Research Article
Hybrid Features and Classifier for Classification of ECG Signal

K. Muthuvel and L. Padma Suresh
Department of Electrical and Electronics Engineering, Noorul Islam University, Kumarakoil
629180, India

Abstract: In this research, we have proposed an efficient technique to classify beat from ECG database. The proposed technique is composed into three stages, 1) pre processing 2) Hybrid feature extraction 3) hybrid feature classifier. The beat signals are initially taken from the physiobank ATM and in the pre-processing stage the beat signals are made suitable for feature extraction. For efficient feature extraction we use hybrid feature extractor. The hybrid feature extraction is done in three steps, i) Morphological based feature extraction ii) Haar wavelet based feature extraction iii) Tri-spectrum based feature extraction. Once the feature is extracted the hybrid classifier is used to classify the beat signal as normal or abnormal. Beat classification studies are conducted on the MIT-BIH Arrhythmia Database using three efficient features like as morphological, wavelet and trispectrum. The beat classification system based morphological information gives an accuracy of 68%, wavelet information gives an accuracy of 78%, trispectrum information gives an accuracy of 70%, combined morphological with wavelet information gives an accuracy of 77%, combined morphological with trispectral information gives an accuracy of 70%. By combining the evidence from both the morphological, wavelet and trispectrum features, an accuracy of 91% is obtained, indicating that ECG beat information is present in the hybrid features.

Keywords: Artificial bee colony, genetic, haar wavelet, hybrid classifier, hybrid feature extraction, tri spectrum

INTRODUCTION

The cardinal function of the Electrocardiograph is the sound management of the electrical activity of the central of the blood circulatory system i.e., the heart. Supervised by keeping sensors at the limb boundaries of the subject, Electrocardiogram (ECG) is a documentation of the source and the Circulation of the electrical potential throughout cardiac muscles (Owis *et al*., 2002). An ECG beat signal contains important information that can help medical diagnosis, reflecting cardiac activity of a patient, if it is normal or failing heart that has certain pathologies (Castillo *et al*., 2012). Now-a-days, Electrocardiogram (ECG) is one of the most effective diagnostic tools to detect heart diseases. Normally, an ECG waveform consists of five basic waves P, Q, R, S and T waves and sometimes U waves. As a rule, an ECG waveform comprises mainly five basic waves P, Q, R, S and T waves, though at times it is complemented by U waves. The P wave symbolizes atria depolarization, Q, R and S waves generally known by the name QRS complex signifies the ventricular depolarization and T wave stands for the re-polarization of ventricle Maglaveras *et al*. (1998).

The ECG beat signal investigation is applied in the identification of many heart ailments such as ischemia, arrhythmias and Myocarditis, or disorder of heart beat or rhythm, or modification in the morphological model and for monitoring drug effects or pacemaker action. Automated analysis of the ECG has been the area of severe research during the last three decades and is recognized as an effective clinical tool in the physiological measurement field. The reason behind the research interest for ECG analysis comes from its role as a proficient non-invasive investigative method that offers valuable information for the detection, diagnosis and treatment of cardiac diseases (Filho *et al*., 2009). The ECG beat-by-beat analysis and classification can provide important information regarding the subject's cardiac condition (Exarchos *et al*., 2007). Therefore, it is desirable to develop computer-based diagnosis tools.

In addition to the noise, the main problem in computer-based classification of heart beats in ECG recordings is the wide variety in the shape of beats belonging to a given class and the similarity in the shape of beats belonging to different classes (Osowski and Linhm, 2001). This phenomenon makes it all the more essential for the algorithms related to computer-based diagnosis to proceed normally through three major stages such as ECG beat recognition, mining of advantageous characteristics from beats and categorization. A number of methods have been developed for beat detection (Afonso *et al*., 1999; Kohler *et al*., 2002). Feature extraction can be done in

Corresponding Author: K. Muthuvel, Department of Electrical and Electronics Engineering, Noorul Islam University, Kumarakoil 629180, India

the time domain (Chazal and Reilly, 2003), in the frequency domain (Acharya *et al.*, 2004), by multi-scale decomposition (Prasad and Sahambi, 2003), by multi-fractal analysis (Ivanov *et al.*, 2009), or by statistical means (Osowski and Linh, 2001). Accordingly, several features like, QRS duration, RR interval, amplitude of P, Q, R, S and T points are extracted by several authors. On the other hand, statistical features like cross correlation, correlation, ICA components were used by the authors to train the classifier. Classification can be performed using neural networks (Ozbay *et al.*, 2006; Osowski *et al.*, 2008; Hu *et al.*, 1993), linear discriminants (Senhadji *et al.*, 1995), learning vector quantization (Hu *et al.*, 1997), a mixture of experts model (Yu and Chou, 2007), or a switchable scheme (Hu *et al.*, 1993).

In this study, a technique is proposed to solve the classification problem. The main objective of this research is to classify beat signal from ECG database. It consists of twofold:

- Three effective features like morphological, wavelet and Tri-spectrum are hybridized and used to feature extraction purpose, which (hybrid) improved the performance than individual feature extraction system.
- Two artificial intelligent techniques like Artificial Bee Colony (ABC) and Genetic algorithm are hybridized and used to classification of the ECG beat signal, which proved as better connection to find optimal weight in feed forward neural network.

LITERATURE REVIEW

In the literature review, several methods have been proposed for the automatic classification of ECG beat signals. Among the most recently published works are those presented as follows: Jiang and Kong (2007) have presented evolvable Block-based Neural Networks (BbNNs) for personalized ECG heartbeat pattern classification. Network structure and connection weights were optimized using an EA that utilized evolutionary and gradient-based search operators. An adaptive rate modification system that demonstrated the efficiency of an operator in turning out superior robustness for the healthy persons in relation to prearranged consistent rates. The GDS operator significantly scaled up the execution and the optimization pace. The BbNN demonstrated a potential to classify ECG heart-beat patterns with a high accuracy for personalized ECG monitoring.

Ubeyli (2009) has described the application of Adaptive Neuro-Fuzzy Inference System (ANFIS) model for classification of Electrocardiogram (ECG) beat signals. Decision making was performed in two stages: feature extraction by computation of Lyapunov exponents and classification by the ANFIS trained with the back propagation gradient descent method in combination with the least squares method. Four types of ECG beats (normal beat, congestive heart failure beat, ventricular tachyarrhythmia beat and atrial fibrillation beat) obtained from the PhysioBank database was classified by four ANFIS classifiers. To improve diagnostic accuracy, the fifth ANFIS classifier (combining ANFIS) was trained using the outputs of the four ANFIS classifiers as input data. The proposed ANFIS model combined the neural network adaptive capabilities and the fuzzy logic qualitative approach. Some conclusions concerning the saliency of features on classification of the ECG beat signals were obtained through analysis of the ANFIS. The efficiency of the ANFIS configuration was assessed in relation to guidance execution and categorization precision. The results confirmed that the proposed ANFIS model has potential in classifying the ECG beat signals.

Kim *et al.* (2010) have proficiently propounded an innovative ECG beat signal processing technique with Quad Level Vector (QLV) for the ECG holter system. The ECG processing consisted of the compression flow and the classification flow and the QLV was presented for both flows to achieve better performance with low computation complexity. The compression algorithm was executed by means of ECG skeleton and the Huffman coding. Unit block size optimization, adaptive threshold adaptation and 4-bit-wise Huffman coding techniques were executed to scale down the processing overhead without any compromise on the beat signal excellence. The heart beat segmentation and the R-peak detection methods were employed for the classification algorithm. The performance was evaluated by using the Massachusetts Institute of Technology-Boston's Beth Israel Hospital Arrhythmia Database and the noise robust test is also per-formed for the reliability of the algorithm.

Ka (2011) has smartly put forward an Electrocardiogram (ECG) beat categorization technique founded on waveform resemblance and RR intervals. The technique categorized six kinds of heart beat, such as normal beats, atrial early beats, paced beats, early ventricular contractions, left bundle branch block beats and right bundle branch block beats. The ECG beat signal was initially cleared of noise by means of wavelet-transform-based methods. Heart beats, sampled at 128 points centered on the R peak, were removed from the ECG beat signal. The number of samples per heart beat was then decreased to 16 to comprise a trait. The RR intervals surrounding the beat were also employed as a feature. A database of interpreted beats was constructed for the classifier for waveform assessment with unidentified beats.

Iftikhar *et al.* (2012) have developed an arrhythmia disorder classifier using Feed forward Back propagation neural network. The supervised network was trained based on the features extracted from the ECG databases

of MIT-BIH. The trained network classified the beats into Premature Atrial Ventricular Contraction (PAC/PVC), Left/Right Bundle Branch Block (LBBB/RBBB), paced beat and normal beat.

Yeh (2012) has proposed a Principal Component Analysis (PCA) and Fuzzy Logic to analyze ECG beat signals for effective determining heartbeat case. It could precisely categorize and recognize the divergence between regular heartbeats (NORM) and irregular heartbeats. An irregular heartbeat comprised Left Bundle Branch Block (LBBB), Right Bundle Branch Block (RBBB), Ventricular Premature Contractions (VPC), Atrial Premature Contractions (APC) and Paced Beat (PB). Assessment of the ECG beat signals proceeds through three significant phases such as:

- Recognizing the QRS waveform
- The qualitative traits choice
- Heartbeat case assessment

The study used Principal Component Analysis for selection of qualitative features and determination of heartbeat case was carried out by fuzzy logic. Records of MIT-BIH database were employed for the purpose of efficiency estimation.

