IJournals: International Journal of Software & Hardware Research in Engineering
ISSN-2347-4890
Volume 4 Issue 5 May, 2016

ECG Beat Classification Using Adaptive Thresholding and Artificial Neural Networks

Mr.Arbaj Hussain Mizwani
M.Tech Scholar
Astral Institute of Technology and Research,
Indore, M.P., India
arbajmizwani@gmail.com

Ms.Megha Nigam
Assistant Professor
Astral Institute of Technology and Research,
Indore, M.P., India
meghanigam19@gmail.com

ABSTRACT

Heart diseases are among the top three reasons for human deaths. Among several heart abnormalities, cardiac arrhythmia is a major heart abnormality which occurs due to abnormal rate or rhythm of the heart. Classification of cardiac arrhythmia is a difficult task. One of the ways to diagnose cardiac arrhythmia is to use electrocardiogram (ECG) signals.[1] The ECG is the most important bio-signal used by cardiologists for diagnostic purposes. Therefore analysis and classification of the ECG signal has remained one of the most sought after fields in applications of computer engineering in the field medicine and computational biology. In this paper, we have proposed a novel method comprising of filtering and adaptive thresholding for the feature extraction of ECG signals. Subsequently we have employed artificial neural networks and the Euclidean classifier for the classification of the different ECG signals. Finally it has been shown that the proposed methodology achieves higher values of sensitivity and accuracy compared to previously existing methods. It is expected that the proposed work will not only augment the knowledge of the academia in the field of computational medicine but will also prove pivotal in the exactness of diagnosis of patients with different heart abnormalities.

Keywords: - Electrocardiogram (ECG), feature extraction, adaptive thresholding, artificial neural network (ANN), Euclidean classifier, baseline drift.

1. INTRODUCTION

According to a recent survey, it is being revealed that, by 2030, almost 23.6 million people will die from Cardiovascular Diseases (CVD), mainly from heart disease and stroke. They are found to remain the main cause of death. One of the CVD risk factors is Cardiac Arrhythmia.

Cardiac arrhythmia is a major type of abnormal heart activity. An arrhythmia is a problem with the rate or rhythm of the heartbeat. During an arrhythmia, the heart can beat too fast, too slow, or with an irregular rhythm. A heartbeat that is too fast is called tachycardia. A heartbeat that is too slow is called bycardia .Most arrhythmias are harmless, but some can be serious or even life threatening. During an arrhythmia, the heart may not be able to pump enough blood to the body. Lack of blood flow can damage the brain, heart, and other organs.

Classification of cardiac arrhythmia is a difficult task. One of the ways to diagnose cardiac arrhythmia is to use electrocardiogram (ECG) signals. The ECG is the most important bio-signal used by cardiologists for diagnostic purposes. The ECG signal provides key information about the electrical activity of the heart. The early detection of the cardiac arrhythmias can prolong life and enhance the quality of living through appreciates treatment.

2. THE CARDIAC CYCLE

The regular motion of the human heart is often referred to as the cardiac cycle. The presence of sodium and potassium ions in the blood stream produces very weak electrical signals (voltages) when blood flows in and out of the heart. It has been observed that the ECG signals follow a repetitive or periodic pattern. Based on the trajectory of the ECG curve, certain fundamental features have been identified. The section that follows explains the cardiac cycle. ECG is the graphical representation of the cyclic rhythm of contraction and relaxation activity generated by the heart. An ECG is composed of the P wave, QRS complex, T and U waves. They are denoted by the capital letters P, Q,R,S, and T and U. The P wave is the contraction of the atria, while the QRS complex is associated with the contraction of the ventricles. The T wave is due to the relaxation of the ventricles. The P, Q, R, S, T and U waves of the ECG signal contain all the important features that characterize the activity in the heart. A typical ECG signal waveform of a normal heart beat is shown in the figure.

