ARRYTHMIA RECOGNITION AND CLASSIFICATION USING ECG MORPHOLOGY AND SEGMETATION

S.Nithyaselvakumari[1], M.C. Jobin Christ[2], B A Gowri Shankar[3]
[1]Assistant professor, Dept of Biomedical Engineering, Sri Jayaram Institute of Engineering and technology, Nithyaselvakumari.s@gmail.com
[2]Professor Department of Biomedical Engineering, Rajalakshmi Engineering college, jobinchrist.mc@rajalakshmi.edu.in
[3] Assistant Professor, School of Chemical & Biotechnology SASTRA Deemed-to-be University, arungowrishankar@gmail.com

ABSTRACT:
Cardiac arrhythmia can be identified using abnormal electrical activity of heart, this is a great menace to humans. In order to diagnose cardiac problems ECG signal is widely used. When the background noise is rejected from the ECG signal we obtain a QRS component. This QRS component consists of high frequency and high energy waves that are very easy to detect and study. Once QRS component is obtained, it is further spited into various classes that can aid in diagnosing the abnormalities. Previously extracted features are compared to find the heart abnormalities. In this paper Feed-Forward neural network is selected and data base are used to store and analyze the data.

KEYWORDS: ***Neural Network, ECG signal, Feature extraction, Classification, Test and Training***

1. INTRODUCTION

In order to automate the study of classification of arrhythmias and diagnosis of heart diseases electrocardiography-based signals (ECG) is widely used over in the recent years. This can be executed by monitoring a person's heart beat rate over a reasonable amount of time during normal routine and when the person is engaged in cardiac activity. Various arrhythmia classification made using the ECG signals. This could be tedious and difficult; hence this is carried out using the machine learning oriented algorithmic techniques. Irrespective of the classification used to diagnose some steps that need to be performed are very critical to design the exact approach to find out arrhythmia

Depending on the initial pre-processing phase performed, the value of classification required to deal with ECG signal is determined. To complete this task an important step is carried out, that is deduction of QRS, particularly the R waveform. Most part of the techniques to detect and segment heart beat is based on the position of such deflection. The QRS is easier to detect and segment, because of the angular coefficient amplitude of the R wave. Final step is the classification of ECG signal which is accomplished in a supervised fashion. Machine language is used for this type of analysis and conclusion; the widely used techniques are Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs). This has a drawback of pattern recognition technique with respect to the parameters, hence fine tuning prior to the analysis is required on the sample [1].

Harmonics are present and harmonics are analyzed to find that there has been a time varying harmonics present. The assumption is such that the components are varied less. Hence a combination of spectral and probabilistic method is required; this is called as the evolutionary spectrum. The techniques applied on Fourier Transform method assume the linearity and stationary signals. While distorted waveforms are constantly changing, this results in the traditional probabilistic approach. In order to carry out proper diagnosis it is necessary to get a harmonic free waveform. The issue of getting a harmonic free waveform is analyzed with a number of time vs. frequency techniques. Here we will discuss on using the wavelet transform as a tool to analysis harmonics. But this mainly deals with finding the equivalent coefficient and does not satisfy the engineer's physical understanding of harmonic distortion. In general, for a steady state signal harmonics can be considered as a minor issue. Whereas when the transform are unstable it is difficult to simplify, for a non stationary signals it is difficult to perform Fourier method analysis. The magnitude of fourier transform does not fluctuate between positive and negative but is somewhat accurate positive envelope in Fourier domain. Fourier transform is absolutely shifting without variation, with a simple linear phase offset coding the shift. The coefficient is not aliased and a complicated aliasing cancellation property to reconstruct the signal, the sinusoidal M-D Fourier basis are highly directional plane waves. The DWT is a real-values oscillating wavelet, whereas the Fourier transform is a complex-valued oscillating sinusoids [2].

The below figure shows the time-varying waveform with help of two signals Figure 1.

The first signal is a steady state distorted wave where it has harmonics varying with time. Over a period of time the harmonics component of the second signal is monitored. This requires continuous supervision in systems like the power system.