Discussion about the literature review: By considering the above literature review, the researchers have developed several distinct methods to classify the ECG signal. Among them, Jiang and Kong (2007) have developed block based neural networks for classifying the heart beats. The weights are optimized by using evolutionary and gradient based search operators. Ubeyli (2009) introduced Adaptive Neuro-Fuzzy Inference System (ANFIS) model for classifying the ECG beats. In this method four types of ECG beats are classified by using four ANFIS classifier and the fifth ANFIS classifier is trained using the results of four classifiers. Kim *et al*. (2010) have presented quad level vector technique to classify the ECG beat signal. This method involves two processes:

- Compression flow
- Classification flow

For compression algorithm ECG skeleton and Huffman coding is used. For classification algorithm heart beat classification and R-peak detection method was used. Ka (2011) described ECG classification techniques using wavelet resemblance and RR intervals. In this method neighbouring beats of the RR interval is considered as features. Iftikhar *et al*. (2012) have developed a technique to detect the arrhythmias by using arrhythmias classifier and feed forward back propagation algorithm. This method has some deficiency in identifying the P and T waves. Yeh (2012) has proposed principal component analysis and fuzzy logic method to classify the normal and abnormal beats

in ECG. This technique provides accurate classification only for portable devices. Obtaining a good quality model for ECG beat classification depends on many factors, like as feature extraction and classification. Various techniques proposed in the literature have met with only limited success due to complexity of feature extraction and classification. By analyzing the demerits of the related works we develop a proposed technique to classify the beats from the ECG signal. In our proposed techniques we use hybrid feature extraction and hybrid classifier for effective classification of ECG beat. For hybrid feature extraction we use three extraction techniques:

- Morphological feature extraction
- Wavelet feature extraction
- Spectrum based feature extraction which helps to extract the features so accurately from the ECG signal

For effective classification we use two optimization algorithms:

- ABC algorithm
- Genetic algorithm in the hybrid classifier

CLASSIFICATION OF ECG BEATS BY USING HYBRID FEATURE AND HYBRID CLASSIFIER

The ECG beat classification is mainly necessary to categorize the normal and abnormal beat. This classification of normal and abnormal beats helps to spot the heart diseases. Several techniques are followed to classify the ECG beats. In our proposed method we use a hybrid technique for both feature extraction and classification. This involves three major steps they are:

- Pre-processing
- Hybrid feature extraction
- Hybrid feature classification

In the pre processing stage the input ECG beats are taken from physio bank ATM. In hybrid feature extraction stage, features are extracted by three processes:

- Morphological based feature extraction
- Wavelet based feature extraction
- Spectrum based feature extraction

Once the feature is extracted then the hybrid classifier is used to classify the ECG beat signal. In the hybrid classifier, ABC algorithm is combined with genetic algorithm for training the neural network. The trained neural network by the hybrid ABC and genetic algorithm is used to classify the beat. The overall diagram of proposed ECG beat classification is depicted in Fig. 1.
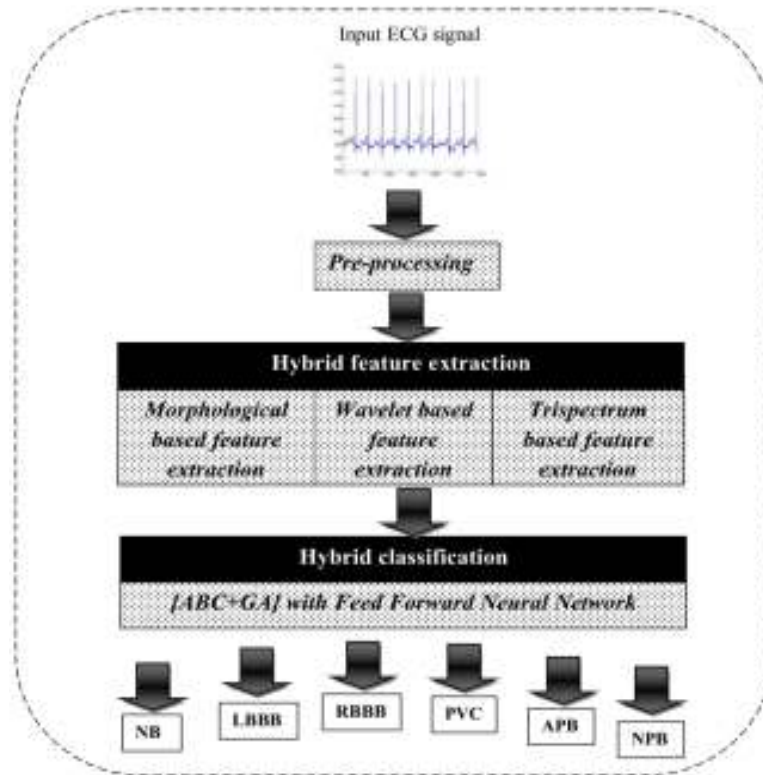
Fig. 1: Overall block diagram of proposed technique

Fig. 2: Snapshot representation of physiobank ATM

Pre-processing: In this pre processing step, we make the beat signals suitable for feature extraction. In this step the ECG beat signal input dataset is taken from the physiobank ATM (http://physionet.org/cgi-bin/atm/ATM). To generate the dataset, in ATM input control panel select an input beat signal from the database. For our work we have chosen MIT-BIH Arrhythmia Database. After choosing the database, beat signal is selected by choosing corresponding record value. Once we select the type of database, the annotations type gets changed automatically. Annotations are labels that used to describe the time of occurrence and the type of individual heart beats at a particular point within the recording. After choosing the beat records change the length in the output panel as 'the end'. Then choose the tool in the toolbox such as 'Show annotations as text', 'export beat signal as .mat'. Then for each tool store the corresponding dataset. Similarly for five abnormal and normal beat signals the dataset is obtained by changing the record value in the input panel. Then obtained input dataset values are transferred for feature extraction. The snapshot representation of physiobank ATM[1] is illustrated in Fig. 2.
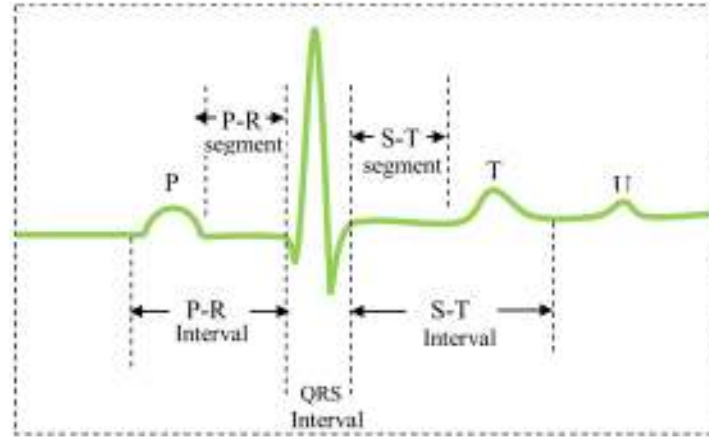
Fig. 3: ECG beat signal labelled by P, Q, R, S, T

Hybrid feature extraction: Computer-based diagnosis algorithms have usually three stages, like as ECG beat detection, extraction of useful features from beats and classification.

ECG beat detection: After pre-processing step, the MATLAB files are given into the input of feature extraction for PQRST detection. After we point out P, Q, R, S and T points of ECG signal based on the peak value. The PQRST characteristics points are recognized by peak detector and following ways:

- The normal deflection of the P wave is upright (positive).
- The first deflection, if it is negative (downward), is labeled the Q wave.
- The first positive (upright) deflection is labeled the R wave, whether it is predicated by a Q wave or not.
- A negative deflection following an R wave is labeled an S wave.
- Any subsequent waveforms are labeled "primes". Positive being R', R" etc.... Negative being S', S".
- The direction of the normal adult T wave is upright.

After PQRST point's detection, useful features are extracted from beats to classify the ECG. Hybrid feature techniques are used for effective feature extraction which involves:

- Morphological feature extraction
- Wavelet based feature extraction
- Spectrum based feature extraction

Morphological based feature extraction: After characteristic point's detection, the ECG beat signals are subjected to morphological centred feature extraction. Morphological feature extraction is a method to extract features that helps to describe the representations. Generally there are five deviations in ECG beat signal they are P, Q, R, S and T as shown in the Fig. 3. Morphological feature extraction is done in three steps. First we need to find the standard deviation of RR interval, PR interval, PT interval, ST interval, TT interval, QT interval. Second the maximum values of P, Q, R, S, T peaks are obtained and finally the number of R peaks count is taken.

