ISSN: 2348 - 2117

International Journal of Engineering Technology and Computer Research (IJETCR)
Available Online at www.ijetcr.org
Volume 4; Issue 3; May-June-2016; Page No. 05-11

DWT Feature Extraction and Classification Using ANN for Detection of Cardiomyopathy

Rabiya Begum[1], Manza Ramesh R[2]

[1]Department of Computer Science and Information technology, Dr.B.A.M.U University, Aurangabad, 431001(M.S), India

rabiya.altaf@gmail.com

[2] Department of Computer Science and Information technology, Dr.B.A.M.U University, Aurangabad, 431001(M.S), India

manzaramesh@gmail.com

Abstract

The Proposed System Detects a Heart Disease called Cardiomyopathy from ECG signal by using Image and Signal processing techniques Discrete Wavelet Transform and Artificial Neural Network. Cardiomyopathy is a disease of Heart muscle in which the muscle get enlarged or thickened and the heart lacks it's working. Electrocardiogram or ECG is a diagnostic tool used for finding any heart related problem the ECG is Visually Analyzed by Cardiologist as a Human he is prone to make wrong interpretation that may leads to wrong diagnosis hence automation is required to help medical practitioners therefore the system is developed to Analyze the ECG signal, extract the valuable information and detects the disease automatically. ECG signals for Cardiomyopathy and Healthy are obtained from PTB diagnostic ECG database. DWT is used for feature Extraction and Artificial Neural Network is used for classification. Classification accuracy obtained with Levenberg-Marquardt algorithm is 93.65% which is better when compared with gradient descent giving 88.71%.

Keywords: *Cardiomyopathy, Discrete Wavelet Transform, Artificial Neural Network, Classification accuracy*

1. INTRODUCTION:

Cardiomyopathy is a progressive disease of the myocardium, or heart muscle. In most cases, the heart muscle weakens and is unable to pump blood to the rest of the body as well as it should. There are many different types of cardiomyopathy caused by a range of factors, from coronary heart disease to certain drugs. These can all lead to an irregular heartbeat, heart failure, a heart valve problem, or other complications. [1] the symptoms of cardiomyopathy includes Breathlessness with exertion or even at rest , Swelling of the legs, ankles and feet , Bloating of the abdomen due to fluid buildup ,Cough while lying down , Fatigue , Irregular heartbeats that feel rapid, pounding or fluttering , Chest pain Dizziness, lightheadedness and fainting. [2] Dilated Cardiomyopathy is the most common form of Cardiomyopathy, in this the pumping ability of the heart's main pumping chamber — the left ventricle — becomes less forceful. The left ventricle becomes enlarged (dilated) and can't effectively pump blood out of the heart. Although this type can affect people of all ages, it occurs most often in middle-aged people and is more likely to affect men. [2].

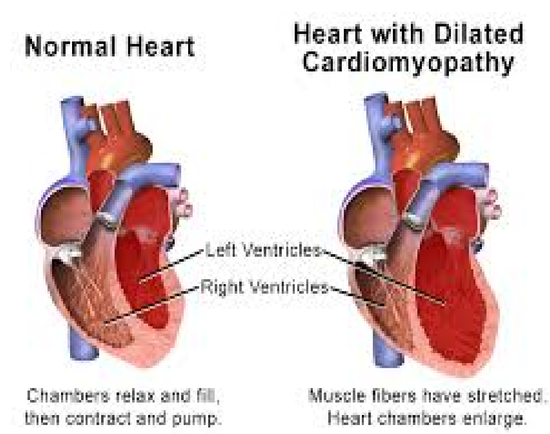


Figure 1: Heart with Dilated Cardiomyopathy

The electrical activity associated with any muscle action travels through various tissues and ultimately reaches the surface of the body where it can be detected by electrodes applied to the skin. The record obtained from the depolarisation and repolarisation voltages of the heart muscle is called an electrocardiogram or ECG.[3]. A typical ECG tracing is a repeating cycle of three electrical entities: a P wave (atrial depolarization), a QRS complex (ventricular depolarization) and a T wave (ventricular repolarization). The ECG is traditionally interpreted methodically in order to not miss any important findings. [4] The diagram below shows the ECG Normal Sinus Rhythm.

