Artificial Neural Network Based ECG Feature Extraction Using Wavelet Transform

Priyanka Agrawal[1], Vanya Arun[2] and Amlan Basu[3]

[1,2] Department of Electrical & Electronics Engineering, IILM University, UP, India
[3]Centre for Ultrasonic Engineering, University of Strathclyde, Glasgow, U.K.

Abstract. In this research, Automatic techniques to detect diseases have been developed because of requirements of continuous attention to patient having heart diseases. This research deals with implementation of Artificial neural network methods for analyzing ECG (Electrocardiogram) signals with a focus on early and accurate detection. Feature extraction of ECG signal plays vital role in cardiovascular diseases. ECG signal is decomposed using wavelet transform and then feature extracted of decomposed ECG signal are given as input to Neural Network. The wavelets used for decomposition are Daubechies and Symmetric. The selection of detail coefficient d4 had been done based on the following important parameters i.e. Energy, Frequency and Correlation. The overall of detection using db6 and sym11 were 96.65% and 84.37%. In this work, study of the classification of ECG signal has been done in detail by using computational methods effectively for early cardiovascular diagnosis. Coefficients of discrete wavelet transforms are used for analyzing ECG signals in conjunction with the Artificial Neural network (ANN). Three different types of ECG data have been used normal sinus rhythm, supra ventricular arrhythmia and atrial fibrillation. Decomposition and Classification of ECG signals using discrete wavelet transform and Artificial Neural Network have been successfully designed. The method has been implemented on 18 subjects. The results show that proposed method is effective for classification of normal and cardiac arrhythmia with an overall accuracy of 97.5%

Keywords: ECG, Neural Network, Wavelet, Feature Extraction, Morphological Features, Statistical Features.

1 Introduction

1.1 A Subsection Sample

This paper focuses to extract morphology and statistical parameters from the ECG signal using the Discrete Wavelet Transform and Artificial Neural Network technique. ECG signal is decomposed using discrete wavelet transform then features of Decomposed ECG signal are extracted and these features are trained using back propagation neural network.

ECG is a diagnostic methodology that records the biological signals activity of heart from the chest. It's a common and painless test used to quickly detect heart problems and monitor the heart's health. ECG is used to find heart problems and health of heart in very quick way. This is non- invasive and painless test. ECG may be used to find

This is a peer-reviewed, accepted manuscript of the following paper: Agrawal, P., Arun, V., & Basu, A. (in press).
Artificial Neural Network Based ECG Feature Extraction Using Wavelet Transform. Paper presented at Second International Conference on Emerging Wireless Technologies and Sciences-2024, Dehradun, India.

Arrhythmias, coronary artery diseases and heart attack. During the process of diagnosis, many electrodes are used on human limbs [1]. These electrodes in pairs are called lead. Analysis of ECG signal systems have a vital role in assisting medical staff by for fast treatment in hospitals and enhancing quality of life by finding abnormality at an early age. ECG signals have been used by doctors to find diseases like arrhythmia and myocardial infarction An ECG signal has P, QRS complex, and T waves as shown in Fig. 1. The P wave shows atrial depolarization. The QRS complex reflects ventricular depolarization. The T wave shows ventricular repolarization. Cardiac diseases are detected by analyzing these waves. Spikes can be seen in waves due to artifacts. Such artifacts are in waves due to movements in body of patients and electrode and interferences [2]. So, these noise and artifacts should be filtered from ECG signals so that accurate analysis of ECG signal can be done. For removing noise, different time and frequency domain transforms have been used. Previous research involves Discrete Wavelet Transform, Hermitian Basis and other methods but now a days Wavelet transform is most widely used [3].

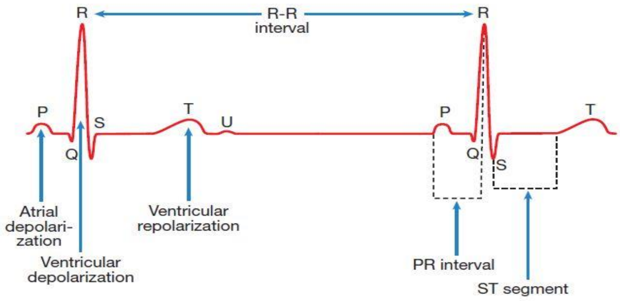

Fig. 1. ECG Waveform.

2 Methodology

Various methodologies of automatic diagnosis have been done. Neural network and wavelet tool box has been used for decomposition of ECG and its classification .Data acquisition has been done from MIT-BIH data base available at physionet.org for performance evaluation. For the analysis of signal various tools are used such as wavelet, neural network. During the analysis of ECG signals, variations in features of ECG signals is the main problem. Such variations occur not only in different patients or group of patients but also with in the same patient . In our present research, different ECG samples of different diseases have been taken.In order to classify ECG signals based on diseases, following ECG signal analysis pipeline will be done:Selections of wavelet, particular detailed coefficient, feature extraction, process of classification using ANN.