P-R interval: The initial modulation point in the ECG beat signal is the P wave. The P-R interval is attained by computing the phase from the starting off P wave to the QRS wave. The maximum rate of the P-R interval is 200 msec. We mainly extract two features from the P-R interval:

- Mean value
- Standard deviation

Let the P-R interval values are denoted as P_R and the mean value of P-R intervals is given as:

$$\mu_{PR} = E[P_R] \tag{1}$$

where,
μ_{PR} = The mean value of P-R intervals
$E[P_R]$ = The expected time value of P-R interval

From the mean value the standard deviation of the P-R interval can be calculated using the following equation:

$$\sigma_{PR} = \sqrt{E[P_R^2] - (E[P_R])^2} \tag{2}$$

where, σ_{PR} is the standard deviation of P-R intervals.

P-T interval: The P-T interval is attained by computing the duration from the starting off P wave to the T wave. Mean value and standard deviation are two

features that are mainly extracted from the P-T interval. The mean value of the P-T intervals is given as:

$$\mu_{PT} = E[\mathrm{P}_T] \tag{3}$$

where,
μ_{PT} = The mean value of P-T interval
P_T = The randomly obtained P-T intervals
$E[P_T]$ = The expected period of P-T intervals

Therefore the standard deviation of P-T intervals is given as:

$$\sigma_{PT} = \sqrt{E[\mathrm{P}_T{}^2] - (E[\mathrm{P}_T])^2} \tag{4}$$

where, σ_{PT} is the standard deviation of P-T interval.

S-T interval: The negative swerve next to the R wave in the ECG beat signal is S wave, where T wave is the refraction next to the S wave. The S-T interval is obtained by evaluating the period from the offset of QRS wave to the start off T wave. The S-T interval usually get varies from 80 to 120 msec. Mean value and standard deviation are two features that are mainly extracted from the S-T interval. The mean value of the S-T intervals is given as:

$$\mu_{ST} = E[\mathrm{S}_T] \tag{5}$$

where,
μ_{ST} = The mean value of S-T interval
S_T = The randomly obtained S-T intervals
$E[S_T]$ = The expected period of S-T intervals

Therefore the standard deviation of S-T intervals is given as:

$$\sigma_{ST} = \sqrt{E[\mathrm{S}_T{}^2] - (E[\mathrm{S}_T])^2} \tag{6}$$

where, σ_{ST} is the standard deviation of S-T interval.

Q-T interval: The initial negative change in the ECG beat signal is Q wave which helps to sense the depolarization and re-polarization of left and right ventricles. The interval between the Q and T wave is obtained by evaluating the phase from the inception of Q wave to the offset of T wave. Then the standard deviation of the Q-T interval is obtained by taking the square root of the mean value, so, to find the standard deviation we need to calculate the mean value. The mean value can be calculated by taking the expected value of the intervals:

$$E[\mathrm{Q}_T] = \mu_{QT} \tag{7}$$

where,
μ_{QT} = The mean value
$E[Q_T]$ = The expected time value of Q-T intervals:

$$\sigma_Q = \sqrt{E[\mathrm{Q}_T{}^2] - (\mathrm{E}[\mathrm{Q}_T])^2} \tag{8}$$

where, σ_Q is the standard deviation of Q-T intervals.

Maximum peak values: The change of the waves from the positive to negative side gives the peak value. Generally each cardiac cycle of heart has five changes namely P, Q, R, S, T. P wave is generated in the ECG beat signal when the electrical vector shift from the right to left atrium. Similarly transferring of electrical beat signals through myocardial creates Q wave, which help to spot the depolarization in inter-ventricular septum. R wave has the ultimate peak value in the ECG beat signal which mainly helps to spot the arrhythmia. The negative change next to the R wave in the ECG beat signal is S wave. The re-polarization of ventricles produce T wave which helps to sense the coronary ischemia, left ventricular hypertrophy etc. After identifying the P, Q, R, S, T waves in ECG beat signal, the maximum peak values P_{peak}, Q_{peak}, R_{peak}, S_{peak} and T_{peak} wave are calculated.

R-peak count: After calculating the maximum peak values of P, Q, R, S, T wave, number of R peak count R_N in the ECG beat signal is taken. This is used for extracting the morphological features. The features extracted from the morphological feature extraction are detailed in the Table 1:

$$R_N = N[R_{peak}] \tag{9}$$

where, N $[R_{peak}]$ is the number of R peaks count.

Wavelet based feature extraction: The processing of the information by the heart is reflected in dynamical changes of the electrical activity in time, frequency and space. Mostly features in time (Jekova *et al.*, 2008) and frequency (Khadra and Binajjaj, 2005) were extracted and combined with efficient classifiers. Time domain features fail to detect arrhythmias contaminated with noise. Frequency domain features are suitable to pursue variations in different kinds of ECG beats; however no information of time localized features can be presented by them. Time-frequency domain features are of great importance today, which can give information in both time and frequency domain. They are also useful in investigation of non-stationary signals similar to ECG. Haar wavelet transform is capable of detecting and characterizing specific phenomena in time and frequency planes. The haar wavelet is a smooth and quickly vanishing oscillating function with a good localization in both frequency and time. In wavelet transform the feature extractions were made in two steps:

Table 1: Steps of morphological based feature extraction

Morphological based features extraction

| Steps	Feature	
Step 1	P-R interval	Standard deviation of P-R interval σ_{PR}
	P-T interval	Standard deviation of P-T interval σ_{PT}
	S-T interval	Standard deviation of S-T interval σ_{ST}
	Q-T interval	Standard deviation of Q-T interval σ_{QT}
Step 2	P wave	Maximum peak value of P wave P_{peak}
	Q wave	Maximum peak value of Q wave Q_{peak}
	R wave	Maximum peak value of R wave R_{peak}
	S wave	Maximum peak value of S wave S_{peak}
	T wave	Maximum peak value of T wave T_{peak}
Step 3	R peak count	Number of R peaks count R_N
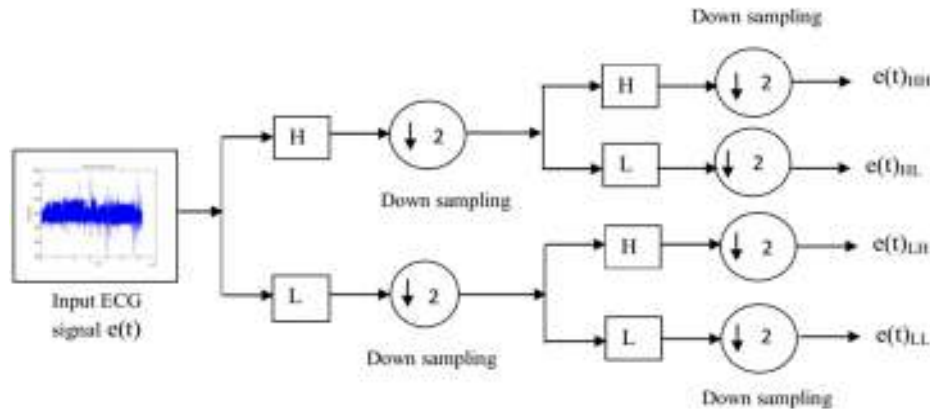
Fig. 4: Architecture of two dimensional haar wavelet decomposition occurring at level 1

- Depending on distinct frequency sub bands the ECG beat signals are decomposed.
- The disintegrated beat signals at distinct frequency sub bands are evaluated using numerous resolutions.

For the ECG beat signal $e(t)$ the wavelet transform is given as:

$$W(p,q) = \int_{-\infty}^{\infty} e(t).\Psi p,q(t)dt \qquad (10)$$

where, $\Psi p,q(t)$ is the wavelet function.

In our proposed technique we use two dimensional Haar wavelet transform because it reduces the computational time and also it extracts more features. For the t input beat signal u_t the haar wavelet transform v_t is given as:

$$v_t = H_t u_t \qquad (11)$$

In Haar wavelet the ECG beat signals are get disintegrated into coarse approximation and detail information. For this disintegrated two filters:

- Low pass filter
- High pass filter are used

First the ECG beat signals are passed to Low pass filter which screen the low frequency beat signals less than the cut off frequency. Second the ECG beat signals are passed to High pass filter which screen the high frequency beat signals beyond the cut off frequency. The resultant beat signal from the Low pass filter is down sampled by 2 gives coarse coefficients and the resultant beat signal from the high pass filter is down sampled to produce detail coefficients, which is shown in the Fig. 4.

Then the mean value of the coarse coefficients is calculated by taking the average of the coarse coefficient:

$$E[a_t] = \mu_{a_t} \qquad (12)$$

where, μ_{a_t} is the mean value for approximation coefficient.

From the mean value the standard deviation of the coarse coefficients is measured by taking the square root of the mean value:

$$\sigma_{a_t} = \sqrt{E[a_t] - (E[a_t])^2} \qquad (13)$$

where, σ_{a_t} is the standard deviation for approximation coefficient.