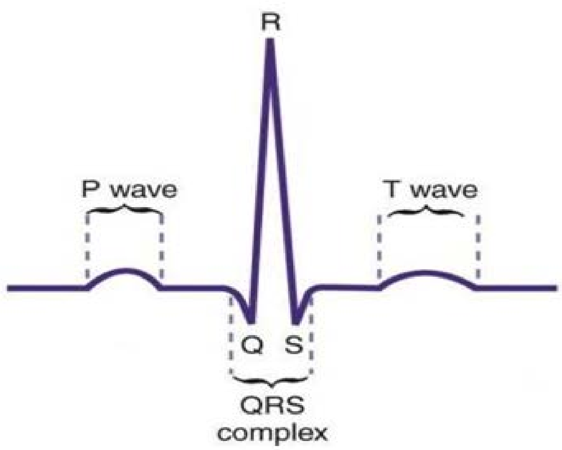


Figure 2: ECG Normal Sinus Rhythm

With the growing needs of medical diagnosis many researchers developed algorithms for automated ECG analysis. The various techniques for ECG Feature Extraction and classification, Methods includes Adaptive Thresholding, Linear Discriminant Analysis (LDA), Principal Component Analysis, Wavelet Transforms, FFT, Artificial Neural Networks, Support Vector Machines and Neuro Fuzzy Approach. Etc.

2. METHODS AND PROCESS

The proposed system for Cardiomyopathy detection with ECG analysis has the steps as follows.

A. Signal Acquisition
B. Preprocessing
C. Feature Extraction
D. Classification
E. Performance Evaluation

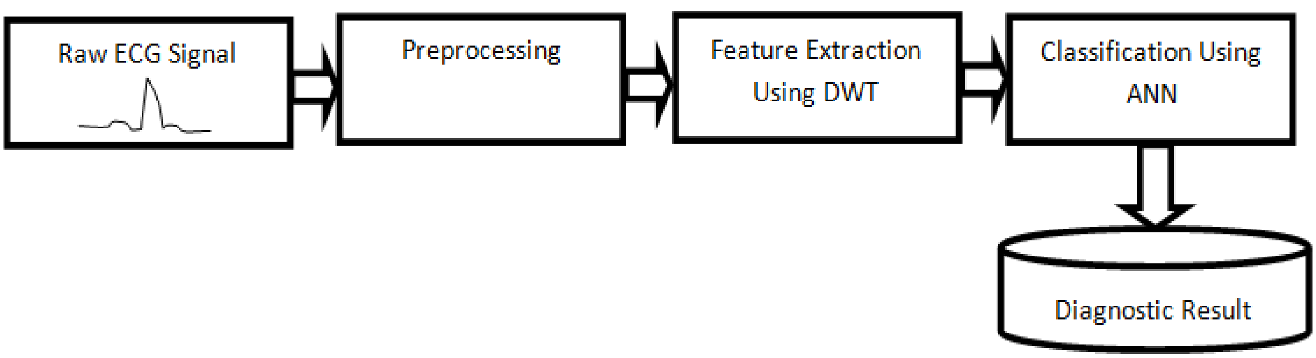


Figure 3: Block diagram of the System

A. Signal Acquisition:

ECG signals are acquired from PTB diagnostic ECG databasee provided by the National Metrology Institute of Germany, for research, algorithmic benchmarking or teaching purposes to the users of Physionet. The database contains 549 records from 290 subjects (aged 17 to 87, mean 57.2; 209 men, mean age 55.5, and 81 women, mean age 61.6; ages were not recorded for 1 female and 14 male subjects). Each subject is represented by one to five records. There are no subjects numbered 124, 132, 134, or 161. Each record includes 15 simultaneously measured signals: the conventional 12 leads (i, ii, iii, avr, avl, avf, v1, v2, v3, v4, v5, v6) together with the 3 Frank lead ECGs (vx, vy, vz). Each signal is digitized at 1000 samples per second, with 16 bit resolution over a range of ± 16.384 mV. [5].for this work we are working on 63 records of the database out of which 49 are Healthy and the remaining 14 are of cardiomyopathy.