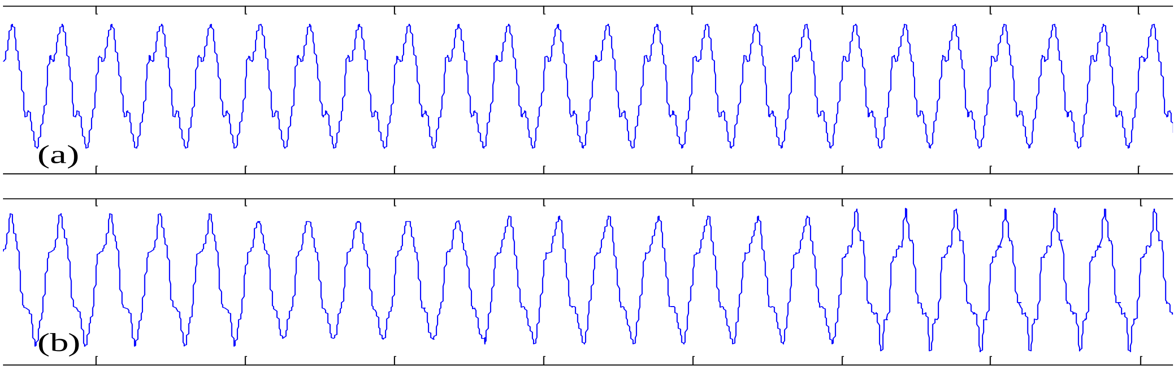

Fig1: Steady state distorted waveform; (b) time-varying waveform distortion

2. RELATED WORK

This deal with deducting heart beat perfectly even under the presence of severe background noise. Here an adaptively timed multi-layer perceptron formation is used to model the non-linear time varying background noise. The noise is removed from subtracting the predicted noise component from original signal waveform. The research makes it clear that the ANN based approach is most critical for the development of automated on-line ECG arrhythmia monitoring system.

Based on neural network and Prony's modeling algorithm another method of identifying the abnormalities in heart beats is possible. Signals can be written as exponentially changing form of finite sum of series based on the ECG signal[3]. The beats of a multi layer feed forward neural network based propagation are categorized into five types. They are ventricular bigeminy (VB), normal sinus rhythm (NSR), ventricular fibrillation (VF), ventricular tachycardia (VT), and ventricular couplet (VC).

PVC (Premature Ventricular Contraction) and PAC (Premature Atrial Contraction) are applied by the system to perform and classificated based on P-QRS-T peak value without. This extraction feature extraction done using deep NN. Every parameter used in PVC and PAC classification as part of NN learning is updated using P-QRS-T peak without feature value. Since this method has a validation of performance, it is highly accurate than the previous studies. In recent times various research algorithms are developed to extract the features. For the P, QRS and T waves a classification that is automated generates heart beats depending on morphological properties as shoen on figure 2