Extraction of nonlinear activities through tri spectrum: The objective of this section is to apply tri-

spectral analysis for classification of ECG beat. Generally, the Electrocardiogram (ECG) signal is the recording of the bioelectrical activities of the cardiac system. Therefore, it is believed that several sources that are individually coherent will emit their electrical impulses at the instant of arrhythmia. The overall response of these electrical activities will be accumulated as an ECG signal that is composed of incoherent sources. Analyzing the nonlinear activities of the ECG beat by utilizing the ability of high order statistic techniques like as tri-spectrum in revealing hidden details of these beats could improve their detection and further detect the instance of their occurrences. Therefore, a nonlinear time invariant process is used to classify the ECG beat. Here, we have used tri-spectrum to extract phase coupling information or features from the ECG signal for detecting and classifying different types of beats. Tri-spectrum is type of statistics used to identify the output beat signal which is not directly proportional to the corresponding input beat signal. Tri spectrum can also obtain by taking Fourier transform for fourth cumulant function of random process u_n (Ge *et al*., 2002). Let v_1, v_2, v_3 be the three frequencies then the Tri spectrum of three frequencies is given as:

$$T_s(v_1, v_2, v_3) = E[U(v_1)U(v_2)U(v_3)U*(v_1 + v_2 + v_3)] \qquad (14)$$

where,
v = The frequency
$U(v)$ = The Fourier transform of random process u_n
$E[]$ = The expectation factor
$*$ = The complex conjugation

In the tri-spectrum method, expectation of the frequency is obtained by taking the average of the three frequencies. It is a three dimensional structure. The resultant image is obtained in 512*512 pixels. The tri-spectrum was applied to obtain a tri-spectrum plot and magnitude tri-spectrum.

Steps involved in the feature extraction of tri spectrum:

- By using the number of pixel present in the row, column and both diagonals we can calculate the centre point of the pixel.
- The properties of the region near to the centre point of tri-spectrum can be calculated by using the region props function in the mat lab. The region props function helps to evaluate the orientation, eccentricity, solidity, extent and perimeter values of the region.

By using tri spectrum based feature extraction we extract 12 features from the input ECG beat. To estimate the tri-spectrum of a signal the following processing steps are performed.

Location of maximum values: The location of maximum values can be obtained by computing the maximum values from both row and column. Thus we get two maximum values from both horizontal r_{max} and vertical line c_{max}.

Sum of diagonal values: In this we calculate the sum of all the values present in the diagonals from both the direction i.e., left diagonal Ld_{sum} and right diagonal Rd_{sum}.

Sum of centre column Cc_{sum} and centre row Cr_{sum} values: This value can be evaluated by computing the sum of the values present in the centre row Cr_{sum} and the values present in the centre column Cr_{sum}.

Orientation of region O_r: The orientation of the image O_r is obtained by measuring the angle between the horizontal axis and the major axis of the ellipsoid.

Eccentricity of region E_t: The eccentricity of the region E_t can be calculated by taking the relation of distance between the foci of ellipse and the major axis.

Solidity of region S_o: Solidity is the ratio of region area to the convex area of the region. The solidity of the region is calculated using the formula:

$$S_o = A/CA \qquad (15)$$

where,
A = The Area of the region
CA = The convex area of the region
S_o = The solidity

Extent of region E_x: Extent of the region E_x is measured by taking the relation of region area to the bounding box area which is given by the formula:

$$E_x = A/BA \qquad (16)$$

where,
A = The Area of the region
BA = The bounding area of the region
E_x = The solidity

Perimeter of region P_r: The perimeter of the region P_r can be measured by taking the number of neighbouring pixel of the region and calculating the space between the adjacent pixels which lies in the border of the region.

Entropy of tri spectrum T_s: Entropy is determined by taking the probability of a process or information content. The entropy for tri spectrum T_s is calculated using the formula:

Table 2: 12 different features using tri-spectrum analysis

Spectrum based features extraction

Steps	Features
Location of maximum values	Maximum values for row r_{max}
	Maximum values for column c_{max}
Sum of diagonal values	Sum of left diagonal Ld_{sum}
	Sum of right diagonal Rd_{sum}
Sum of centre column and centre row values	Sum of centre column Cc_{sum}
	Sum of centre row Cr_{sum}
Orientation	Orientation of region O_r
Eccentricity	Eccentricity of region E_t
Solidity	Solidity of region S_o
Extent	Extent of region E_x
Perimeter	Perimeter of region P_r
Entropy	Entropy of tri spectrum

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i) \qquad (17)$$

where, $p(x_i)$ is the probability mass function.

The features extracted from the tri spectrum based feature extraction are detailed in the Table 2.

ECG beat classification using hybrid classifier: After extracting the features from ECG beat signal we use hybrid classifier is used to classify the ECG beat signal. In the hybrid classifier we use both ABC algorithm and genetic algorithm to train the beat signals in the neural network. For training purpose five abnormal beat signals are used along with the normal beat signal. The five abnormal beat signals includes Left Bundle Branch Block Beat (LBBB), Right Bundle Branch Block Beat (RBBB), Premature Ventricular Contraction (PVC), Atrial Premature Beat (APB) and Nodal (junctional) Premature Beat (NPB). The hybrid classifier involves the following steps.

Feed forward neural network layer generation: In the artificial neural network number of neurons required in the output layer depends on the target solutions in each sequence. Initially we generate an output layer model to optimize the weights. The output layer model is given as:

$$v_p^{(1)} = t \sum_{p=1}^{n} w_{pq}^{(1)} u_p + \theta_p$$

$$v_p^{(2)} = t \sum_{p=1}^{n} w_{pq}^{(2)} v_p^{(1)} + \theta_p$$

. .
. .
. .

$$v_p^{(m)} = t \sum_{p=1}^{n} w_{pq}^{(m)} v_p^{n-1}$$

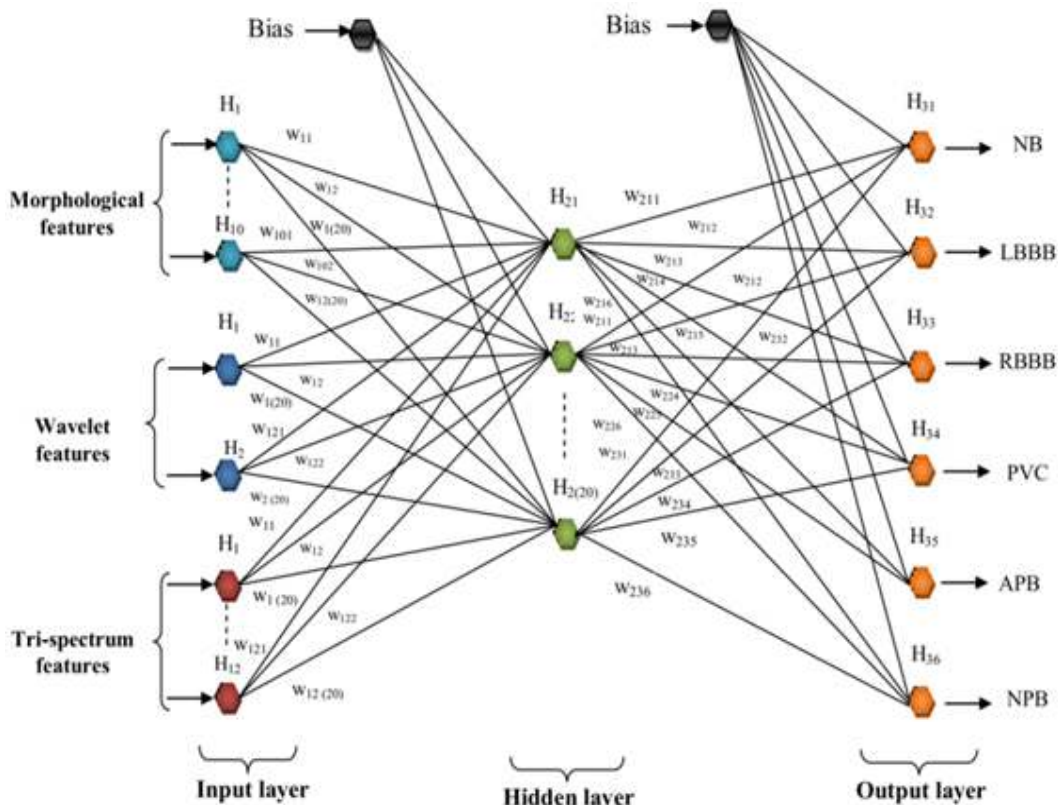
Fig. 5: Proposed feed forward neural network model

where,

p = The input model
q = The hidden mode
v_p = The output of the mode
t = The transfer function
θ = The threshold value
w_{pq} = The weight between the input and hidden node

Training phase of FFNN: In our proposed method five abnormal beat signals are trained along with normal beat signals by using both ABC and genetic algorithm. Most of the neural network uses back propagation algorithm for training but it has many disadvantages such as it consumes more time to find the minimum error. To avoid these defects we use both ABC and genetic algorithm. In our proposed method first we train the FFNN by using the ABC algorithm. Then we use genetic algorithm to train the FFNN by initializing the optimized weight obtained from ABC algorithm. For training the neural network 24 features extracted from the hybrid feature extraction is given as input layer. After optimizing the weights used in the hidden layer the six beats of ECG beat signals are get grouped in the output layer, which is described in the following Fig. 5.