B. Preprocessing:

ECG signal Artifacts are the noise induced to ECG signals that result from movements of electrodes. This in turn causes deformation and change in the electrical characteristics of the skin under and around the electrodes. These electrical changes appear in the ECG as motion artifacts and baseline drifts. [6] These artifacts have to be removed for correct diagnosis. Baseline removal is an important requirement; the frequency content of the baseline wander is usually in a range well below 0.5Hz. This baseline drift can be eliminated without changing or disturbing the characteristics of the waveform. We use the median filters (200-ms and 600-ms) [7] to eliminate baseline drift of ECG signal. The process is as follows-[8]

a) The original ECG signal is processed with a median filter of 200-ms width to remove QRS complexes and P waves.
b) The resulting signal is then processed with a median filter of 600-ms width to remove T waves. The signal resulting from the second filter operation contains the baseline of the ECG signal.
c) By subtracting the filtered signal from the original signal, a signal with baseline drift elimination can be obtained.
The raw/Noisy signal, baseline removed signal and preprocessed signal are shown in figure 5, 6 and 7 of Result section.

C. Feature Extraction:

The feature extraction process selects and retains relevant information from original signal. The Feature Extraction stage extracts diagnostic information from the ECG signal. In the proposed algorithm the ECG signal after processing is decomposed up to level 8 using Db 6 wavelet thus separating approximate and detail coefficients. After Wavelet decomposition up to level 8 it is observed that small scales represents the high frequency components and large scales represents the low frequency components of the signals. The first and eighth level reconstruction coefficients represent high frequency and low frequency contents of the ECG waveform respectively which in most of the cases appear to be high and low frequency noises. It is clear that most energy of the QRS complex is concentrated at decomposition level 3, 4 and 5.thus d3 , d4 and d5 coefficients are selected for detection of QRS complex.[9]

$e1 = d3 + d4 + d5$ ..... *(1)*

Although the QRS region is properly captured but it is difficult to identify R peak due to its oscillatory nature. So, a function $e2$ is defined as,

$e2 = \frac{d4X(d3+d5)}{2^n}$ ..... (2)

Where *n* is the level of decomposition.

Then the modulus of $e1 \times e2$ is taken. The accuracy of the entire feature extraction work mainly depends upon the identification accuracy of R peak. [9][10].The coefficients for R peak are shown in figure of Result section.

After the R peak coefficients selection the coefficients for remaining waves Q and S, QRS complex and P and T are calculated by using following equations.

For Q and S waves
$e3 = d2 + d3 + d4 + d5$ .....(3)

For P and T waves
$e4 = d6 + d7$ ... (4)

Thus from wavelet decomposition we have extracted four features coefficients for R peak, Q and S waves, QRS complex and p and T waves from lead I and lead II. The Coefficients from level 1 to 8 are shown in figure 8 and the extracted coefficients are presented in figure 9 of Results section.

D. Classification

Classification is the process of partitioning data into groups based on some characteristics feature. Through an extensive review, it was discovered that the ANN maintains its efficiency as a fundamental yet powerful neural network-based arrhythmia classifier. This in particular, was made possible through implementation of innovative learning algorithms. Hence, this paper proposes the ANN network for analysis and detection of cardiomyopathy disease. The work focuses on classification performance of feed forward backpropagation neural network structure with two learning algorithms. Learning algorithms plays an integrative function in any neural network-based classifiers. Apart from modifying the network performance, different algorithms will directly affect the convergence rate during the learning process, hence influencing the training speed. To date, the gradient descent algorithm remains the most fundamental approach in the back-propagation weight update procedure. The algorithm can be expressed mathematically by (5) [11]

$\mathbf{w_{k+1} = w_k - a_k g_k}$ ....(5)

Where w_k is a vector of present weights and biases, a_k is the learning rate, while g_k represents the current gradient.

In an effort to get better performance the network is also trained with The Levenberg-Marquardt algorithm which can be mathematically expressed by (6)

$$W_{k+1}=w_k-[J^TJ-\mu I]^{-1}J^Te \quad \dots (6)$$

Where J is the Jacobian matrix, e is the vector of network errors and µ is a constant

The work of classification is expressed in the following figure

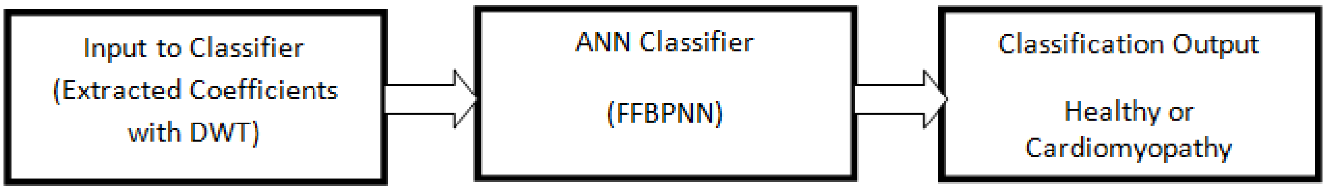


Figure 4: Classification Process

E. Performance Evaluation:

The performance of classifier is evaluated using three statistical Measures those are accuracy, sensitivity, positive predictive. These parameters are defined using 4 measures True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) [12].