2.1 Data Acquisition

It is mandatory that Standard database should be selected to analyze ECG signal. ECG signals used in this paper were taken from Physionet MIT-BIH.ECG signals are denoted by header file and a mat file [4]. Header file has detailed information about sampling frequency, number of ECG leads, and information about patient.ECG signals used in this paper are: MIT BIH Normal, Atrial fibrillation and super ventricular arrhythmia. 6 subjects of Normal ECG, Atrial fibrillation and Supra ventricular is taken, half of them have been used for training and remaining for testing. All classes, sampling frequency of each class, number of samples are shown in Table 1. Data base contains total 18 records of 1 min recording. 1024 samples of each record have been taken in this paper.

Table 1. Data

Classes	Sampling frequency	Gain (mv)
N (Normal)	128 Hz	200
SV (Supra Ventricular Ar-rhythmia)	128 Hz	200
AF (Atrial Fibrillation)	250 Hz	200

2.2 Wavelet Transform

Wavelet transform deals with processing method for non-stationary signals with enough scale values and shifting in time. The wavelet transform represents signals in different resolutions by dilating and compressing in different functions. The wavelet transform is considered a valuable tool to analyse data from different fields such as biomedical, mathematics, mechanics etc. It has now been applied to a different types of biomedical signals: the ECG, EMG, EEG, image compression. It is very important to select mother wavelet for preprocessing of ECG. There are different basis wavelets for example Symmlet, Harr, Daubechies wavelets etc. The main advantage of wavelet method is the availability of wavelet functions, so making easy to choose the most suitable for the signal under investigation.

There are some predefined criteria to select optimal wavelet transform. Suitable wavelet for ECG signal is Daubechies wavelet. This wavelet families have similarity with QRS complex and energy spectrums of Daubechies wavelets are concentrated around low frequencies[5]. The generated cross correlation coefficient of the ECG signal with various wavelet filters is shown in Fig. 2. Results strengthens the suitability of Daubechies wavelet filter of order 8 is an optimal one for the ECG signal processing. So, in this paper, Daubechies wavelet is used for ECG signal processing.

After the selection of a particular wavelet, ECG signal is decomposed using discrete wavelet transform. As ECG signal is decomposed in to six detailed coefficients then Entropy criteria is used in this paper to select one detailed coefficient. The detailed coefficient which shows highest energy with lowest entropy is chosen for feature extraction and classification.

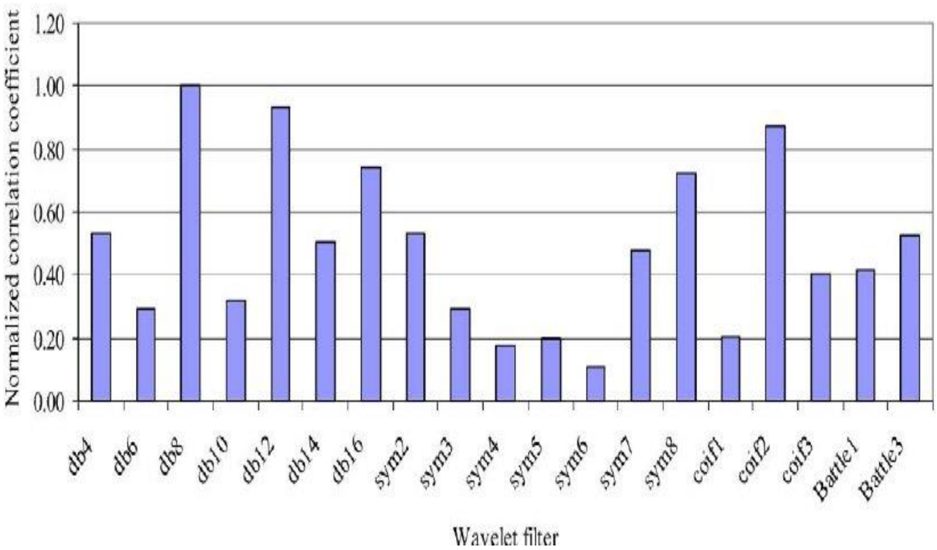

Fig. 2. Correlation between different wavelet and ECG signal

2.3 Feature extraction

Human can perform the process of analysing and extracting Morphology and statistical features of ECG signals. Wave peaks (amplitude) as well as time durations(intervals) consist of clinically useful information present in ECG signal [6].
Morphological features were found as the most important for arrhythmia classification. Morphological features are calculated by signal processing methods while statistical features are calculated mathematically. In this paper, features which have been extracted are R wave amplitude, RR interval, standard deviation, variance, etc. These features are calculated by given formulae.