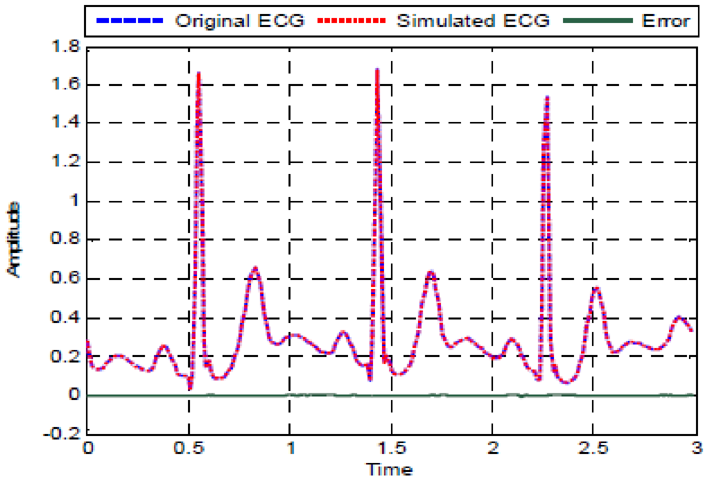

Fig. 2. A patient ECG and simulated ECG

A extraction of ECG signal based on morphological and time interval features is also difficult to analyze and describe the patterns that do not have any clear P and T waves, hence not appropriate for finding the types of heartbeat. For the purpose of classification ECG signals are taken based on the frequency domain, where the mean, variance and power spectrum of the frequency were taken as a feature of the ECG signal. This methods lags in finding the accurate results as compared to other methods due to poor performance in classifying heartbeats and robustness of noise[4]. After detecting QRS complexes the multi resolution wavelet transforms feature was set and the coefficients for the discrete wavelet transforms are detected.

At low frequencies a long width window and at high frequencies a short width window can be used overcome the above stated issues with wavelet transforms. The method based on wavelet transform is suitable as this provides more signal information in both time and frequency domains[6]. From the investigations it is clear that the maximum resolution is profived by wavelet transform in both high and low frequencies as shown in figure 3 and figure 4.

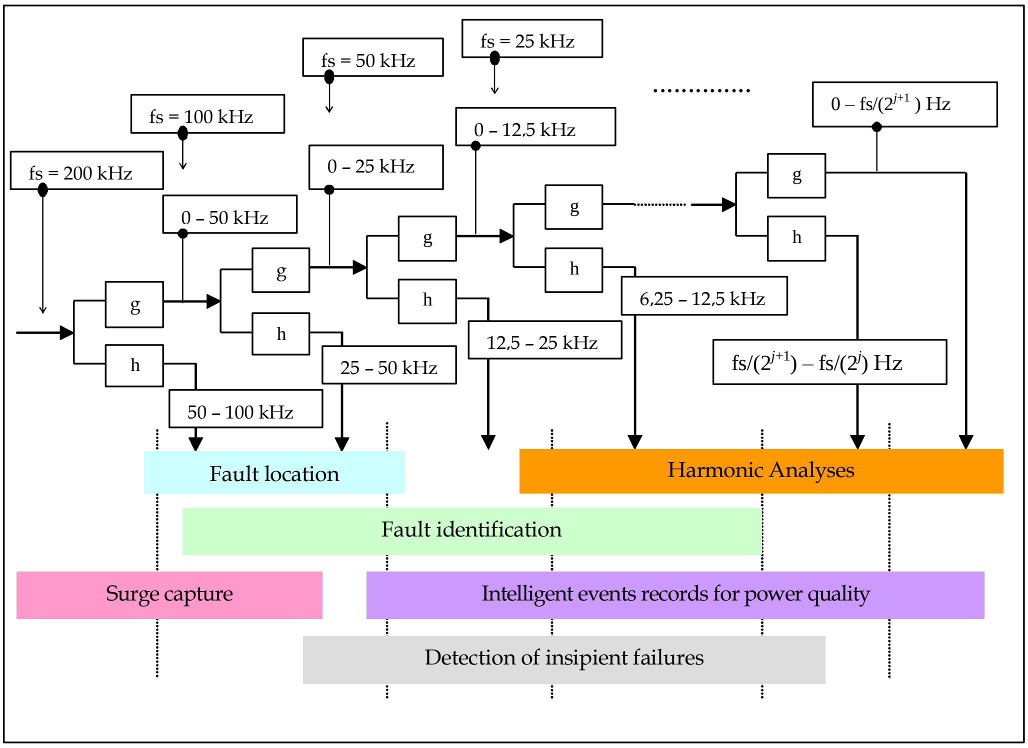

Fig 3: Multi-Resolution Analysis and Applications in Power Systems

For the above demonstration MRA is performed and then decomposed to visualize the signal composed by 0.3 pu of 7th harmonic, 1 pu of 60 Hz 0.12 pu of 13th harmonic and some noise[5]. Harmonic content is present at all times and the sampling happens at the rate of 10 kHz.

