. Sarfraz, O. Farooq, “Epileptic Seizure Detection Using Heart Rate Variability”, *IEEE International Conference on Signal Processing and Communication (ICSC)*, Vol. 16, pp. 250-254, 2016.
- [12] M. Sarfraz, F. Li, and M. Javed, “A Comparative Study of ICA Algorithms for ECG Signal”, *ACM Press The International Conference*, Vol. 11, pp. 135–138, 2011.
- [13] L. Goldberger, “PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals”, 20 Jun. 2000. [Online]. Available: [physionet.net](http://physionet.net).
- [14] A. Holmes, B. Illowsky, and S. Dean Introductory Business Statistics, *Kansas: 12th Media Services*, 2018.
- [15] C. Mathworks, “Statistics and Machine Learning Toolbox TM User’ s Guide”, 2020. [Online].
- [16] M. Jonathan and C. Sofia, “Generalized Morse Wavelets as a Superfamily of Analytic Wavelets”, *IEEE Transactions on Signal Processing*, Vol. 1, pp. 11-15, 2012.
- [17] T. Mane and V. Kulkarni, “A Survey on Supervised Convolutional Neural Network and Its Major Applications”, *Int. J. Rough Sets Data Anal.*, pp. 71–82, 2017.
- [18] B. G. Tóth, M. Osváth, D. Papp, and G. Szűcs, “Deep Learning for Plant Classification and Content-Based Image”, *Cybernetics and Information Technologies*, Vol. 19, No. 1, pp. 88-100, 2019.
- [19] S. Guo, S. Chen, and L. Yanjie, “Face Recognition Based on ConVolutional Neural Network and Support Vector Machine”, *IEEE Int. Conf. Inf. Autom. IEEE ICIA*, No. August, pp. 1787–1792, 2017.



- [20] M. Coşkun, A. Uçar, Ö. Yıldırı, and Y. Demi, “Face Recognition Based on Convolutional Neural Network”, *IEEE, International Conference on Modern Electrical and Energy Systems (MEES)*, pp. 376–379, 2017.
- [21] O. Rioul and P. Duhamel, “Fast Algorithms for Discrete and Continuous Wavelet Transforms”, *IEEE Trans. on Information Theory*, Vol. 38, No. 2, pp. 569-586, 1992.