$$HR = 60 / RR \text{ Interval} \quad (1)$$

$$RMS = \sqrt{\frac{1}{N}\left(\sum x^2(n)\right)} \quad (2)$$

$$Varience = \frac{1}{N}\left(\sum_{n=1}^{N}\left[x(n) - x\right]^2\right) \quad (3)$$

$$SD = \sqrt{\frac{1}{N}\left(\sum_{n=1}^{N}\left[x(n) - \bar{x}\right]^2\right)} \quad (4)$$

2.4 Artificial Neural Network (ANN)

ANN used in this paper has 9 neurons in input layer, 1 hidden layer with 15 neurons and 3 neurons in output layer. Learning rate of 0.2, momentum factor of 0.5, error goal

of 10^{-5}, 100 epochs and batch size of 1 were used during the training. Mean Square Error (MSE) has been used as the loss function of the ANN.
Once the features are extracted, normalization is done in such a way that each data should fall within ranging from 0 to 1[7]. As the values of features may vary, it is important to normalize all the features to same level and to obtain optimal results. Formula for normalization is given by equation (5),

$$Normalized\ Value = 0.1 + \left[\frac{\{(Actual\ Value - Minimum\ Value)\ X\ 0.8\}}{\{Maximum\ Value - Minimum\ Value\}}\right] \quad (5)$$

Back propagation algorithm is used for training the ANN. MSE is reduced between the real output (actual output) and desired output by adjustment of weights using gradient descent search by back propagation. Flow chart of back propagation algorithm is shown in Fig. 3.

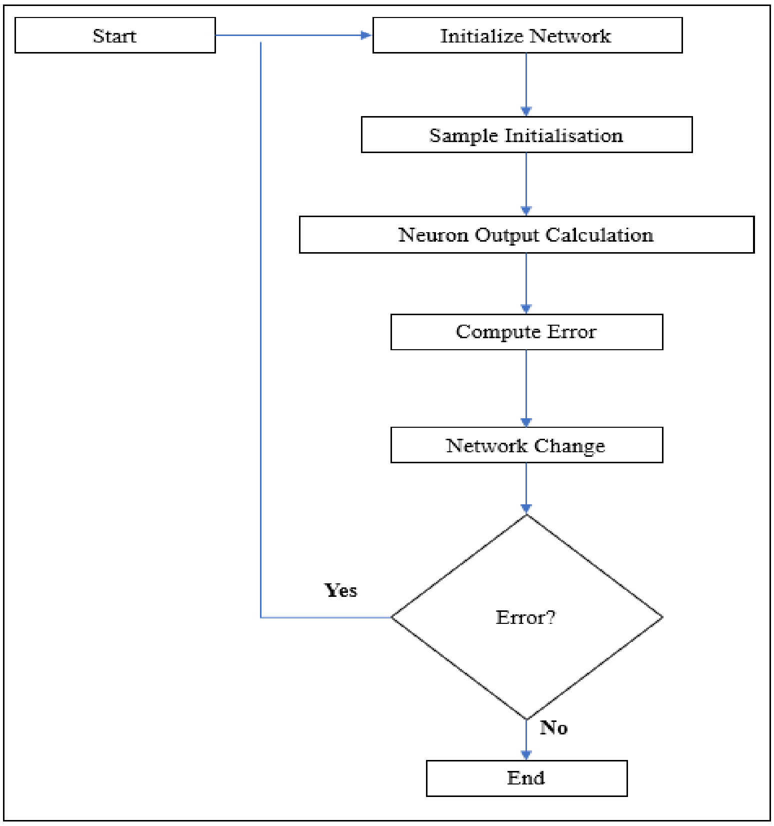

Fig. 3. Flow Chart of EBP algorithm

The basic steps used in back propagation algorithm are described below:

1. Find the different learning parameters based on predefined network structure, involving input, output and hidden layer, error measurements and learning rate[8].

2. Provide input and target vector.
3. Start learning and follow given steps:

Amount proportional to multiplication of error signal on the unit k and output of unit j adjust the weight of connection given below:

$$\Delta_p w_{jk} = \gamma \delta_k^p y_j^p \quad (6)$$

Error signal for output unit is given by equation:

$$\delta_o^p = (d_o^p - y_o^p) F'(s_o^p) \quad (7)$$

If sigmoid is the activation function then

$$y^p = F(s^p) = \frac{1}{1+e^{-s^p}} \quad (8)$$

Find the derivation of activation function

$$F'(s^p) = y^p(1-y^p) \quad (9)$$

In such way that error signal is given for output unit

$$\delta_o^p = (d_o^p - y_o^p) y_o^p (1-y_o^p) \quad (10)$$

a. Error signal for hidden layer is calculated in term of error signals of directly connected units and their weights. For sigmoid function:

$$\delta_h^p =' (s_h^p) \sum_{o=1}^{N_o} \delta_o^p w_{ho} = y_h^p (1-y_h^p) \sum_{o=1}^{N_o} \delta_o^p w_{ho} \quad (11)$$

b. If we make changes in weight dependent of past weights with the help of adding momentum then oscillations can be avoided.

$$\Delta w_{jk}(t+1) = \gamma \delta_k^p y_j^p + \alpha \Delta w_{jk}(t) \quad (12)$$

Every time new sample is entered till error requirement is not achieved. Re-random order is needed for input order during training [9]. When network response is obtained then error of the last output layer is found only, this error is recalculated according to instructions given on previous layers. The base of Neural Networks is back propagation algorithm[10]. Computation of error is done again and again for fast and accurate results. Error function is minimized in this algorithm.