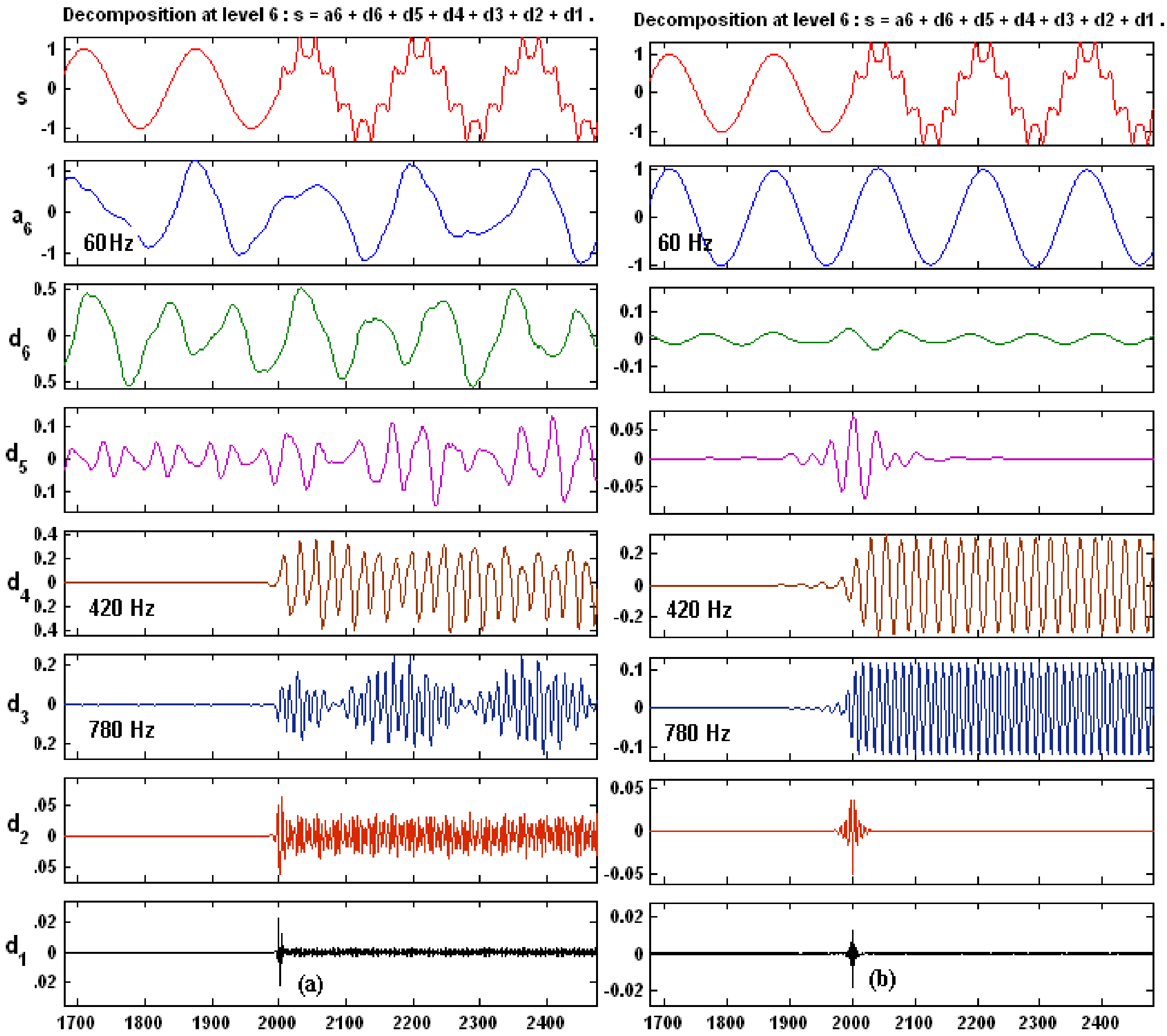

Fig 4 – MRA with six level of decomposition

3. METHOLODGY

Cardiovascular diagnosis is mainly based on the ECG signal. ECG signals cab get contaminated by different features of noise like Power line disturbance, Baseline wandering and Electromyogram signals. After filtering out the noises the signals are fed into next stage for further processing. Feature extraction and classification are the two divisions in identification of ECG signals.

