Conférence Internationale en Automatique & Traitement de Signal (ATS-2017)
Proceedings of Engineering and Technology – PET
Vol.22 pp.67-72

Artificial Neural Network classifier for heartbeat arrhythmia detection

Hèla LASSOUED #1, Raouf KETATA *2

Physical Engineering and Instrumentation Department,
Energy, Robotics, Control and Optimization Laboratory
National Institute of Applied Sciences and Technology, Tunisia

1 laswed.hela@gmail.com
2 raouf.ketata@insat.rnu.tn

***Abstract*— In this paper, we have presented a diagnostic system for arrhythmia classification using a machine learning approach based on the Artificial Neural Network (ANN). We have selected 44 files of one minute recording from the MIT-BIH arrhythmia database, where 25 files are considered as normal class and 19 files are considered as arrhythmia class. Feature sets were based on ECG morphology (heartbeat intervals and RR-intervals) and features calculated from the Discrete Wavelet Transformer (DWT). Afterwards, we have discussed the appropriate Neural Network structure and the suitable training algorithm in order to properly classify ECG recordings into normal and arrhythmia classes. We have compared then the cascade Forward Network and the Multi-layered Perceptron (MLP) neural network architectures. By only using MLP structure, we have compared two training algorithms, based on backpropagation approach, which are Resilient Backpropagation (RPROP) and Gradient Descent with Momentum (GDM).**

The ANN performance is evaluated in terms of Mean Square Error (MSE) and Accuracy(ACC). The model reached a null MSE and 99% as ACC.

***Keywords*—ECG, ANN, MIT-BIH, MLP, Rpop, GDM.**

I. INTRODUCTION

It is estimated that Cardio Vascular Diseases (CVD) will become the foremost cause of death worldwide [1]. Thus, an early detection of these abnormalities can save the human's life. Cardiologists, instead of relying on their professional experience, need consequently a clinical decision support system (CDSS), especially when the analysis require a carefully inspection of long electrocardiogram (ECG) recordings. Hence, they need CDSS to make correct heart diagnosis as quickly as possible to improve the quality and the speed of medical services.

CDSS can be categorized into two main types. The first type consists of systems with a knowledge base which apply rules to patient data. The second type consists of system without a knowledge base relying on machine learning to analyse clinical data.

This study investigated on the second group approaches which includes arrhythmia heartbeats classification models.

Currently, since the recognition of ECG arrhythmias has become an active research area, several artificial intelligent algorithms have been developed.

These methods include mostly Wavelet coefficient [2], Support Vector Machines [3], Neural Networks [4], fuzzy c-means clustering techniques [5] and many other approaches. These methods would mostly ameliorate the performance of arrhythmia classification systems.

In the same purpose, an Artificial Neural Network (ANN) classifier is presented in this study to classify the ECG recordings into normal and abnormal classes. Indeed, we have discussed mainly the appropriate Neural Network structure and its suitable training algorithm.

This paper is organized as follows. In Section II, the proposed Arrhythmia classification methodology is revisited. Here, we start by introducing MIT-BIH arrhythmia database of one minute ECG recordings. Then we present respectively ECG preprocessing, feature extraction and feature selection stages. In Section III, we detail the ANN classifier. Finally in Section IV, using Mean Square Error (MSE) and Accuracy (ACC) performances, two comparative studies are discussed. Primarily, we have compared the cascade Forward neural network and the Multi-layered Perceptron (MLP) neural network architectures. Then, by only using MLP topology, a brief comparison between two training algorithms, based on backpropagation approach (Resilient Backpropagation (RPROP) and Gradient descent with momentum (GDM)), is done.

II. MATERIAL AND METHODS

The block diagram of the arrhythmia classification methodology is described in figure1. It consists of two main parts which are ECG pre-processing and Neural Network arrhythmia classification.

First of all, the input signals from the MIT-BIH database are presented for pre-processing part. At this stage, ECG artefacts are removed and a set of feature which best characterize the original signal is extracted. Then, in order to reduce the feature vector size, we have applied a feature

ISSN 2356-5608

selection approach. Finally, the processed signals are classified into normal and abnormal heartbeats by an ANN algorithm.