3 Results and Discussion

3.1 Selection of detailed coefficient

In this paper, initially ECG signal is decomposed in to different detailed coefficients using DWT.A particular detailed coefficient is chosen on the basis of entropy criteria. Figure 4, 5 & 6 shows the entropy of different detailed coefficients for Normal, Supra ventricular (SV) and Atrial fibrillation (AF) respectively. It also indicates that detailed coefficient of level 3 for normal, level 1 for supra ventricular and level 3 for atrial fibrillation should be selected to extract features as it represents minimum entropy.

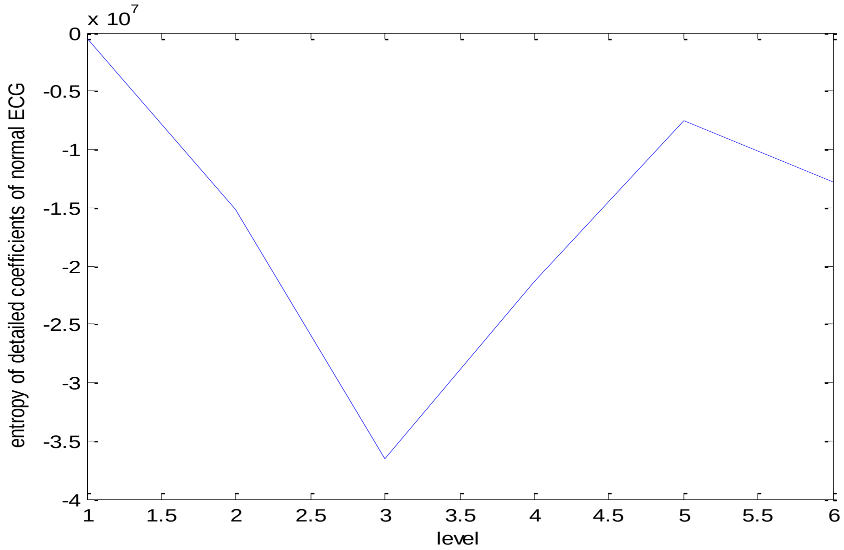

Fig. 4. Entropy Plot of Detailed Coefficient of Normal ECG

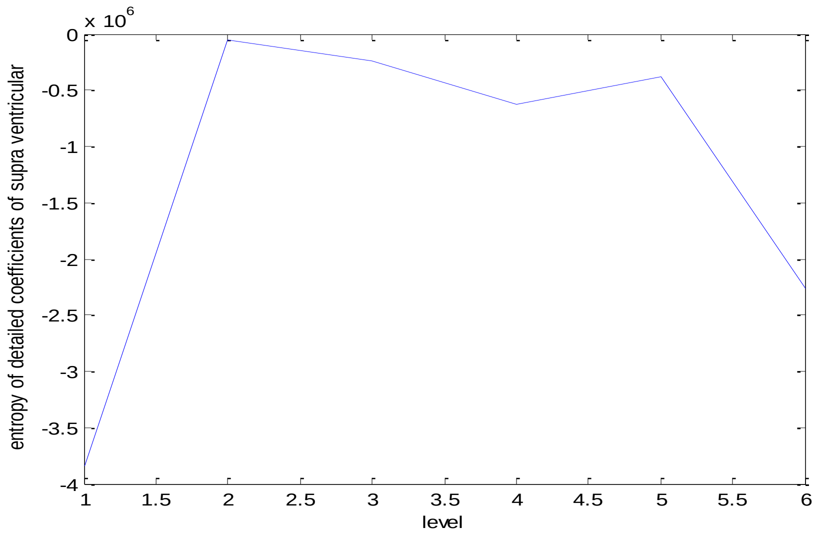

Fig. 5. Entropy Plot of Detailed Coefficient of Supra Ventricular

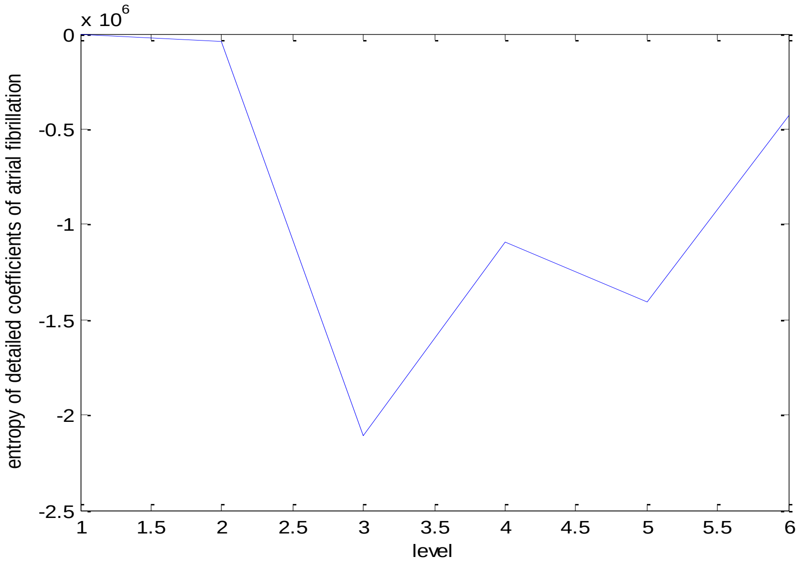

Fig. 6. Entropy Plot of Detailed Coefficient of Atrial Fibrillation

3.2 Feature Extraction & Normalization

There are two types of features (Statistical and morphological features) have been extracted for all three types of signals. After that they are normalized to obtain the optimum results. Table 2 shows the different normalized statistical features. These features provide information regarding cardiac health condition[11]. All these parameters are useful for discrimination between normal and arrhythmia patients. Different statistical features are calculated from normal, atrial fibrillation and supra ventricular ECG signals. Different statistical features calculated in this paper are maximum peak value, minimum peak value, mean, standard deviation, variance,entropy and mode.

Table 2. Normalized Statistical Features

Selected-detailed coefficients for Different ECG Signal	Mean (mv)	Mode	entropy
Detailed coefficient of third level for normal ECG signal 1	0.1139	0.1012	0.2665
Detailed coefficient of third level for normal ECG signal 2	0.1134	0.1095	0.8686
Detailed coefficient of third level for normal ECG signal 3	0.1135	0.1000	0.2661
Detailed coefficient of third level for normal ECG signal 4	0.1135	0.1072	0.7745

Detailed coefficient of third level for normal ECG signal 5	0.1133	0.1085	0.8696
Detailed coefficient of third level for normal ECG signal 6	0.1134	0.1088	0.8813
Detailed coefficient of first level for SV ECG signal 1	0.1135	0.1084	0.1000
Detailed coefficient of first level for SV ECG signal 2	0.1136	0.1119	0.7965