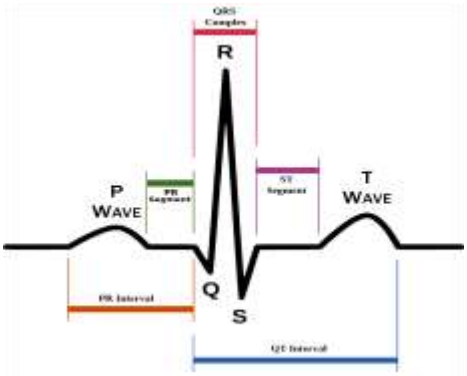


Fig 1.ECG signal showing P, Q, R, S, T and U waves.[4]

The ECG signal is measured through a number of electrodes that are normally attached to a patient's body. ECG recordings usually contain high and low frequency noise. Amplitudes within beats vary from person to person. That is why the effective manual detection ECG arrhythmia is very important, but it is tedious and time consume. Due to the ECG signal, monitoring may have to be carried out over several hours because the volume of the ECG data is enormous. This difficulty turns out a very high possibility of the analyst missing (or misreading) vital information. Therefore, computer-based analysis and detection of diseases can be very helpful in cardiologist's diagnoses specially in the intensive care units the electro-cardiogram (ECG) care units recordings, especially in real-time long-term (24h) monitoring.. This paper proposes an algorithm to detect and classify the ECG arrhythmia.

A successful ECG arrhythmia classification usually involves three important procedures: signal preprocessing, feature extraction, and classifier construction. Feature extraction is the important procedure that usually influences the classification performance of any ECG arrhythmia classification system.

Therefore, the extraction of relevant features to achieve optimal classification results has become primary tasks for the ECG arrhythmia classification problems.

2. ECG PREPROCESSING[2],[3]

Prior to the feature extraction stage, proper pre processing stage in very crucial for the correct extraction of features. In some ECG signals the noise level is very high and it is not possible to recognize it by single recording, it is important to gain a good understanding of the noise processes involved before one attempt to filter or preprocess a signal. The ECG signal is very sensitive in nature, and even if small noise mixed with original signal the characteristics of the signal changes. The most difficult problem faced by an automatic ECG analysis is the large variation in the morphologies of ECG waveforms, it happens not only for different patients or patient groups but also within the same patient. Since the ECG signal is the most affected by 50-60 Hz power line noise also called baseline drift, therefore we need to employ high pass filtering for its removal.

3. FEATURE EXTRACTION

This stage consists of extraction of salient features which can give conclusive results for different heartbeat cases.. The heartbeat detection module attempts to locate all heartbeats .The feature extraction module forms a feature vector from each heartbeat. The feature extraction modules are required, because greater classification performance is often achieved if a smaller number of discriminating features are first extracted from the ECG.[7].[9] The Feature Extraction Parameters:

- RR interval evaluation.
- SS interval evaluation.
- QQ interval evaluation.
- QRS complex evaluation.

ECG Feature Extraction plays a significant role in diagnosing most of the cardiac diseases. One cardiac cycle in an ECG signal consists of the P-QRS-T waves. This feature extraction scheme determines the amplitudes and intervals in the ECG signal for subsequent analysis. The amplitudes and intervals value of P-QRS-T segment determines the functioning of heart of every human. The following figure illustrates various ECG features.

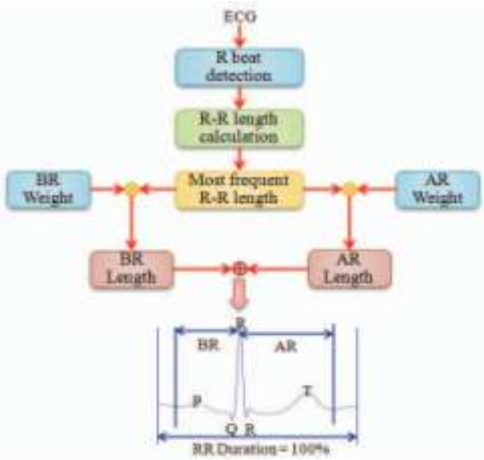


Fig2. Various ECG features.