TP (True Positive): Cardiomyopathy detection matches with decision of physician.

TN (True Negative): both classifier and physician advised absence of Cardiomyopathy.

FP (False Positive): system labels a healthy case as a Cardiomyopathy.

FN (False Negative): system labels a Cardiomyopathy as healthy

Accuracy: Accuracy is the ratio of number of correctly classified Samples, and is given by

Accuracy= (TP+TN) / N ×100 .... (7)

Total number of cases are N

Sensitivity: Sensitivity refers to the rate of correctly classified positive. Sensitivity may be referred as a True Positive Rate. Sensitivity should be high for a classifier.

Se = TP / (TP+FN) ×100 ... (8)

Positive predictive: Positive predictive is probability that disease is present when test is positive, which is by how much amount disease is correctly predicted.

Pp = TP/ (TP+FP) ×100 .. (9)

3. RESULTS AND DISCUSSION

The Feed forward backpropagation neural network structure was created having eight input layers , 10 hidden layers and one output layer , this network was trained with gradient descent (GDX) and Levenberg-Marquardt (LM) algorithms**.** From a total of 63 subject’s data, 70% were used for network training, while 30% were used for validation and testing purposes. The classification results obtained are shown in table I. the classification accuracy with Levenberg-Marquardt (LM) is 93.65% which is better when compared with gradient descent and (GDX) giving 85.71%.