Detailed coefficient of first level for SV ECG signal 3	0.1134	0.1118	0.5470
Detailed coefficient of first level for SV ECG signal 4	0.1134	0.1110	0.9000
Detailed coefficient of first level for SV ECG signal 5	0.1134	0.1120	0.8392
Detailed coefficient of first level for SV ECG signal 6	0.1135	0.1125	.8331
Detailed coefficient of third level for AF ECG signal 1	0.1134	0.1095	0.8632
Detailed coefficient of third level for AF ECG signal 2	0.1134	0.1085	0.8629
Detailed coefficient of third level for AF ECG signal 3	0.1134	0.1053	0.7669
Detailed coefficient of third level for AF ECG signal 4	0.1134	0.1103	0.8809
Detailed coefficient of third level for AF ECG signal 5	0.1134	0.1119	0.8959
Detailed coefficient of third level for AF ECG signal 6	0.1134	0.1087	0.8471

Selected-detailed coefficients for different ECG signals	**Max. (mv)**	**Min. (mv)**	**SD (mv)**	**Var. (mv)**
Detailed coefficient of third level for normal ECG signal 1	0.123	0.101	0.116	0.754
Detailed coefficient of third level for normal ECG signal 2	0.117	0.109	0.114	0.163
Detailed coefficient of third level for normal ECG signal 3	0.125	0.100	0.117	0.900
Detailed coefficient of third level for normal ECG signal 4	0.119	0.107	0.115	0.282
Detailed coefficient of third level for normal ECG signal 5	0.116	0.108	0.114	0.195
Detailed coefficient of third level for normal ECG signal 6	0.116	0.108	0.114	0.165
Detailed coefficient of first level for SV ECG signal 1	0.117	0.108	0.114	0.152

Detailed coefficient of first level for SV ECG signal 2	0.117	0.111	0.114	0.136
Detailed coefficient of first level for SV ECG signal 3	0.115	0.111	0.113	0.119
Detailed coefficient of first level for SV ECG signal 4	0.116	0.111	0.113	0.125
Detailed coefficient of first level for SV ECG signal 5	0.115	0.112	0.113	0.118
Detailed coefficient of first level for SV ECG signal 6	0.118	0.112	0.114	0.139
Detailed coefficient of third level for AF ECG signal 1	0.118	0.109	0.114	0.163
Detailed coefficient of third level for AF ECG signal 2	0.119	0.108	0.114	0.183
Detailed coefficient of third level for AF ECG signal 3	0.120	0.105	0.115	0.301
Detailed coefficient of third level for AF ECG signal 4	0.116	0.110	0.114	0.152
Detailed coefficient of third level for AF ECG signal 5	0.115	0.111	0.113	0.120
Detailed coefficient of third level for AF ECG signal 6	0.118	0.108	0.114	0.205