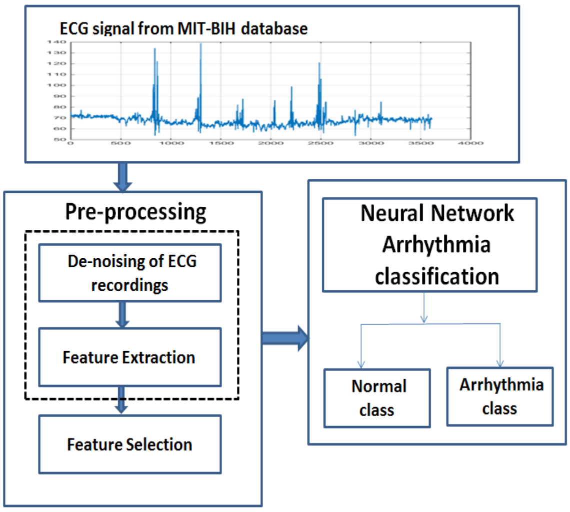

Fig. 1 Block diagram of the adopted methodology

A. MIT-BIH database

For a fair comparison of the methods aiming on the automatic heartbeat classification, we have used the public and the standard arrhythmia database MIT-BIH [6]. It is recommended by the Association for the Advancement of Medical Instrumentation (AAMI) [7]. Besides, it is used frequently by all groups aiming to arrhythmia classification. Indeed, MIT-BIH contains 48 records of heartbeats at 360Hz for approximately 30 minutes of 47 different patients. Each record has two ECG leads (lead A and lead B) which are depending on the electrodes configuration on the patient's body [6]. The database includes approximately 109,000 beat labels. Besides, it contains 25 normal records, 19 abnormal records and four paced beats records (102, 104, 107and 217) which were excluded in this study.

In this study, we have used only the first one minute ECG recordings. Referring to AAMI, we have divided MITBIH database into two datasets of 22 samples as they are shown in Table I, where Dataset1 is designed for training and Dataset2 is considered for evaluating the ANN classifier.

TABLE II. DISTRIBUTION OF MIT-BIH DATABASE RECORDING.

Database	Records number
Dataset1	101- 106-108 -109 -112- 114- 115- 116- 118- 119 122 -124- 201- 203- 205- 207- 208- 209- 215- 220 223- 230
Dataset2	100 103 105 111 113 117 121 123 200 202 210 212 213 214 219 221 222 228 231 232 233 234

B. Pre-processing

ECG input signals from the MIT-BIH database represent the electrical activity of the heart muscle. They are constituted basically by five successive waves as it is shown in figure2: P, QRS complex and T waves. They present respectively: atrial depolarization, ventricular depolarization and ventricular repolarization [8]. Due to biological and instrumental sources, different noise structure distress the ECG signal, which are basically skin resistance, respiration, muscle contraction, base line drift and power line interference[8] . Therefore, to filter out all kinds of noise, we have started by the pre-processing stage. Unfortunately, in this study, we have combined de-noising of ECG recordings and feature extraction stages by applying a robust algorithm based on wavelet Transform [9].

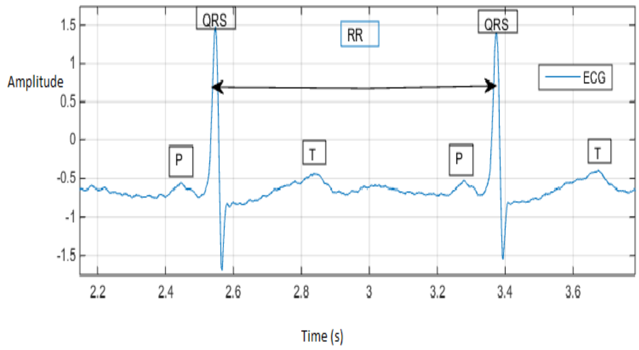

Fig. 2 ECG normal beat

C. Feature Extraction

The most important step in the automatic heartbeat classification is the feature extraction. In fact, to prepare feature vector, it is important to follow the cardiologists' arrhythmia classification practice. They focus particularly on ECG rhythm and morphology analysis [8].