Step 1: Initialization of population: In neural network data are mainly trained to optimize the weight and to detect the minimum error. To optimize the weights, ABC algorithm initially creates arbitrary population of solution:

$$u_p^q = u_{\min}^q + rand\,(0,1)\left(u_{\max}^q - u_{\min}^q\right),$$
$$q \in \{1, 2, ..., w\} \tag{18}$$

where,
u_p = The random solution
w = The weight of the network

Step 2: Fitness evaluation: After the generating the initial population, the fitness of the solution is evaluated. The fitness of the solution is determined by calculating the error between the target and the output obtained:

$$E_v = (\mathrm{v}_p)t\,\arg et - (\mathrm{v}_p)output \tag{19}$$

where,
E_v = The error
$(v_p)t\,\arg et$ = The target value
$(v_p)output$ = The output

Step 3: Modification of food source by employed bee: After initializing the solution, employed bees are allowed to search the neighbouring

food source. Employed bee examine the nectar quantity i.e., fitness of the new food source. Based upon the nectar quantity and the visual information employed bee update the food source. The search of new solution is performed by the following equation:

$$v_p^q = u_p^q + \phi_p^q\left(u_p^q - u_r^q\right), q \in \{1, 2, ..., w\},$$
$$r \in \{1, 2, ..., \mathrm{N}\} \tag{20}$$

where,
$r \neq p, \mathrm{r}, \mathrm{q}$
u_{p}^q = The new food source
ϕ_p^q = The random number between the range (-1, 1)

Step 4: Fitness selection by onlooker bee: After collecting the information of new food source all the employed bees return to the hive and reveal the data's to the onlooker bees. The onlooker bees examines all the food source data's provided by the employed bees and select an food source based upon the nectar quantity and distance. Then onlooker bee stores the new food source in its memory and forgets the exiting data. Mean while the employed bees check the selected food source and updates its memory.

Step 5: Generation of new food source by scout bee: After little iteration if there is no change in the food source location, the scout bees are allowed to find the new food source. If the nectar amount of the new food source is high, then the onlooker bee forget the old one and store the new food source in the memory. If the nectar quantity is not high no modification is made. In our method to generate random solution we use cross over technique. In this step we randomly generate solution based upon on the optimized weight obtained from ABC algorithm. To generate random solution pair we use cross over technique in which the cross over rate is multiplied with the length of the solution.

Step 6: Cross over: In this step we randomly generate solution based upon on the optimized weight obtained from ABC algorithm. To generate random solution pair we use cross over technique in which the cross over rate is multiplied with the length of the solution. In our proposed work we take chromosome length as 10 and cross over rate as 0.2. After generating random solution the fitness value of the solution is examined and the best solution is given for testing.
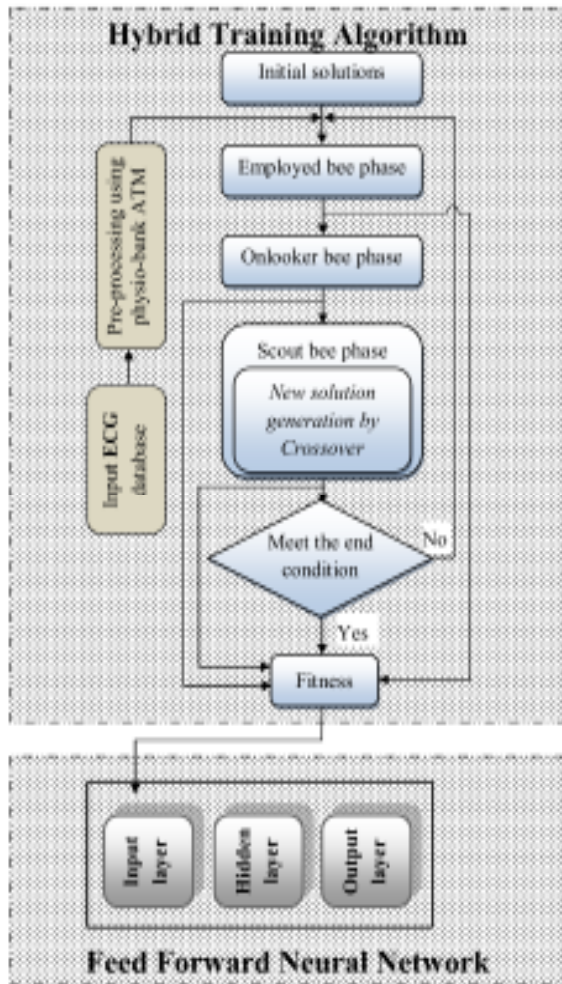
Fig. 6: Training and testing phase of proposed hybrid classifier

Testing phase of FFNN: The selected solution from the training is taken as input to the feed forward neural network. In this step the ECG beat signal is compared with the trained beat signal and classified based upon the feature. The training and testing of the ECG beat signal is shown in the Fig. 6.

RESULTS AND DISCUSSION

Dataset description: In our proposed study we have chosen MIT-BIH Arrhythmia Database from the physiobank ATM. MIT-BIH Arrhythmia Database is the collection of 4000 long term Holter recording taken from the arrhythmia laboratory during the years 1975 to 1979. About 60% of the recording was taken from inpatients. Every single of the 48 recording in the database is over 30 min long. The source of the ECG beat signal in the database is taken from 25 men aged from 32 to 89 years and 22 women aged from 23 to 89 years. The Arrhythmia Database contains almost

109,000 beats. In our proposed work we have taken six beats including normal beat, abnormal beat such as Left Bundle Branch Block Beat (LBBB), Right Bundle Branch Block Beat (RBBB), Premature Ventricular Contraction (PVC), Atrial Premature Beat (APB) and Nodal (junctional) Premature Beat (NPB). For our proposed work we have taken 45 beat signals in the dataset.

Experimental setup: For implementing the propose technique we have used MATLAB version (7.12). This proposed technique is done in windows machine having Intel Core i5 processor with speed 1.6 GHz and 4 GB RAM. For comparing the performance MIT-BIH Arrhythmia Database is used.

Evaluation metrics: The evaluation of proposed ECG beat classification technique in MIT-BIH Arrhythmia Database are carried out using the following metrics as suggested by below equations.

Sensitivity: The sensitivity of the feature extraction and the feature classification is determined by taking the ratio of number of true positives to the sum of true positive and false negative. This relation can be expressed as:

$$S_t = \frac{T_p}{T_p + F_n}$$

where,
S_t = The sensitivity
T_p = The true positive
F_n = The false negative

Specificity: The specificity of the feature extraction and the feature classification can be evaluated by taking the relation of number of true negatives to the combined true negative and the false positive. The specificity can be expressed as:

$$S_p = \frac{T_n}{T_n + F_p}$$

where,
S_p = The specificity
T_n = The true negative
F_p = The false positive

Accuracy: The accuracy of feature extraction and the feature classification can be calculated by taking the ratio of true values present in the population. The accuracy can be described by the following equation:
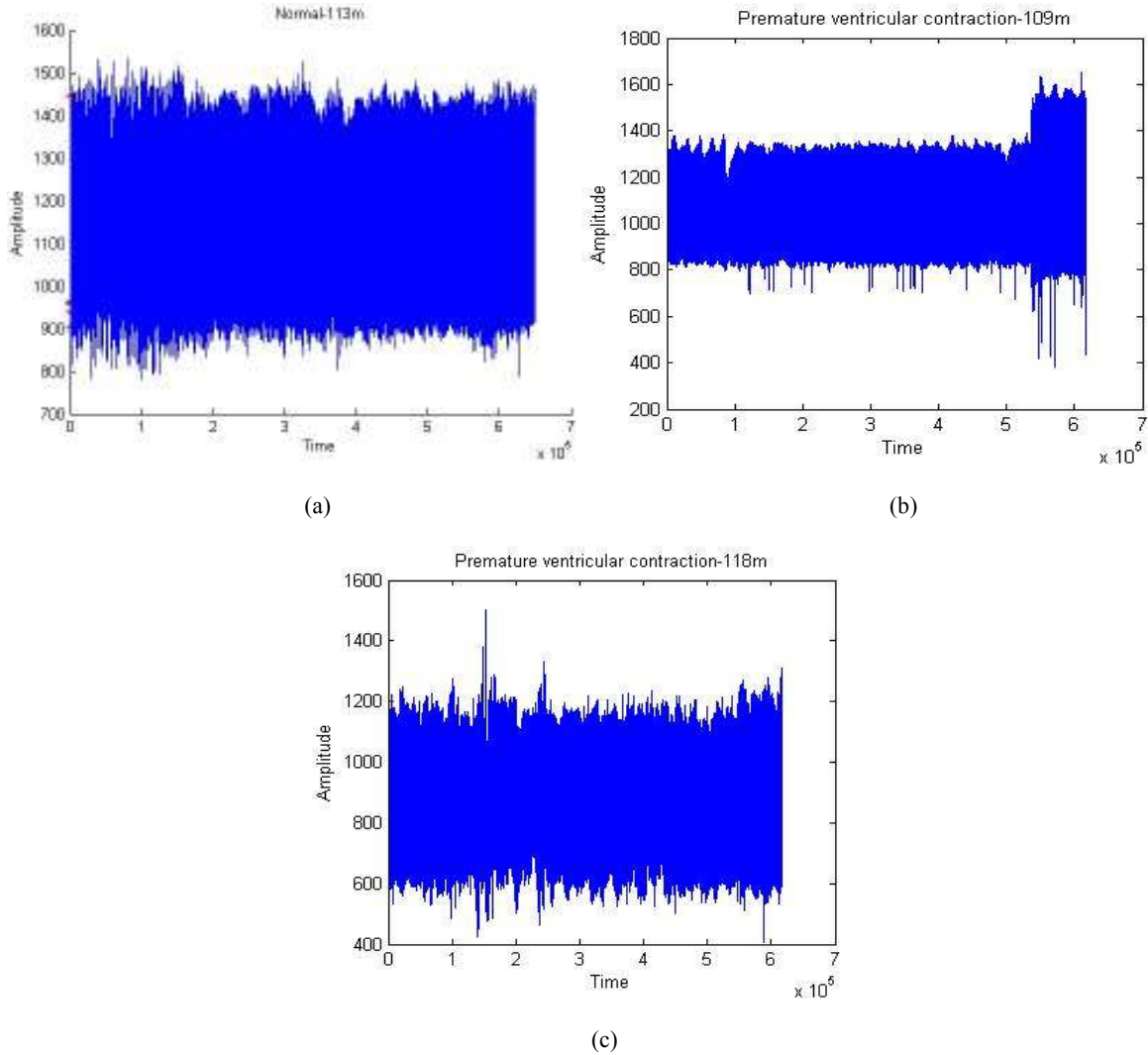
(a)