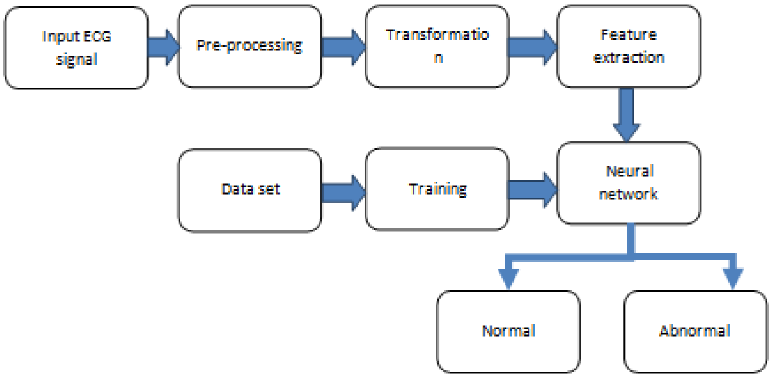

Fig 5: Block Diagram of ECG signal

Block diagram of proposed methodology has been depicted in figure 5.The fundamental requirement of a reliable and accurate ECG monitoring system, a fundamental requirement is to identify all the life threatening symptoms without raising any false information. The ECG signal processing system in off-line has claimed reorganization; on-line monitoring system is most preferred. Particularly the philosophy of design is to identify every possible drawback; the rate of getting a false alarm is way too high. Example, even with 99.3% accuracy rate which can be achieved with our database using QRS detection algorithms, per patient the average of 33 false alarms are recorded. Hence any improvement in the deduction rate will be a most expected reform both economically and clinically. The ECG signal waveforms can be modeled a spike-like QRS complex wavelets for deducting the QRS complex, , which are super imposed on a time varying non linear background noise.

The background noise can be music, normal instrument noise also any of frequency components with P and T waves of ECG signals consisting. An effective method need to be used to remove background noise before a matched filter based on the applied conventional detection algorithm. There are two methods to remove the background noise; first one is made on high frequency components of QRS complex. Second one is having components of the very low frequency. In order to filter out only the main frequency components a band pass filter is popularly used. The output of QRS complex[7] is identified by band pass filter. Removal of background noise is very difficult as per this approach since it varies with time.

Band frequencies are subject to varied with time. Hence we go for data adaptive method. Linear LMS adaptive model was designed to reduce background noise by filtering out noise component.

Here, ECG signal denoted by s(t), QRS complex signal is xt and q(t) and background noise is denoted by n(t) as shown in equation 1.

$$\hat{n}(t) = \sum_{j=1}^{P} h(t,j)\, s(t-j) \tag{1}$$

The **QRS** complex signal then can be estimated as the prediction error as shown in equation 2.

$$\hat{q}(t) = s(t) - \hat{n}(t) \tag{2}$$

In Eq. (l), the filter coefficient h(t,j) is adaptively updated using the LMS algorithm:

$$h(t,j) = h(t-1,j) + \alpha\, \hat{q}(t)\, s(t-j) \tag{3}$$

a constant, is the determinant of step size. When the steepest descent gradient is need to be determined. Main aim is to reduce the value of E(&t), where (t) is periodic spikes in time domain. By proper choice of filter length p, the signal part of time does not have any impact, with any option of h(t,j) and the noise component n(t), will be suppressed. As filter can only able to model the low frequency part of the signal, hence a finite sized window of size p is obtained as shown in equation 3.

In order to remove the high frequency part of the signal, we remove the noise component by considering average of square of $(t) (i.e. Energy) of signal. After the improvement of signal quality, the same is passed into a matched filter for deduction of any QRS complex as shown in figure 6 and figure 7.