Since most of the diagnostic information lies around the R peak of the ECG signal, hence a portion of signal before it and a portion of signal after it are selected for processing[12]. So After the extraction of statistical parameters, different morphological parameters like R wave amplitude, RR Interval are extracted.[13] Now the morphological features are normalized which are shown in table 3.

Table 3. Normalized Morphological Features

Selected detailed coefficient for different ECG signals	R Wave Amplitude (mv)	RR Interval (seconds)
Detailed coefficient of third level for normal ECG signal 1	0.8679	0.1095
Detailed coefficient of third level for normal ECG signal 2	0.4204	0.1259
Detailed coefficient of third level for normal ECG signal 3	0.900	0.1101
Detailed coefficient of third level for normal ECG signal 4	0.5503	0.1095
Detailed coefficient of third level for normal ECG signal 5	0.4641	0.1058
Detailed coefficient of third level for normal ECG signal 6	0.4600	0.1080

Detailed coefficient of first level for SV ECG signal 1	0.5145	0.1000
Detailed coefficient of first level for SV ECG signal 2	0.4537	0.1012
Detailed coefficient of first level for SV ECG signal 3	0.5551	0.1055
Detailed coefficient of first level for SV ECG signal 4	0.5076	0.1061
Detailed coefficient of first level for SV ECG signal 5	0.4378	0.1070
Detailed coefficient of first level for SV ECG signal 6	0.7434	0.1043
Detailed coefficient of third level for AF ECG signal 1	0.5044	0.1022
Detailed coefficient of third level for AF ECG signal 2	0.5076	0.1026
Detailed coefficient of third level for AF ECG signal 3	0.6173	0.1033
Detailed coefficient of third level for AF ECG signal 4	0.6066	0.1003
Detailed coefficient of third level for AF ECG signal 5	0.4759	0.1041
Detailed coefficient of third level for AF ECG signal 6	0.5139	0.1019

R Peaks have been detected in normal ECG, super ventricular ECG and Atrial Fibrillation ECG and they have been marked by star using MATLAB.