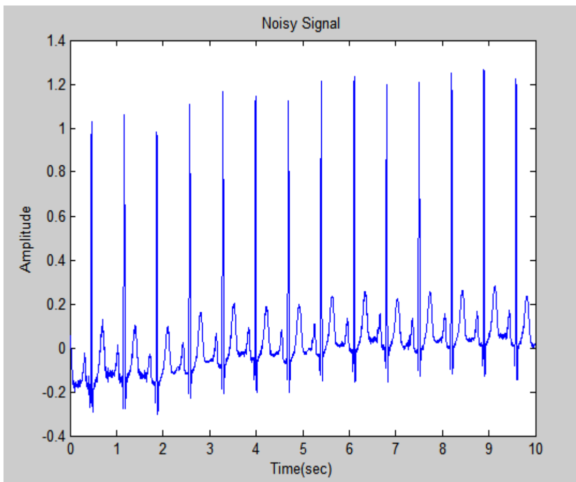


Figure 5: Noisy signal

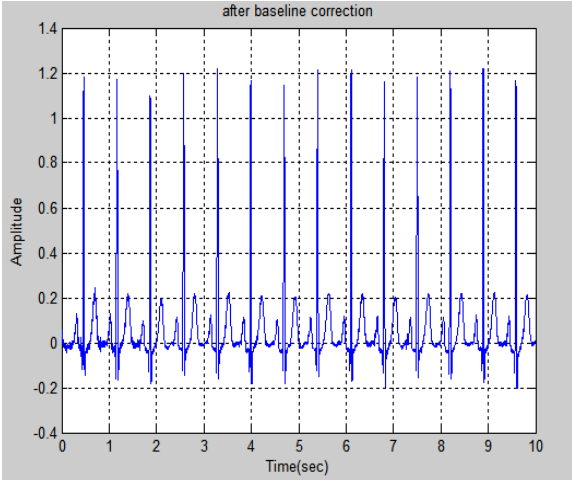


Figure 6: Baseline removes signal

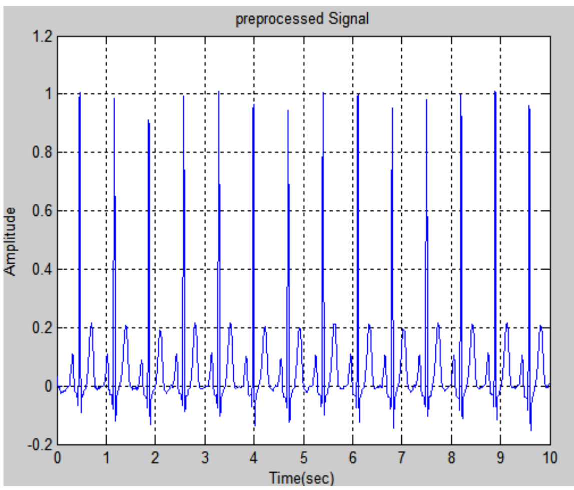


Figure 7: Preprocessed Signal

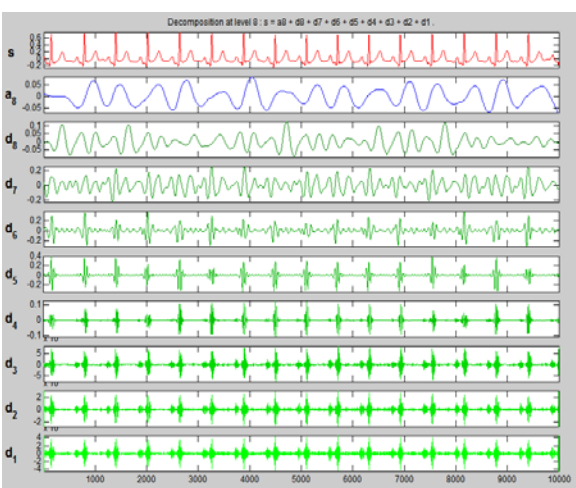
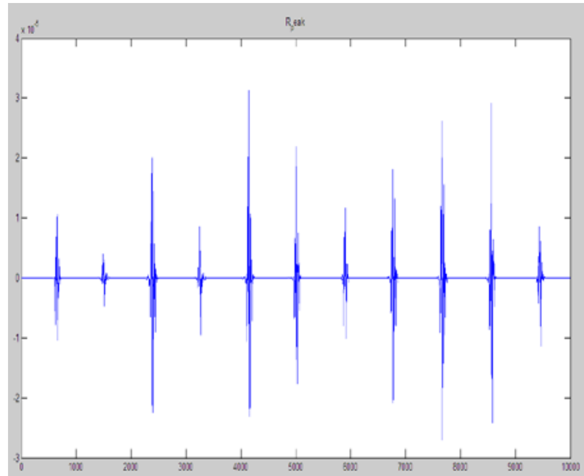
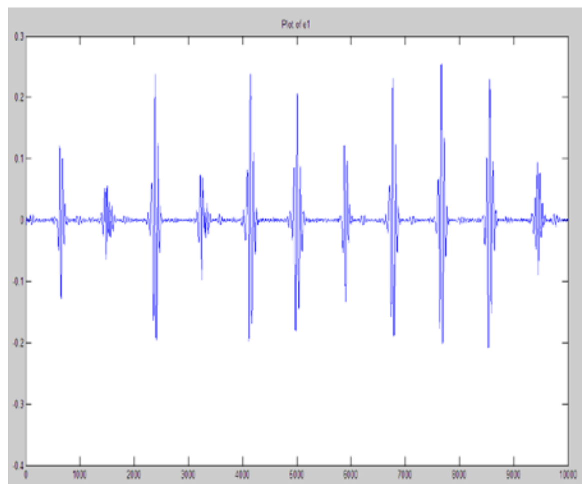


Figure 8: Wavelet Coefficients for level 1 to 8

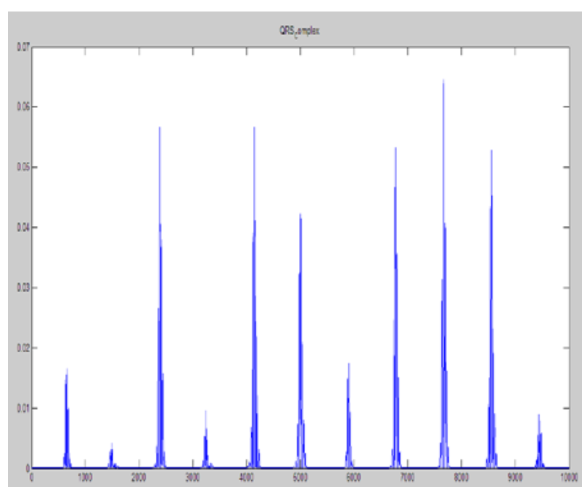
Selected Coefficients for R peak (a), for Q and S wave (b), for QRS complex (c) , for p and T waves (d)