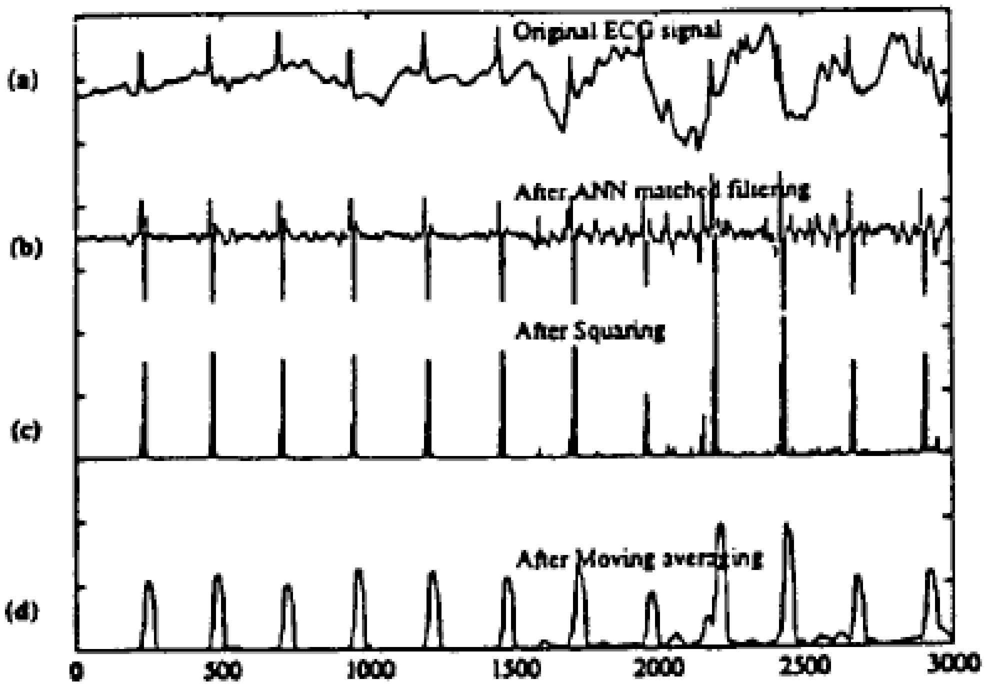

Fig 6: adaptive, nonlinear noise removal filtering result

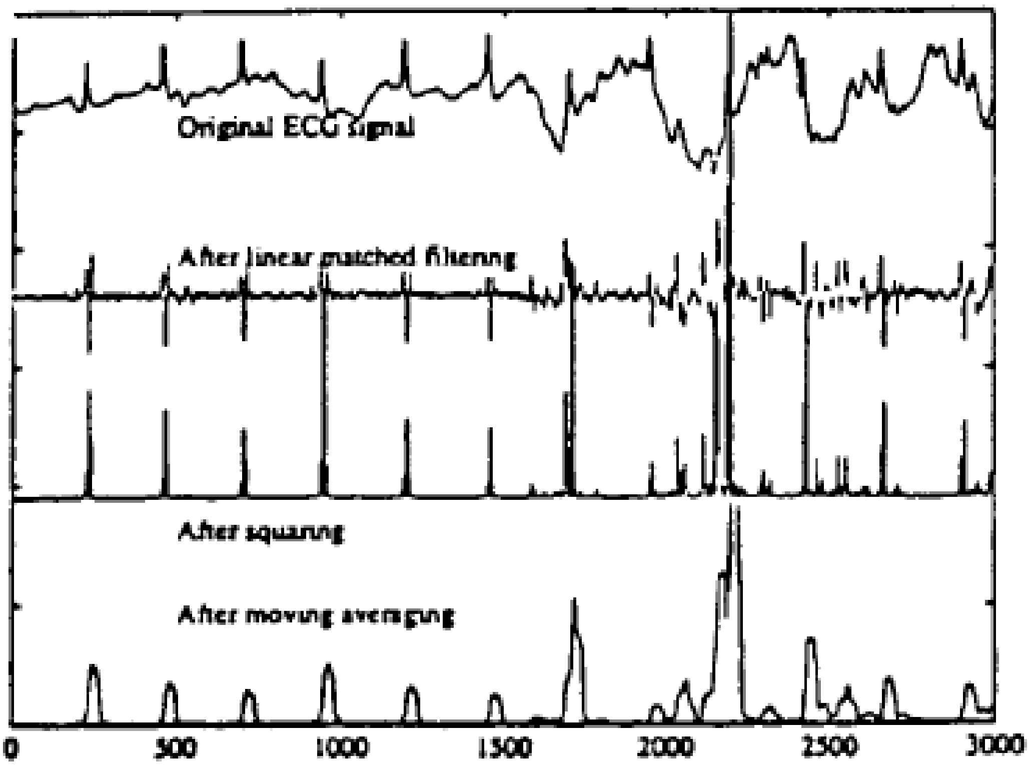

Fig7: Adaptive linear noise removal filtering result

3.1 DUAL-TREE COMPLEX WAVELET TRANSFORMS (DT-CWT)

The recent development to the discrete wavelet transform (DWT) is a dual-tree complex wavelet transform (CWT). This type of signal is invariant with respect to shift and selective towards a direction in two or higher dimension. For any dimensional signal the redundancy factor is lower than CWT. It is of 2d for DWT, where d is the dimensional