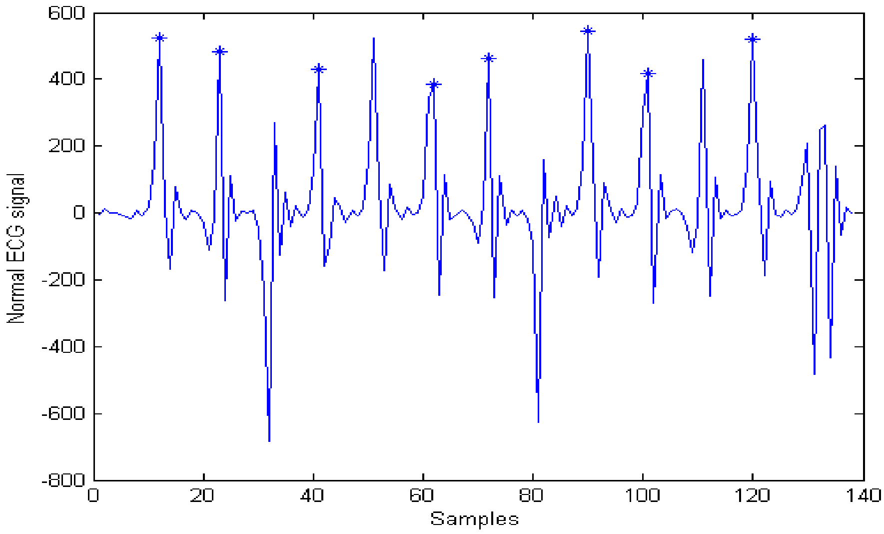

Fig. 7. Detection of R peak in Normal E

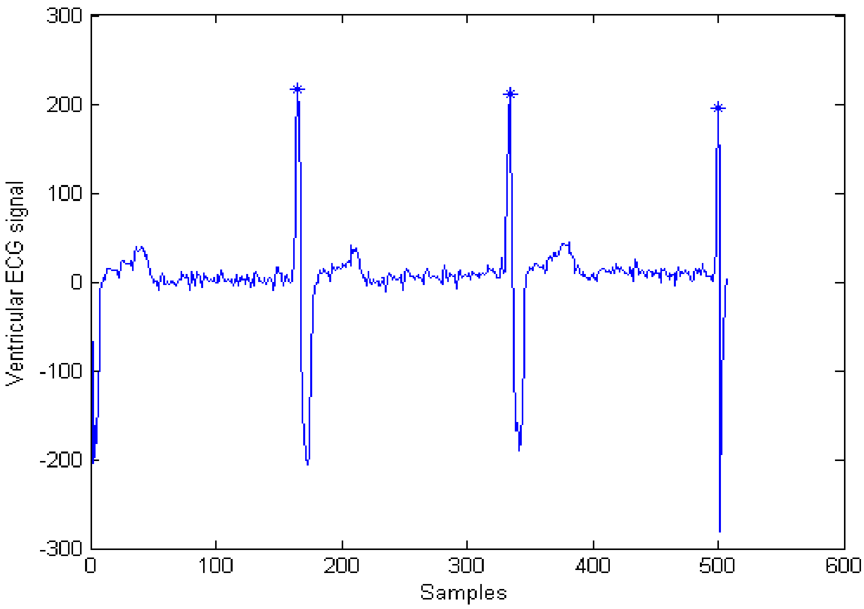

Fig. 8. Detection of R peak in Super Ventricular ECG

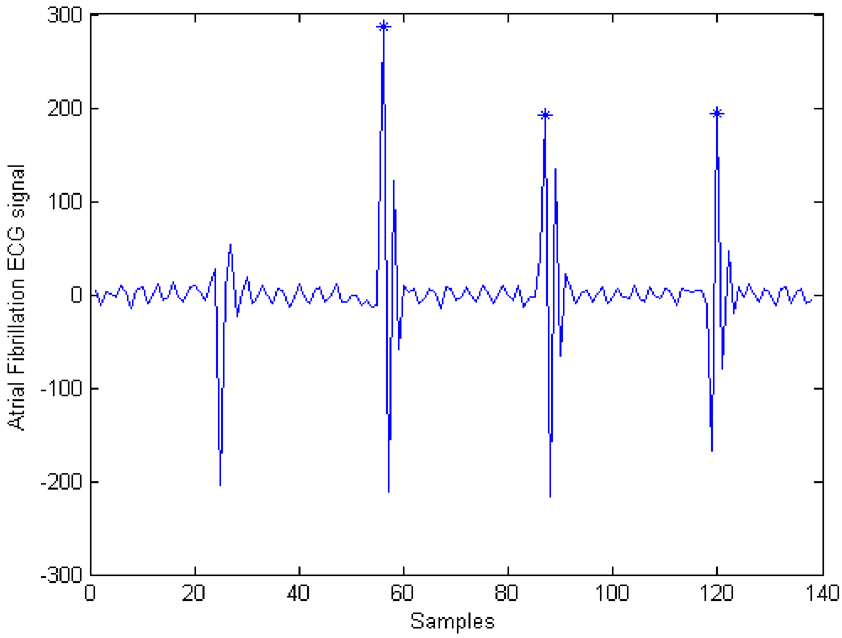

Fig. 9. Detection of R peak in Atrial Fibrillation ECG

3.3 Back Propagation (BP) Algorithm

Multi layered feed forward network with BP algorithm is used during training and testing phase. Mean square error (MSE) is employed as error function [14]. But in this paper, Network is trained with momentum factor 0.5, learning rate 0.2, number of iterations 1000 and error goal 10^{-5}. As total selected decomposed coefficients are 18 and 9 features of each coefficient is determined so 9x9 matrix (9 features and 9 signals) are given as input of neural network for training purpose and another 9x9 matrix are used for testing purpose. Variation in MSE provides training performance with different iterations. Figure 10 shows the performance graph:

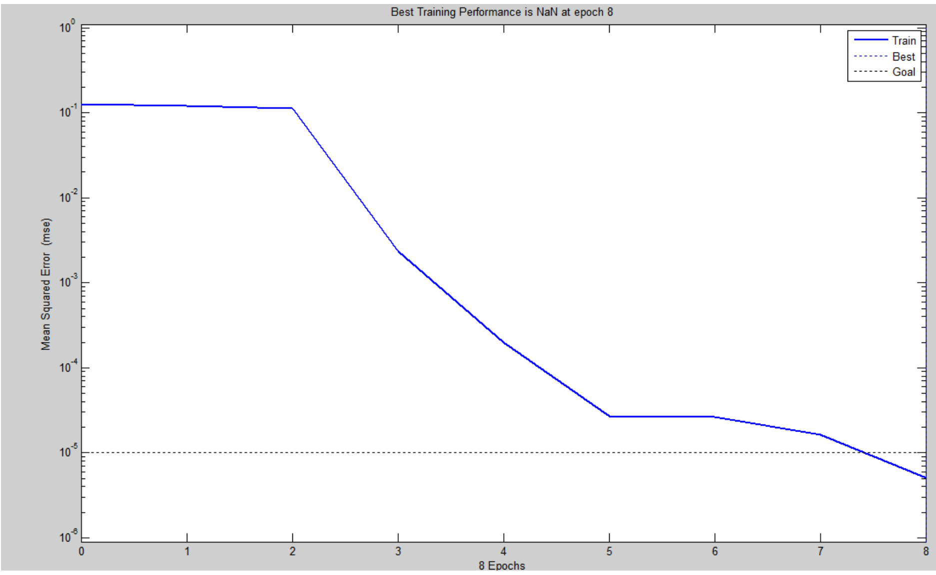

Fig. 10. Performance Analysis

The accuracy during the testing of the network is found to be 97.5% (only 2 out of 81 returned the negative results). The error performance goal is obtained at the eight epochs as shown in above figure. Accuracy can be increased by varying the network parameters (iterations, momentum factor, learning rate, number of inputs to the network and error goal).

4 Conclusion

The methodology used in this paper has been applied on 18 patients. The results of this work produced an accuracy of 97.5% .The result obtained in the paper shows that data base can be classified using ANN. Architecture for feature extraction of 18 ECG signals and their classification have been developed successfully.

This paper presents the use of wavelet transform and Multi-Layer neural networks for classification of normal and disease subjects. A classifier expert system for ECG signals can also be developed by using Fuzzy logic and Genetic algorithm. Features extracted from the program can be used for training purpose, concludes the dissertation indication for its utility in the field of Biomedical & future scope for further work. One step can also be added that features will be extracted by decomposition of ECG signal by WT, the feature selection will be done by Genetic Algorithm then more reduced features will be given to input of ANN for classification. Further, research is required in order to enhance in the field of clinical applications of signal processing methodology.

References

1. Zhu, Huaiyu, Yun Pan, Fan Wu, and Ruohong Huan. "Optimized electrode locations for wearable single-lead ECG monitoring devices: A case study using WFEES modules based on the LANS method." Sensors 19, no. 20 (2019)
2. Hamidi, A.A., Robertson, B. and Ilow, J., 2023. A new approach for ECG artifact detection using fine-KNN classification and wavelet scattering features in vital health applications. Procedia Computer Science, 224, pp.60-67. (2023)
3. Singh, Ritu, Rajesh Mehta, and Navin Rajpal. "Efficient wavelet families for ECG classification using neural classifiers." Procedia computer science, pp. 11-21. (2018)
4. Goldberger, A., L. Amaral, L. Glass, J. Hausdorff, P. C. Ivanov, R. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley. "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. e215–e220." (2000).
5. Das, Maumita, and Bikash Chandra Sahana. "Optimized orthogonal wavelet-based filtering method for electrocardiogram signal denoising." Journal of The Institution of Engineers (India): Series B, pp. 1-14. (2022)
6. Montenegro, Larissa, et al. "Human-assisted vs. deep learning feature extraction: an evaluation of ECG features extraction methods for arrhythmia classification using machine learning." Applied Sciences (2022)
7. Rahul, Jagdeep, and Lakhan Dev Sharma. "Artificial intelligence-based approach for atrial fibrillation detection using normalised and short-duration time-frequency ECG." Biomedical Signal Processing and Control (2022).
8. Hammad, Mohamed, Shanzhuo Zhang, and Kuanquan Wang. "A novel two-dimensional ECG feature extraction and classification algorithm based on convolution neural network for human authentication." Future Generation Computer Systems, pp. 180-196. (2019)
9. Pandey, Saroj Kumar, and Rekh Ram Janghel. "ECG arrhythmia classification using artificial neural networks." Proceedings of 2nd International Conference on Communication, Computing and Networking: ICCCN 2018, NITTTR Chandigarh, India. Springer Singapore, (2019).

10. Bishop, A. J., Nehme, Z., Nanayakkara, S., Anderson, D., Stub, D., & Meadley, B. N. (2024). Artificial neural networks for ECG interpretation in acute coronary syndrome: A scoping review. *The American Journal of Emergency Medicine*, *83*, 1-8.
11. Asgharzadeh-Bonab, Akbar, Mehdi Chehel Amirani, and Alaeddin Mehri. "Spectral entropy and deep convolutional neural network for ECG beat classification." Biocybernetics and Biomedical Engineering 40.2 (2020): 691-700.
12. Mandal, Saurav, Pulak Mondal, and Anisha Halder Roy. "Detection of Ventricular Arrhythmia by using Heart rate variability signal and ECG beat image." Biomedical Signal Processing and Control 68 (2021): 102692.
13. Fan, W.; Liu, K.; Liu, H.; Ge, Y.; Xiong, H.; Fu, Y. Interactive reinforcement learning for feature selection with decision tree in the loop. IEEE Trans. Knowl. Data Eng. 2021, 35, 1624–1636.
14. Alotaiby, T. N., Alrshoud, S. R., Alshebeili, S. A., & Aljafar, L. M. ECG-based subject identification using statistical features and random forest. Journal of Sensors, 2019, 1-13. (2019)