Based on the values of the features extracted from the ECG signals, the different ECG signals are classified.

4. ARTIFICIAL NEURAL NETWORK

Work on artificial neural network has been motivated right from its inception by the recognition that the human brain computes in an entirely different way from the conventional digital computer.[5],[12] The brain is a highly complex, nonlinear and parallel information processing system. It has the capability to organize its structural constituents, known as neurons, so as to perform certain computations many times faster than the fastest digital computer in existence today. The brain routinely accomplishes perceptual recognition tasks, e.g. recognizing a familiar face embedded in an unfamiliar scene, in approximately 100-200 ms, whereas tasks of much lesser complexity may take days on a conventional computer. A neural network is a machine that is designed to model the way in which the brain performs a particular task. The network is implemented by using electronic components or is simulated in software on a digital computer. A neural network is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experimental knowledge and making it available for use. It resembles the brain in two respects:

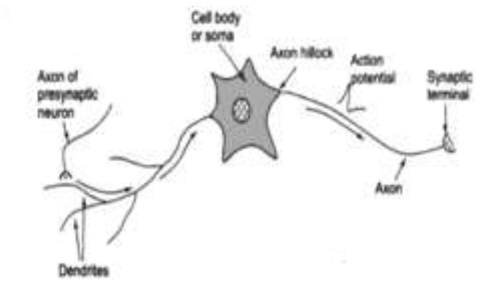
1. Knowledge is acquired by the network from its environment through a learning process.

2. Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge.

Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. Other advantages include:

1. Adaptive learning**:** An ability to learn how to do tasks based on the data given for training or initial experience.
2. Self-Organization**:** An ANN can create its own organization or representation of the information it receives during learning time.
3. Real Time Operation**:** ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.

The biological model of the neuron is shown in the figure. It consists of the cell body, axon hillock, action potential, synaptic terminal, axon of pre synaptic neuron and dendrites. Signals from different parts of the body travel through different parts and reach the neuron where the neuron processes it and produces an output. It should be noted though that the output of a neuron may also be fed to another neuron. A collection of such neurons is called a neural network. The neural network can perform simple to complex tasks depending on the structure of the neural network.



ISSN-2347-4890

Volume 4 Issue 5 May, 2016

Fig.3 Biological model of neuron

After studying the basic biological model of the neural network, a mathematical model is envisaged to be designated. The mathematical model for such a neural network is given by:

$$\sum_{i=1}^{n} XiWi + \Theta$$

Where Xi represents the signals arriving through various paths, Wi represents the weight corresponding to the various paths and Ө is the bias. The above concept can be visualized by the following diagram:

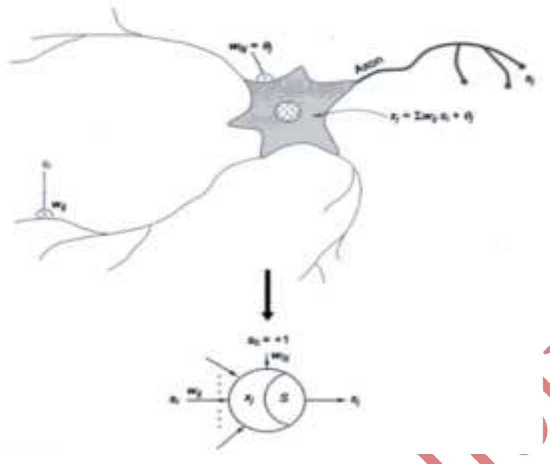


Fig.4 Mathematical model of a neural network

The above diagram exhibits the derived mathematical model of the neural network. It can be seen that various signals traversing different paths have been assigned names X and each path has been assigned a weight W. The signal traversing a particular path gets multiplied by a corresponding weight W and finally the overall summation of the signals multiplied by the corresponding path weights reaches the neuron which reacts to it according to the bias Ө. Neural networks can be used in the classification of ECG data in different categories of diseases or that of normal condition.