In this purpose, many types of features can be extracted from the cardiac morphology in several ways: the time domain, the frequency domain and the time-scale domain [2,10].

The most common features used in the automatic heartbeat classification are RR interval and wavelet decomposition coefficients.

In this work, ECG features are categorized into two different groups: Morphological ECG features and Discrete wavelet transformer (DWT) coefficients [11, 12].

1) Morphological ECG feature :

We have used wavedet algorithm to extract from both ECG leads (A and B) these morphological features [9]. Here, features are divided into three sets as following:

- ECG peaks (P, Q, R, S, T),
- Time duration between waves (PR, PT, ST, QT, TT and QRS),
- ECG rhythm (RR interval).

Accordingly, we have obtained (P, Q, R, S, T) peaks as well as time durations between waves (PR, PT,ST,QT,QRS) since onset and offset of ECG waves were identified[2]. Regarding rhythm features, we have used the difference between the current and the previous QRS fiducial points namely RR interval sequence [8].

ISSN 2356-5608

2) *DWT coefficients features:*

For this feature group, we have used the Discrete Wavelet Transform (DWT) to extract features from ECG spectra in both time and scale domain [11]. Indeed, DWT decomposes original signal into low frequency and high frequency components. It is based on through two opposite filters: high pass and low pass fitters. The high-frequency component at low scales is called details and the low frequency components at high scale are known as approximations. These various components can be reconstructed back to form the original signal without any information loss.

In this study, DWT was applied to MIT-BIH recordings of one minute. Then, we have executed Daubechies 6 (db6) as the mother wavelet which decomposes the signal up to eight levels [13]. Therefore, we have returned for each ECG signal the measured approximations and details wavelet coefficients but we have obtained high dimensional feature vector size.

D. Feature selection

In order to construct the final feature vector with smaller number of features, a feature selection method was used.

In our case, we have calculated the standard deviation of (RR, PR, PT,ST,TT and QT) intervals as well as the maximum values of P, Q, R, S, T peaks and the number of R peaks count. Besides, we have considered the following statistics of the detail and approximation coefficients at each level: arithmetic mean, the variance and the standard deviation. Eventually, we have obtained entirely 60 features composed of 48 wavelet coefficient features and 12 morphological features for each ECG recording.

Then, we have applied the Principal Components Analysis (PCA) as the feature selection techniques to discriminant the most useful features for the ANN classifier [14].Indeed, PCA is one of the main linear dimensionality reduction techniques for extracting effective features from high dimensions. It is done by projecting the data into the feature space and finding the correlation among those features. It computes the principal components as a percentage of the total variability of the data used to select a number of them [14].

Hence, using PCA algorithm, the input matrix (60x44) becomes a matrix (10x44).

III. ARRHYTHMIA CLASSIFICATION

Once the feature vectors were defined, artificial intelligence algorithms can be built for arrhythmia heartbeat classification. In our case, we have used the neural network model which classifies ECG recordings into normal beats and pathological beats.

In this section, we introduce firstly ANN structure where two networks which are the cascade Forward NN and the Multi-layered Perceptron (MLP) NN are presented. Then, we introduce two training algorithms based on backpropagation method which are Resilient Backpropagation and Gradient Descent with Momentum.

A. ANN structure

Various types of NN structure are useful for arrhythmia classification, such as Feed Forward Network (FFN), Radial Basic Function (RBF) network, wavelet neural network, self-organization maps (SOM) and others [15].

In this section, we describe the FFN specially the Multi-Layer Perceptron (MLP) and the cascade-forward network to classify ECG recordings.

1) *MLP neural network*

In MLP network, the information moves in only one direction, forward, from the input layer, through the hidden layer to the output layer (fig.3). MLP Feedforward networks often have one or more hidden layers and use the log-sigmoid transfer function [16].