(b)

(c)

Fig. 7: Input ECG beat signals

Table 3: Hybrid algorithm parameter

Length of the chromosome	Colony size	Q$_i$	Cross over rate	Cycle
10	Hidden neuron * (no of input +1)	0.5	0.2	10

Table 4: Neural network parameters

| NN1	No. of iteration	Error	Training algorithm	No of hidden layer	No of neuron
	10	0.0134	(ABC+GA)	1	20

$$A = \frac{T_p + T_n}{T_p + F_p + F_n + T_n}$$

where,
A = The accuracy
T_p = The true positive
T_n = The true negative
F_p = The false positive
F_n = The false negative

**Experimental result:** In this section we discuss the experimental result obtained by using hybrid feature extraction and classification for grouping the beats in the ECG beat signals. For hybrid feature extraction we use three extraction techniques. Figure 6 shows the three sample input ECG beat signals taken for classification. Figure 7 shows the marked P, Q, R, S, T for input beat signal. The tri spectrum plot of the input beat signal is shown in the Fig. 8. From this tri spectrum the features like location of maximum values, sum of diagonal values, sum of centre column and centre row, orientation, eccentricity, solidity, extent and perimeter are extracted. Table 3 gives the hybrid algorithm parameters used for classification. The parameter used in the neural network for training and testing the ECG beat signal is detailed in the Table 4 (Fig. 9).
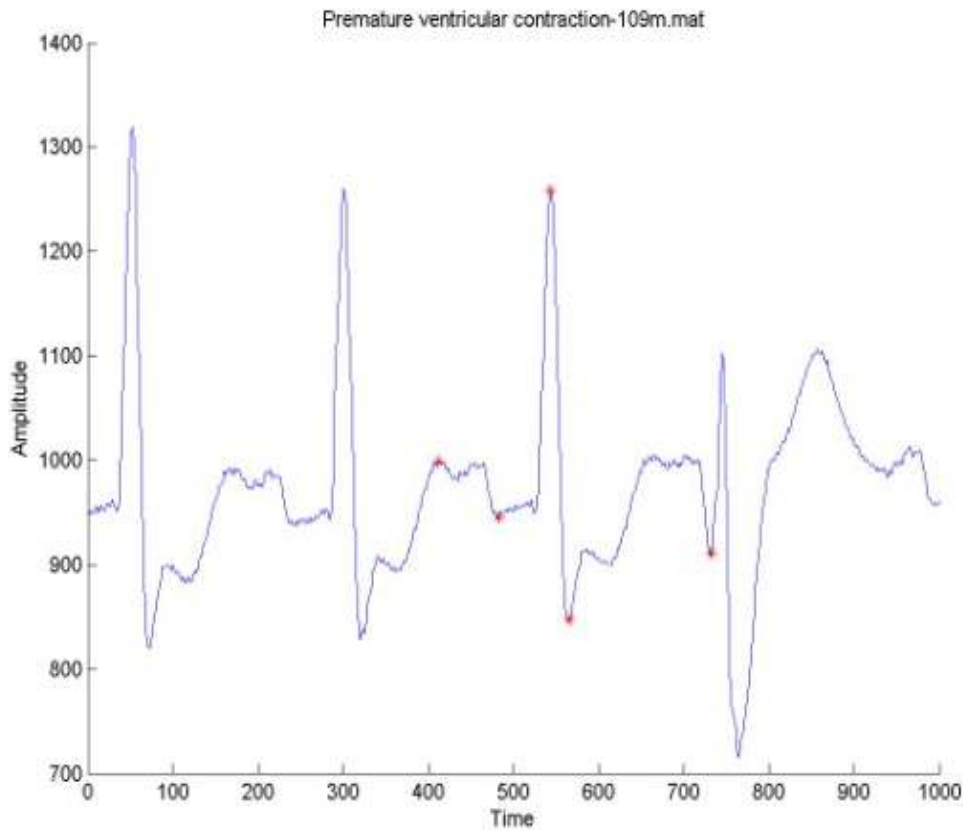
(a)
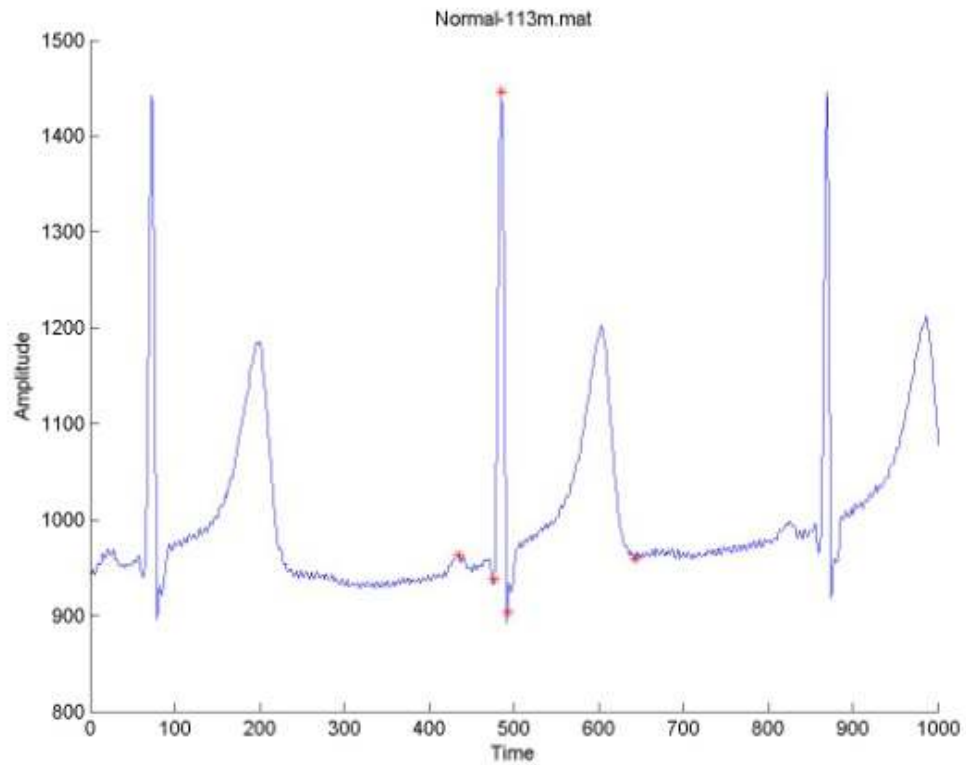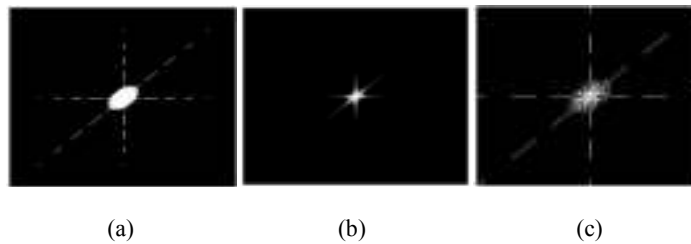