4. CLASSIFICATION OF ECG SIGNALS BASED ON FEATURE ECTRACTION [13],[15]

Based on the obtained feature values, the different ECG samples can be classified using neural networks and the Euclidean distance classifier[15]. The Euclidean distance classifier is based on the mathematical model that if the distance of a particular feature value id the minimum from a particular data set range, then that data sample is considered to be in that category. Mathematically:

If min (d_0, d_1, d_2… d_n) =D ;
Then, X ε C

Where D is the minimum Euclidean distance from a particular data set group X from a category C.
After the classification stage, the parameters used for the accuracy and effectiveness of the proposed algorithm are given below:

1. **True Positive (TP):** It indicates the samples which are affirmative in the test and are also judged so by the algorithm.
2. **True Negative (TN):** It indicates the samples which are non-affirmative in the test and are also judged so by the algorithm.
3. **False Positive (FP):** It indicates the samples which are non-affirmative in the test but are judged affirmative by the algorithm.
4. **False Negative (FN):** It indicates the samples which are affirmative in the test but are judged non-affirmative by the algorithm.

Sensitivity (S_e): It indicates the algorithms ability to correctly detect the samples which test affirmative and belong to a particular category specified by feature values. Mathematically:

S_e= TP/(TP+FN)

Accuracy (Ac): It indicates the algorithms ability to detect correctly whether a sample belongs to a particular data set category or not out of all the possible classification outcomes. Mathematically:

A_c=(TP+TN)/(TP+TN+FP+FN)

Thus the above mentioned parameters help in the evaluation of any proposed algorithm which deals with statistical data and statistical modeling.

6. PROPOSED METHODOLOGY

Next, we focus on the step by step implementation of the proposed methodology. It should be noted here that the ECG data has been obtained in the form of 1 minute samples from MIT BIH library. The format of the files is .mat files. The following diagram explains the proposed methodology.