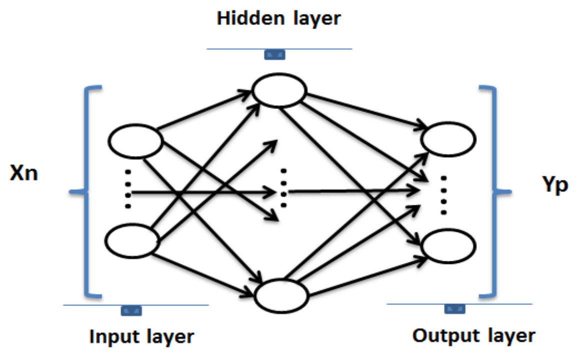

Fig. 3 MLP architecture

2) *Cascade-forward networks*

Cascade-forward networks include a weight connection from the input and every previous layer to following layers (Fig.4). The main symptom of this network is that each layer of neurons related to all previous layers of neurons [17]. Thus, it can learn complex relationships more quickly.

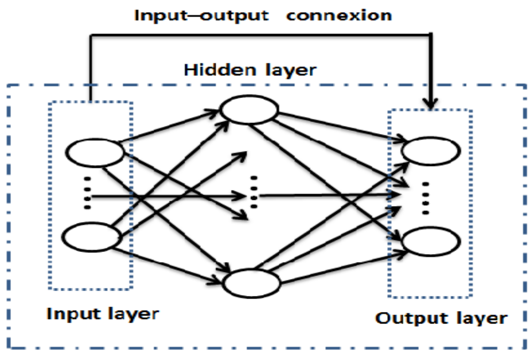

Fig.4 Cascade-forward architecture

B. ANN Training algorithms

The neural network process is to find an optimal set of weight parameters. This is done through a training algorithm. In layered feed-forward ANNs, the backpropagation (BP)

Conférence Internationale en Automatique & Traitement de Signal (ATS-2017)
Proceedings of Engineering and Technology – PET
Vol.22 pp.67-72

algorithm is used. This algorithm is based on Gradient Descent (GD) rule which tends to adjust weights and reduce system error in the network [13].
It can be summarized in four fundamental steps:

- Initialize the connection weights with random values.
- Compute the output of the ANN by propagating each input pattern through the network in a forward direction.
- Compute the Mean Square Error E_k between the desired output t_i and the produced output a_i by the ANN via equation (1).

$$E_k = \sum_{i=1}^{N} (t_i - a_i)^2 \qquad (1)$$

- Adjust the connection weights according to equation (2) where η is the learning rate and $\frac{dE_k}{dw}$ is the gradient.

$$W_{t+1} = W_t - \eta \frac{dE_k}{dw} \qquad (2)$$

The above process is repeated until a stopping criterion is met which can be a desired minimum error or a maximum number of iteration.
Certainty, the choice of the learning rate is important for the method since a high value can cause weight oscillation while a too low value slows the training convergence. In order to avoid oscillation inside the network and to improve the rate of convergence, there are refinements of this backpropagation algorithm.
In the following section, we introduce two BP algorithms variants: GD with Momentum (GDM) and Resilient BP (RPROP).

1) GDM training algorithm:

When BP algorithm has trouble around local optima as can be seen in fig.5 (a), the (GDM) algorithm accelerates GD in the relevant direction and reduces oscillation as in Fig.5 (b) .

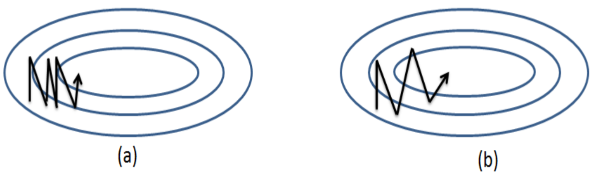

Fig.5 GD without momentum (a)/ GD with momentum (b)

It does this by adding a fraction parameter called the momentum coefficient which controls the influence of the last weight update direction on the current weight update (see equation (3)) where W_t is the momentum factor which is held constant during the entire training process and is usually set to 0.9, W_{t-1} is the last point of weight, W_{t-1} the current weight and Δ^+ is the next weight.