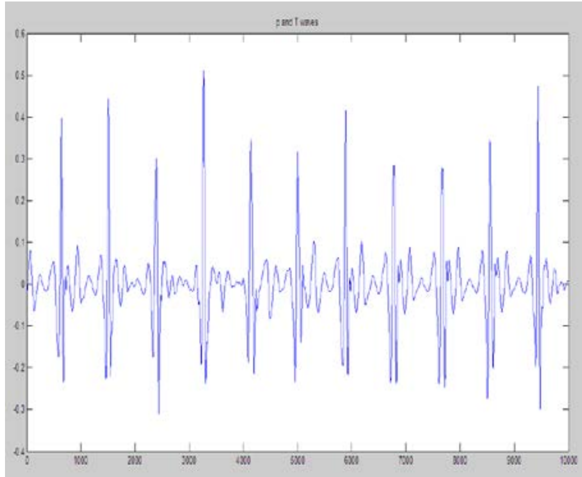
a. R peak Coefficient



b. Q and S Wave Coefficient



c.QRS complx Coefficient



d . P and T wave Coefficient.

Table 1: Classification Results

Learning Algorithm	Average Classification Accuracy	Healthy Acc	Cardiomyopathy Acc	Se	Pp
gradient descent (GDX)	88.71%	95.9%	64.3%	64.28%	81.81%
Levenberg-Marquardt (LM)	93.65%	98.0%	78.6%	78.57%	85.71%

4. CONCLUSION

The System for Detection of Cardiomyopathy is developed that uses DWT for feature Extraction and ANN for classification. The ECG data obtained from PTB diagnostic ECG database. Total 63 subjects data is processed out of which 49 are healthy and remaining 14 are cardiomyopathy. 8 features are extracted from lead I and lead II those are coefficients for R peaks, Q and S waves, QRS Complex and P and T waves. The classification results shown that Feed forward backpropagation neural network when trained with Levenberg-Marquardt (LM) algorithm gives better classification accuracy of 93.65 % Compared with gradient descent (GDX) algorithm with 85.71% classification accuracy.

5. REFERENCES:

1. http://www.healthline.com/health/heart-disease/cardiomyopathy#Overview1
2. http://www.mayoclinic.org/diseases conditions/cardiomyopathy/basics/definition/CON-20026819?p=1
3. http://www.cardionetics.com/introduction-to-the-ecg.
4. https://en.wikipedia.org/wiki/Electrocardiography.
5. https://www.physionet.org/physiobank/database/ptbdb/.
6. N. Siddiah, T.Srikanth and Y. Satish Kumar “Nonlinear filtering in ECG Signal Enhancement “International, ISSN22495789, Volume**:** 02;Issue**:** 01;2012.
7. P. de Chazal, C. Heneghan, E. Sheridan, R.Reilly, P. Nolan, M. O'Malley, “Automated Processing of the Single-Lead Electrocardiogram for the Detection of Obstructive Sleep Apnoea”, IEEE Trans. Biomed.Eng., 50(6): 686-689, 2003.
8. P.D.Khandait , N.G. Bawane, S.S.Limaye “Feature Extraction of ECG signal For Cardiac Arrhythmias” National Conference on Innovative Paradigms in Engineering & Technology(NCIPET -

2012) Proceedings Published by International Journal of Computer Applications(IJCA).

9. Saurabh Pal, Madhuchhanda Mitra ”Detection of ECG characteristics points using Multiresolution Wavelet Analysis based selective Coefficient Method” Elsevier, Measurement 2010;43,255–261

10. Soroor Behbahani, Nader Jafarnia Dabanloo “ Detection of QRS complexes in the ECG signal Using Multiresolution Wavelet and Thresholding Method” Computing in Cardiology 2011 ;38:85-808 ISSN 0276-6574.

11. M.S.A.MegatAli, C.Z.A.CheZainal, A.Husman, M.F.Saaid, M.Z.H.Noor, and A.H.Jahidin ”Detection of Cardiomyopathy Using Multilayered Perceptron Network” 8thInternationalColloquim on Signal Processing and its Applications.2012 IEEE.

12. P. Keerthi Priya, Dr.G.Umamaheswara Reddy” MATLAB Based GUI for Arrhythmia Detection Using Wavelet Transform” International Journal of Advanced Research in Electrical,Electronics and Instrumentation Engineering(An ISO 3297: 2007 Certified Organization) Vol. 4, Issue 2, February 2015.