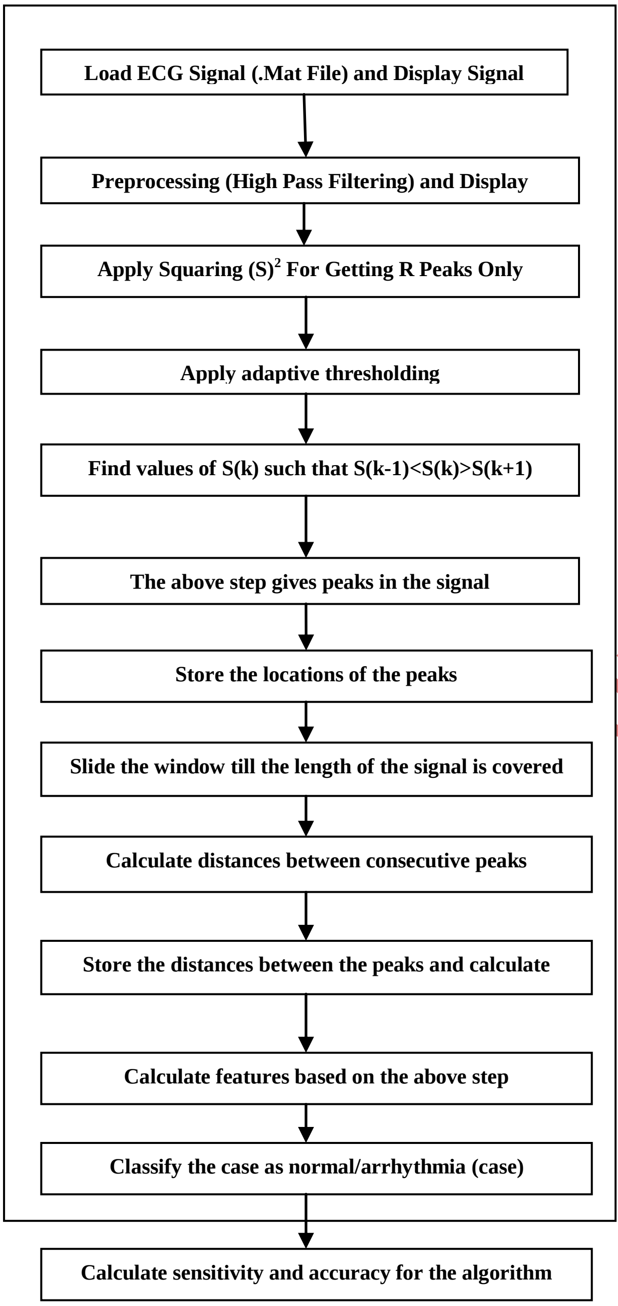


Fig.5 Proposed Methodology

The ECG signal in the form of .mat file is loaded into the MATLAB workspace. Then the ECG signal is displayed. The signal is then passed through a high pass filter the output of which is displayed again. The baseline drift is seen to be removed from the ECG signal due to filtering. Let y(t) denote the output of the filter, x(t) denote the raw ECG signal and h(t) denote the impulse response of the filter. Then:

$$y(t)=x(t)*h(t)$$

where * denotes convolution in the time domain.

It should be noted that the sampling frequency of the filter should follow the Nyquist criteria i.e.

$$Fs >= 2fm$$

Where Fs denotes the sampling frequency and fm denotes maximum frequency of the signal. Subsequently squaring the signal is done to accurately detect R peaks as R peaks are much larger in amplitude compared to other peaks.

$$Sqr\_sig=[y(t)]^2$$

Where Sqr_sig denotes square of the filtered signal. It should be noted that squaring is done only for detection of R peaks as other peaks can not be discriminated after squaring and may introduce errors.

Peaks are detected after setting a threshold which varies adaptively with the concerned peak and signal under consideration.

Peaks are detected using the difference operation that a sample is a peak if it is greater in magnitude compared to previous and subsequent values i.e.

$$S(k-1)<S(k)>S(k+1)$$

Then the locations of the peaks are stored and through subsequent differences, the features are extracted. Finally the ECG samples are classified according to the Euclidean distance:

If $\min(d_0, d_1, d_2\ldots d_n) = D$;
Then, $X \varepsilon C$

and then the Sensitivity and Accuracy are calculated.

7. SIMULATION RESULTS AND DISCUSSION

In this study, the simulations have been carried out using MATLAB platform for different sizes of text

ISSN-2347-4890
Volume 4 Issue 5 May, 2016

The analysis for a .mat file s0017lrem.mat has been shown below. The same process has been employed for all the other .mat files of the MIT-BIH library.