$$W_{t+1} - W_t = -\eta \frac{dE_k}{dw} + \alpha(W_t - W_{t-1}) \qquad (3)$$

2) RPROP training algorithm:

For better weight updates, RPROP only uses the sign of the derivative. If the error gradient for a given weight has the same sign in two consecutive epochs, the update weight is increased by a factor Δ^+ (see equation (4)).

$$W_{t+1} = \Delta^+(W_t - W_{t-1}) \qquad (4)$$

If in the other hand, the sign switched, the update value is decreased by a factor Δ^- (see equation (4)).

$$W_{t+1} = \Delta^-(W_t - W_{t-1}) \qquad (5)$$

RPROP assumes that weights are always changed by adding or subtracting the current step size, regardless of the absolute value of the gradient [13, 14].

IV. EXPERIMENTAL DATA

In this section, we discuss the appropriate NN structure and the suitable training algorithms. We have started by comparing the cascade Forward NN and the (MLP) NN architectures. Similarly, we have done a brief comparison between two training algorithms (RPROP and GDM) based on BP approach using MLP topology.
Both of the comparative studies adopted the (MSE) as mentioned in equation (1) and the ACC as described in the equation (6), in order to evaluate ANN training and testing results quality.

$$ACC_{\%} = \frac{\text{correctly classified sampled}}{\text{total number of samples}} \qquad (6)$$

In this work, the experimental results are carried out in MATLAB software package 14.b. Moreover, Central Processing Unit (CPU) times are given for intel ®Core ™ i5 -2410M CPU(2.30 GHz).

Among 48 ECG recordings each of length 1 min, only 44 non pacemaker recordings from MITBIH database (25 records of normal class and 19 from abnormal class) were used. Therefore, we have selected features from both ECG morphology and DWT coefficients to constrain the neural network input matrix. We have attained 60 features (48 DWT based feature and 12 morphological). Then, we have applied PCA as the feature selection algorithm to reduce the input matrix (60x44) size. Thus, we have obtained 10 most discriminative features.

For the arrhythmia classification, we have applied an ANN to classify ECG recording into two classes normal and abnormal. In fact, we have applied the reduced matrix (10x44) as the input layer. Concerning the network output layer, two neurons were used as (0, 1) and (1, 0) referring to normal and abnormal class. Regarding number of hidden layers ,we have used one hidden layer which was fixed based on application. Both of the NN (MLP and cascade forward) have been used Tan-sigmoid transfer function. Moreover, the system is trained using samples from dataset1 and it is tested using samples from dataset2.

1) Comparison MLP / cascade-forward NN

ISSN 2356-5608

Currently, we have compared the comportment of MLP and cascade-forward networks to classify ECG recordings.
The training and the testing of the NN were carried out with various numbers of neurons from five to fifteen in a one hidden layer. For that, a brief comparison between these ANN networks was done based on MSE and ACC values.
The analysis study is presented in Table II where NHN is the number of neurons in the hidden layer.
As it is shown in Table II, cascade forward network underperforms MLP network. It gives the best result (90.3%) using 10 neurons in the hidden layer. However, MLP network provides good performance especially when the hidden layer is composed of ten neurons. It gives a null MSE and the best result of ACC (100%) for its training.
For the generalization of its training results, MLP network gives also the best result (99%) using testing dataset (dataset2).

TABLE.II MODEL ACCURACY BETWEEN MLP AND CASCADE-FORWARD NEURAL NETWORKS

Dataset	ANN	MLP			Cascade-Forwardnet		
	NHN	5	10	15	5	10	15
Dataset 1	MSE	0.18	0	0.23	0.39	0.21	0.38
	ACC	93	100	83.6	63	67	66
Dataset 2	MSE	0.21	0.03	0.29	0.40	0.13	0.21
	ACC(%)	87	99	93.6	86.4	90.3	86

After comparing MLP and cascade_forward NN performances, we emphasize the use of MLP NN structure. In the following section, we focus on its suitable training algorithm.