(b)
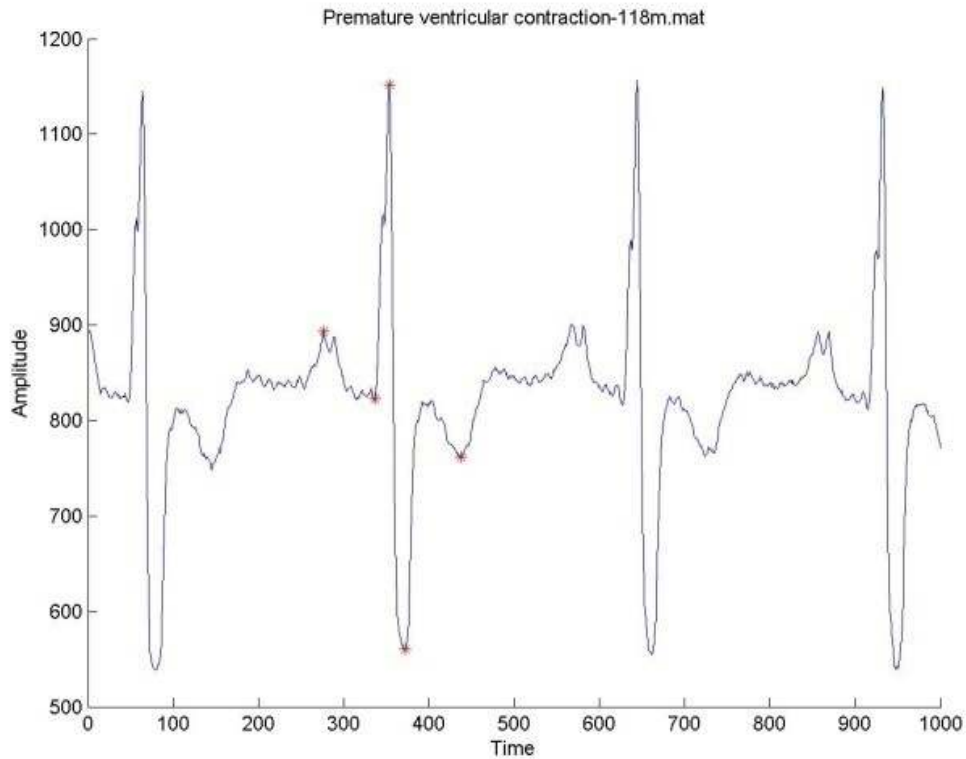
(c)

Fig. 8: Labelled P, Q, R, S, T for input beat signal

(a)   (b)   (c)

Fig. 9: Tri spectrum plot

Comparative analysis: In this section hybrid feature extraction and hybrid classification of our proposed technique is compared separately with the existing method. This comparison is done to prove the efficiency of our proposed method.

Effectiveness of hybrid feature extraction: To prove the effectiveness of hybrid feature extraction, we compare our proposed technique against individual feature extraction method. The evaluation graphs of the sensitivity, specificity and the accuracy graph are shown in Fig. 10. From the Fig. 10a, the proposed ECG beat classification technique achieved the overall accuracy value of 91% which is high compared with the accuracy of existing systems such as morphological feature (Morp + GABC) is achieved 68%, wavelet based feature (Wavelet + GABC) is achieved 78%, tri-spectrum based features (spect + GABC) achieved

70%, combined morphological and wavelet based features (Morp + Wavelet + GABC) is achieved 77%, combined morphological and tri-spectral features (Morp + Spect + GABC) is achieved only 70%. From the graphs it is clear that the hybrid feature extraction shows high accuracy when compared with the existing techniques (Fig. 10a, b and c).

Effectiveness of hybrid classification: To additional effectiveness of our proposed technique, we compare against feed forward with backpropogation classifier. Therefore, we have compared our proposed technique against and existing techniques which are (Hybrid + FFBN), (Morp + FFBN), (wavelet + FFBN), (Spect + FFBN) (Morp + Wavelet + FFBN), (Morp + Spect + FFBN). The evaluation graphs of the sensitivity, specificity and the accuracy graph are shown in Fig. 11. From the Fig. 11a, the proposed ECG beat classification
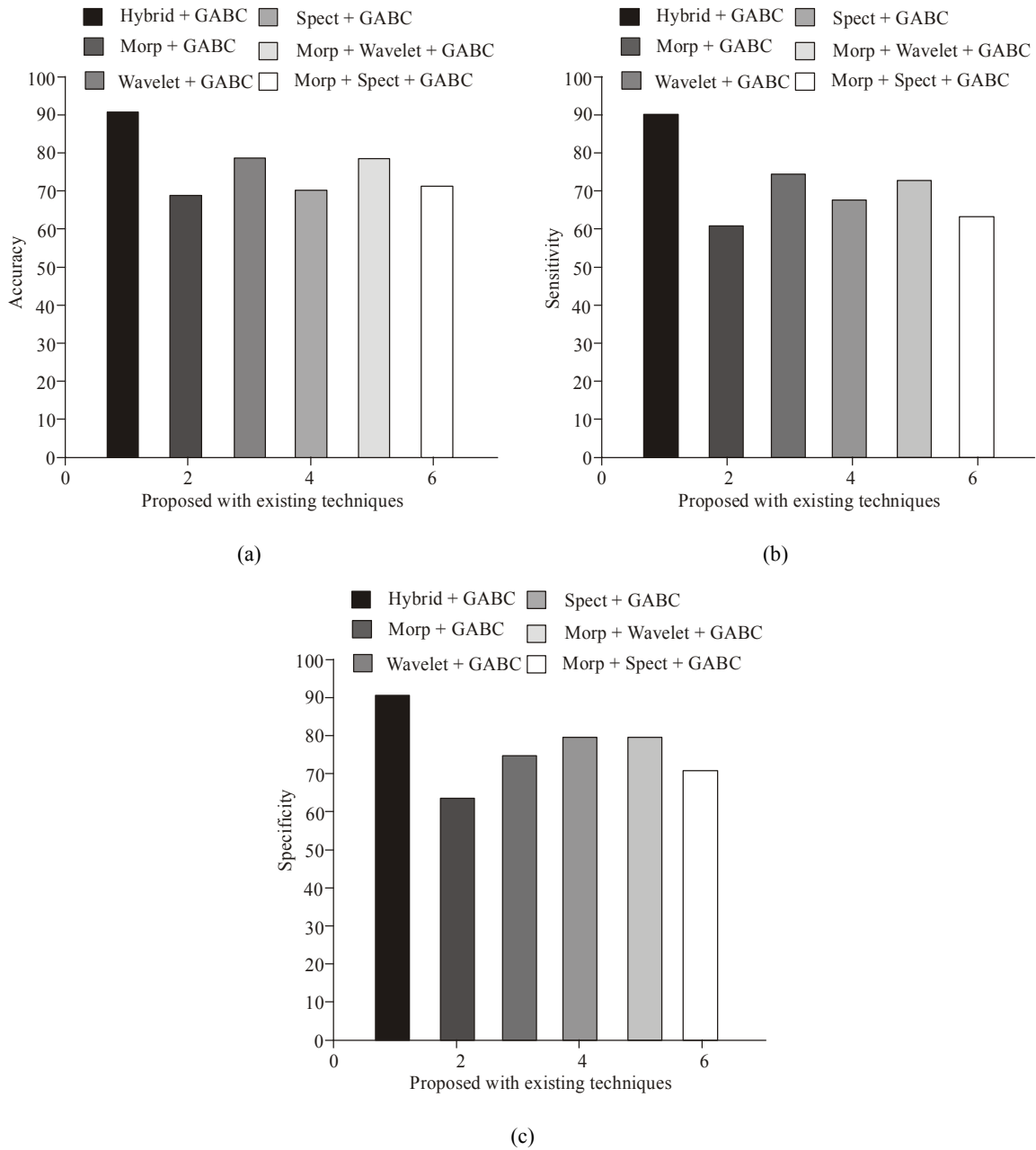
(a)

(b)

(c)

Fig. 10: (a) Comparative analysis graph of accuracy for hybrid feature and existing methods (b) comparative analysis graph of sensitivity for hybrid feature and existing methods (c) comparative analysis graph of specificity for hybrid feature and existing methods

technique achieved the overall accuracy value of 91% which is high compared with the accuracy of existing systems such as (Hybrid + FFBN) is achieved only 78%, (Morp + FFBN) is achieved 62%, (wavelet + FFBN) is achieved 65%, (Spect + FFBN) is achieved 70%, (Morp + Wavelet + FFBN) is achieved 62%, (Morp + Spect + FFBN) is achieved only 73%. Totally, the proposed ECG beat classification technique is achieved better performance when compared existing techniques (Fig. 11a, b and c).