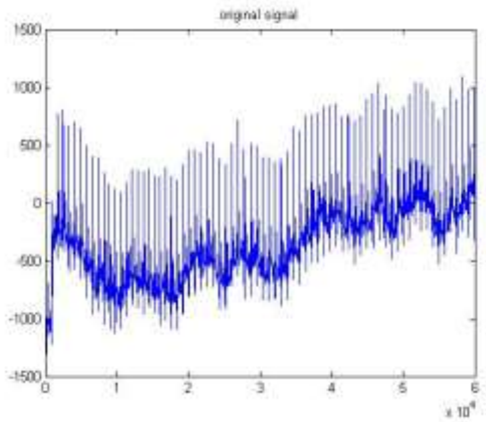


Fig.6 Original ECG signal

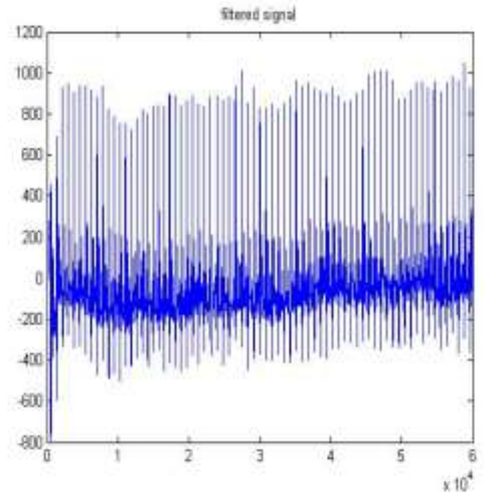


Fig.7 Filtered signal

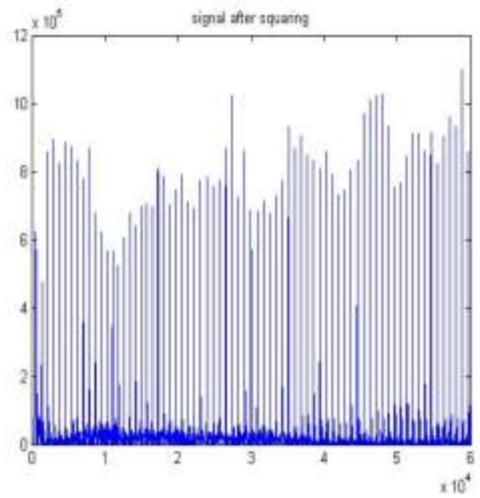


Fig.8 Squared Signal after filtering

S.No	MIT-BIH Tape No.	Mean R	Mean 'Q'	Mean 'S'	Mean 'QRS complex'
1	109m(LBBB)	5.699	2.5763	0.1217	2.6520
	207m	0.0434	1.0600	0.0924	1.0853
2	118m(RBBB)	4.6951	2.3202	0.0985	2.5073
	212m	3.5816	1.8550	0.1015	1.9426
3	203m(PVC)	1.6599	1.1209	0.1062	1.2170
	205m	0.0544	-	0.0756	8.4937
4	104m(APC)	11.8594	7.3544	0.0799	7.3393
	107m	1.6961	2.7214	0.0689	3.4931
5	106m(NORM)	3.6309	0.0278	0.0821	6.7222
	108m	0.0486	14.7418	0.0768	16.2704

Table.1 ECG data after classification

The above table tabulates the values of the different features extracted from different MIT BIH .mat files. It can be seen that the values vary over a particular period.

The table below shows the sensitivity and accuracy of the algorithm for the different .mat files loaded.

It can be seen that the obtained values are better than previously existing work.

Thus the efficacy of the proposed algorithm stands validated against the obtained results tabulated here.

S.No	MIT-BIH Tape No.	Sensitivity(Se)	Accuracy (Ac)
1	109m(LBBB)	99.115	99.1229
	207m	99.6350	99.6364
2	118m(RBBB)	99.1526	99.1579
	212m	99.3549	99.3590
3	203m(PVC)	99.5885	99.5902
	200m	97.1452	97.2244
		-	-
4	104m(PB)	97.4376	97.5016
	107m	98.7807	98.7954
5	106m(NORM)	98.3	96.4706
	108m	97.7514	95.0125

Table.2 Sensitivities and Accuracies after classification

MIT BIH TAPE	Negatives	Positives
S10017lrem.mat (1 minute .mat file sampled at 1000Hz)	3506(TN)	7(FP)
	12(FN)	1007(TP)

Table.3 Confusion Matrix for s10017lrem.mat

Conclusion

It can be concluded from the above mentioned results that the proposed algorithm achieves a higher values of Sensitivity and Accuracy compared to previously existing techniques. The obtained results can be attributed to the accuracy in peak detection due to squaring of the ECG samples after filtering. As R peak detection is the most crucial parameter for ECG beat classification, therefore high accuracy mandatorily needs correctness in R peak detection. Also adaptive thresholding has been employed for non ambiguous detection and classification of peaks except R peaks. This leads to an overall higher value of accuracy and sensitivity of the proposed algorithm.