2) *MLP training algorithm*

Two BP training algorithms which are RPROP and GDM are compared using MLP NN structure
To evaluate these training algorithms, a learning rate equal to 0.15 is used along the study to determine the length of the weight update. As well as, the maximum number of epochs was fixed on 100 epochs.
Regarding RPROP algorithm, $\Delta^{+}$is empirically set to 1.2 and $\Delta^{-}$to 0.5. The total CPU time used by the two training algorithms is around 664s. Results are shown in Table III where NE is the number of epochs.
As it is illustrated in Table III, the GDM algorithm used many epochs (99) with around of 194s of training time.

TABLE III TRAINING ALGORITHMS COMPARISON FOR MLP NETWORK

Dataset	Training Algorithms	ANN performance		
		MSE	NE	CPU time(s)
Dataset1	RPROP	0	15	141.18
	GDM	0.045	99	194.83
Dataset2	RPROP	0	13	138.08
	GDM	0.004	41	191.97

The figure 6 shows similarly the best MSE of the GDM algorithm.

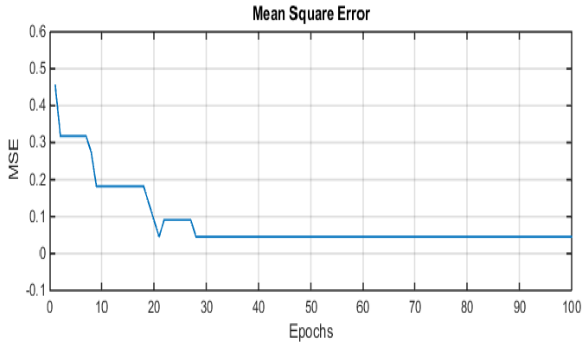

Fig.6 The best MSE of GDM algorithm

However, RPROP algorithm achieved a null MSE only within 141s and 15 as NE. Accordingly; its memory requirements are relatively small in comparison to GDM algorithm. This confirms the faster convergence rate of RPROP algorithm. This could be explicated by the fact that there's no need to store the update values for each weight and bias when RPROP is the training algorithm. In the figure7, we observe the best MSE of the RPROP training algorithm.

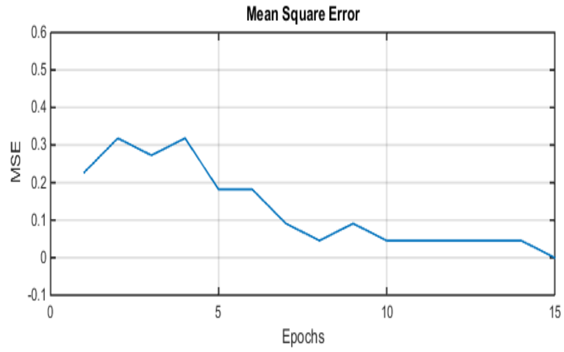

Fig 7 The best MSE of RPROP algorithm

V. CONCLUSION

In this paper, a NN model for ECG arrhythmias classification was proposed. We have used 44 recordings from the MIT-BIH arrhythmias database for training as well as testing the classifier. The proposed system consists of two phases: ECG pre-processing and NN arrhythmia classification.
In the first phase, de-noising of ECG recordings and feature extraction stages are combined by applying a robust algorithm to ECG artefact, based on wavelet Transform. Hence, we have extracted ECG features which are categorized into two different groups: Morphological ECG features and DWT coefficients. Then, in order to reduce the feature vector size, we have applied PCA algorithm as the feature selection approach.

ISSN 2356-5608

In the second phase, we have compared MLP and Cascade _forward NN architecture. The study reveals that the performance of MLP algorithm is better than cascade forward network by comparing the MSE and ACC values.

The use of RPROP algorithm for training data is more efficient than using GDM algorithm. Despite of the choice of MLP neural network structure and RPROP as backpropagation training algorithms, we have to reduce MLP training CPU time.