CONCLUSION

The objective of this research was to demonstrate the complementary nature of tri-spectrum features and wavelet features and to show that this information indeed helps in improving the performance of the conventional systems based on morphological features. It was demonstrated by conducting beat classification experiments on the MIT-BIH Arrhythmia Database. The ECG beat classification system using only
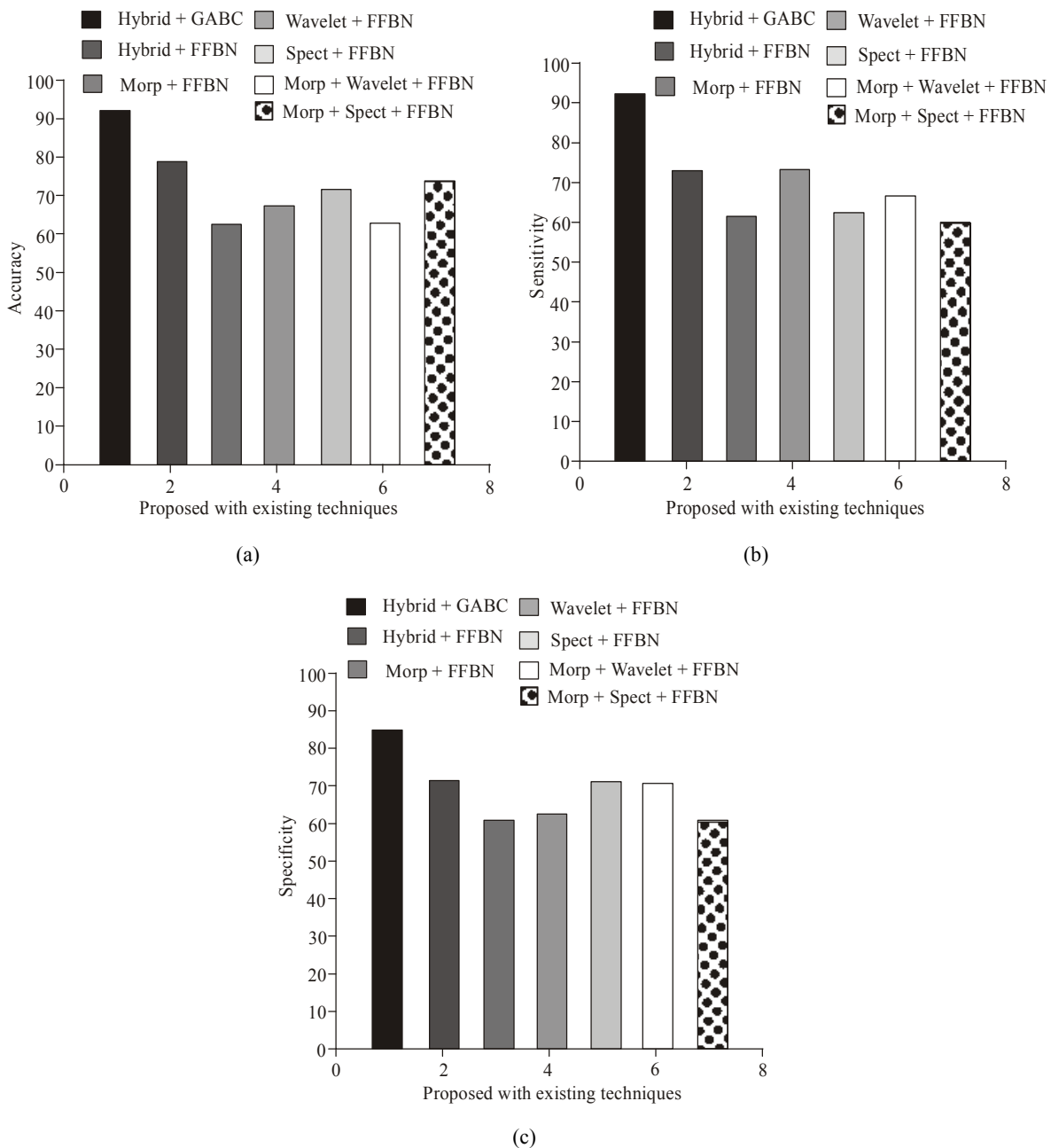
(a)

(b)

(c)

Fig. 11: (a) Comparative analysis graph for accuracy between hybrid classifier and existing methods (b) comparative analysis graph for sensitivity between hybrid classifier and existing methods (c) comparative analysis graph for specificity between hybrid classifier and existing method

morphological information resulted in an accuracy of 68%, wavelet information resulted in an accuracy of 78%, tri-spectrum information resulted in an accuracy of 70%, combined morphological with wavelet information resulted in an accuracy of 77%, combined morphological with tri-spectral information resulted in an accuracy of 70%. However, the combined information led to an accuracy of 91%, which is significantly better than both of the individual information or system.

REFERENCES

Acharya, R., A. Kumar, P. Bhat, C. Lim, S. Iyengar, N. Kannathal and S. Krishnan, 2004. Classification of cardiac abnormalities using heart rate beat signals. Med. Biol. Eng. Comput., 42: 288-293.

Afonso, V.X., W.J. Tompkins, T.Q. Nguyen and S. Luo, 1999. ECG beat detection using filter banks. IEEE T. Bio-Med. Eng., 46: 192-202.

Castillo, O., P. Melin, E. Ramírez and J. Soria, 2012. Hybrid intelligent system for cardiac arrhythmia classification with fuzzy K-nearest neighbors and neural networks combined with a fuzzy system. Expert Syst. Appl., 39: 2947-2955.

Chazal, P.D. and R.B. Reilly, 2003. Automatic classification of ECG beats using waveform shape and heart beat interval features. Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing, 2: 269-272.

Exarchos, T.P., M.G. Tsipouras, C.P. Exarchos, C. Papaloukas, D.I. Fotiadis and L.K. Michalis, 2007. A methodology for the automated creation of fuzzy expert systems for ischaemic and arrhythmic beat classification based on a set of rules obtained by a decision tree. Artif. Intell. Med., 40(3): 187-200.

Filho, E.B.L., N.M.M. Rodrigues, E.A.B.D. Silva, M.B.D. Carvalho, S.M.M.D. Faria and V.M.M.D. Silva, 2009. On ECG beat signal compression with 1-D multiscale recurrent patterns allied to preprocessing techniques. IEEE T. Bio-Med. Eng., 56(3): 896-900.

Ge, D., N. Srinivasan and S.M. Krishnan, 2002. Cardiac arrhythmia classification using autoregressive modelling. Biomed. Eng. Online, 1: 5.

Hu, Y.H., S. Palreddy and W.J. Tompkins, 1997. A Patient adaptable ECG beat classifier using a mixture of experts approach. IEEE T. Bio-Med. Eng., 44: 891-900.

Hu, Y.H., W.J. Tompkins, J.L. Urrusti and V.X. Afonso, 1993. Applications of artificial neural networks for ECG beat signal detection and classification. J. Electrocardiol., 26: 66-73.

Iftikhar, F., S. Ayesha and D. Arfa, 2012. Rhythm disorders-heart beat classification of an electrocardiogram beat signal. Int. J. Comput. Appl., 39: 38.

Ivanov, P.C., Q.D.Y. Ma, R.P. Bartsch, J.M. Hausdorff, L.A.N. Amaral, V. Schulte-Frohlinde, H.E. Stanley and M. Yoneyama, 2009. Levels of complexity in scale-invariant neural beat signals. Phys. Rev. E, 79: 041920.

Jekova, I., G. Bortolan and I. Christov, 2008. Assessment and comparison of different methods for heartbeat classification. Med. Eng. Phys., 30(2): 248-257.

Jiang, W. and S.G. Kong, 2007. Block-based neural networks for personalized ECG beat signal classification. IEEE T. Neural Networ., 18(6).

Ka, A.K., 2011. ECG beat classification using waveform similarity and RR intervals. J. Med. Biol. Eng., 32(6): 417-422.

Khadra, L.A.F.A. and S. Binajjaj, 2005. A quantitative analysis approach for cardiac arrhythmia classification using higher order spectral techniques. IEEE T. Bio-Med. Eng., 52(11): 1840-1845.

Kim, H., R.F. Yazicioglu, P. Merken, C.H. Hoof and H. Yoo, 2010. ECG beat signal compression and classification algorithm with quad level vector for ECG holter system. IEEE T. Inform. Technol. Biomed., 14(1).

Kohler, B.U., C. Hennig and R. Orglmeister, 2002. The principles of software QRS detection. IEEE Eng. Med. Biol., 21: 42-57.

Maglaveras, N., T. Stamkapoulos, K. Diamantaras, C. Pappas and M. Strintzis, 1998. ECG pattern recognition and classification using non-linear transformations and neural networks: A review. Int. J. Med. Inform., 52: 191-208.

Osowski, S. and T.H. Linh, 2001. ECG beat recognition using fuzzy hybrid neural network. IEEE T. Bio-Med. Eng., 48: 1265-1271.

Osowski, S., T. Markiewicz and L. TranHoai, 2008. Recognition and classification system of arrhythmia using ensemble of neural networks. Measurement, 41: 610-617.

Owis, M.I., A.B.M. Youssef and Y.M. Kadah, 2002. Characterization of ECG beat signals based on blind source separation. Med. Biol. Eng. Comput., 40: 557-564.

Ozbay, Y., R. Ceylan and B. Karlik, 2006. A fuzzy clustering neural network architecture for classification of ECG arrhythmias. Comput. Biol. Med., 36: 376-388.

Prasad, G.K. and J.S. Sahambi, 2003. Classification of ECG arrhythmias using multi-resolution analysis and neural networks. Proceeding of Conference on Convergent Technologies for Asia-Pacific Region (TENCON, 2003), 1: 227-231.

Senhadji, L., G. Carrault, J.J. Bellanger and G. Passariello, 1995. Comparing wavelet transforms for recognizing cardiac patterns. IEEE Eng. Med. Biol., 14(2): 167-173.

Ubeyli, E.D., 2009. Adaptive neuro-fuzzy inference system for classification of ECG beat signals using Lyapunov exponents. Comput. Methods Prog. Biomed., 9(3): 13-321.

Yeh, Y., 2012. An analysis of ECG for determining heartbeat case by using the principal component analysis and fuzzy logic. Int. J. Fuzzy Syst., 14(2): 233-241.

Yu, S. and K. Chou, 2007. A switchable scheme for ECG beat classification based on independent component analysis. Expert Syst. Appl., 33: 824-829.

End Note:
[1]http://physionet.org/cgi-bin/atm/ATM.