References

[1] Abdelhaq Ouelli, Belachir Elhadadi, Belaid bouikhalene," *Multivariate Autoregressive Modelingfor Cardiac Arrhythmia Classification Using Multilayer Perceptron Neural Networks*" 2014 , IEEE,978-1-4799-3824-7/14.

[2] Somsanuk Pathoumvanh, Kazuhiko Hamamoto, Phoumy Indahak,"*Arrhythmias Detection And Classification Base On Single Beat Ecg Analysis*", 4th Joint International Conference on Information and Communication Technology, Electronic and Electrical Engineering (JICTEE-2014)

[3] Rashad Ahmed, Samer Arafat,"*Cardiac Arrhythmia Classification Using Hierarchical Classification Model*", 2014 6th International Conference on CSIT.

[4] Pathrawut Klaynin, Waranyu Wongseree, Adisorn Leelasantitham ,Supaporn Kiattisin,"*An Electrocardiogram Classification Method Based On Neural Network",*2013 Biomedical Engineering International Conference (BMEiCON-2013).

[5] Nurul Hikmah Kamaruddin, M.Murugappan, Mohammad Iqbal Omar," *Early Prediction Of Cardiovascular Diseases Using Ecg Signal:*

Review", 2012 IEEE Student Conference on Research and Development.

[6] Manab Kumar Das, Student Member, IEEE, Dipak Kumar Ghosh, Samit Ari, Member, IEEE, *" Electrocardiogram (Ecg) Signal Classification Using S-Transform, Genetic Algorithm And Neural Network",* 2013 IEEE 1st International Conference on Condition Assessment Techniques in Electrical Systems.

[7] Bushra Mehdi, Tahmina Khan, Zain Anwar Ali," *Artificial Neural Network Based Electrocardiography Analyzer*", 2013 IEEE

[8] Shameer Faziludeen1, Sabiq P.V," *Ecg Beat Classification Using Wavelets And Svm*", Proceedings of 2013 IEEE Conference on Information and Communication Technologies (ICT 2013).

[9] Berdakh Abibullaev, Won-Seok Kang, Seung Hyun Lee3 and Jinung An," *Classification Of Cardiac Arrhythmias Using Bi-Orthogonal Wavelet Preprocessing And Svm*".

[10] Shital L. Pingale, Nivedita Daimiwal," *Detection Of Various Diseases Using Ecg Signal In Malab*", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-3, Issue-1, March 2014

[11] F. N. Mohamad, M. S. A. Megat Ali, A. H. Jahidin, M. F. Saaid, and M. Z. H. Noor," *Principal Component Analysis And Arrhythmia Recognition Using Elman Neural Network*", 2013 IEEE 4th Control and System Graduate Research Colloquium, 19 - 20 Aug. 2013.

[12] M. S. A. Megat Ali, and A. H. Jahidin, A. N. Norali," *Hybrid Multilayered Perceptron Network For Classification Of Bundle Branch Blocks*", 2012 International Conference on Biomedical Engineering (ICoBE),27-28 February 2012

[13] Shivaji Rao M. Jadhav, Sanjay L. Nalbalwar, Ashok A. Ghatol, " Generalized Feedforward Neural Network Based Cardiac Arrhythmia Classification From Ecg Signal Data".

[14] S.Karpagachelvi, Dr.M.Arthanari, M.Sivakumar, "*Ecg Feature Extraction Techniques - A Survey Approach*", IJCSIS) International Journal of Computer Science and Information Security,Vol. 8, No. 1, April 2010.

[15] Nahit Emanet," *Ecg Beat Classification By Using Discrete Wavelet Transform And Random Forest Algorithm*", 2009 IEEE.