As a perspective, another study of arrhythmia system by using different NN structures, different transfer function and different training algorithms is required. We propose also a hybrid neuro-fuzzy networks method in order to minimize the problems of MLP, increasing its generalization and reducing its training time.

REFERENCES

[1] Jamison DT, Feachem FG, Makgoba MW, Bos ER, Baingana FK, Hofman KJ and Rogo KO, *Disease and Mortality in Sub-Saharan Africa*. Washington D.C.: World Bank, 2006.

[2] Gutiérrez-Gnecchi, J. A., Morfin-Magaña, R., Lorias-Espinoza, D., del Carmen Tellez-Anguiano, A., Reyes-Archundia, E., Méndez-Patiño, A., & Castañeda-Miranda, "DSP-based arrhythmia classification using wavelet transform and probabilistic neural network,''*Biomedical Signal Processing and Control*, 32, pp. 44-56, 2017.

[3] Garcia, G.Moreira, G.Luz, E.Menotti, and D.David, "Improving automatic cardiac arrhythmia classification: Joining temporal-VCG, complex networks and SVM classifier,'' in *International Joint Conference on Neural Networks (IJCNN)*, 2016, p. 3896-3900.

[4] Gautam, M. K., Giri and V. K, "A Neural Network approach and Wavelet analysis for ECG classification,'' in *IEEE International Conference on Engineering and Technology (ICETECH)*,March.2016, pp. 1136-1141.

[5] Özbay, Y., Ceylan, R., and Karlik, B. ,A, "fuzzy clustering neural network architecture for classification of ECG arrhythmias," *Computers in Biology and Medicine*,pp.376-388,2006.

[6] Moody, G. B., & Mark, R. G., "The MIT-BIH arrhythmia database on CD-ROM and software for use with it," in IEEE *Proceedings Computers in Cardiology,* Sep.1990, pp. 185-188.

[7] Association for the Advancement of Medical Instrumentation, *AAMI standards and recommended practices*, Association for the Advancement of Medical Instrumentation (AAMI), 1986.

[8] L.Sörnmo, P. Laguna, *Electrocardiogram (ECG) signal processing.* Wiley encyclopedia of biomedical engineering, 2006.

[9] Demski, A., & Soria, M. L., "ecg-kit: a Matlab Toolbox for Cardiovascular Signal Processing," *Journal of Open Research Software*, vol. 4, 2016.

[10] F.A Elhaj, N. Salim, A. R.Harris, T. T. Swee and T. Ahmed, "Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals,'' *Computer methods and programs in biomedicine*, 127,pp. 52-63, 2016.

[11] Dewangan, N. K., & Shukla, S. P., "ECG arrhythmia classification using discrete wavelet transform and artificial neural network," in *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT),*May 2016, pp. 1892-1896.

[12] H. Khorrami and M. Moavenian, "A comparative study of DWT, CWT and DCT transformations in ECG arrhythmias classification," *Expert systems with Applications,* vol 37, pp.5751-5757, 2010.

[13] H. M.Rai, A.Trivedi and S.Shukla, " ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier, " *Measurement*, vol.46, pp.3238-3246, 2013.

[14] Shri, T. P., & Sriraam, N., "Comparison of t-test ranking with PCA and SEPCOR feature selection for wake and stage 1 sleep pattern recognition in multichannel electroencephalograms," *Biomedical Signal Processing and Control*, vol.31, pp.499-512, 2017.

[15] Luz, E. J. D. S., Schwartz, W. R., Cámara-Chávez, G., & Menotti, D., "ECG-based heartbeat classification for arrhythmia detection: A survey," *Computer methods and programs in biomedicine,* pp.144-164, 2016.

[16] M. W.Gardner and S. R. Dorling, "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences ", *Atmospheric environment*, vol.32, pp.2627-2636, 1998.

[17] G., Capizzi, G. L. Sciuto, P.Monforte and C.Napoli , "Cascade feed forward neural network-based model for air pollutants evaluation of single monitoring stations in urban areas, " *International Journal of Electronics and Telecommunications*, vol.61, pp.327-332, 2015.