signals. N in case of a multi dimensional signal namely (M-D), Dual-tree CWT can be a separable component based on a computational efficiency; filter bank is represented by FB. This analysis is based on designing complex wavelets with fine characteristics, illustration of wide spectrum of applications in signal and image processing.

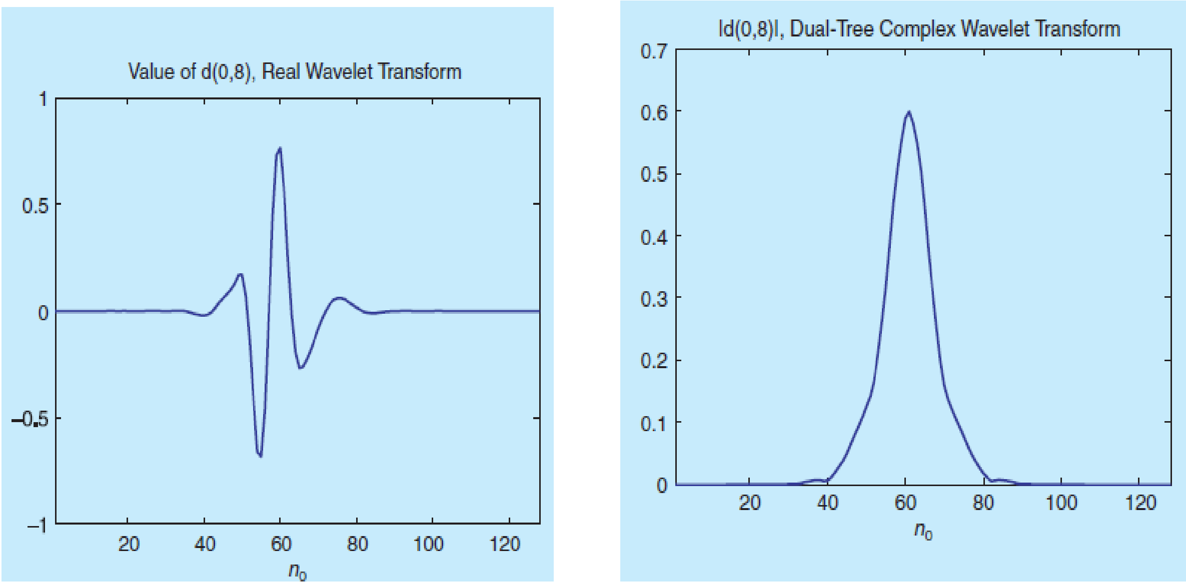

Fig8: The value of wavelet coefficient in "Real-Valued Discrete Wavelet Transform and Filter Banks

In the above figure 8 the real coefficient is *d* (0, 8*), which* is computed based on conventional real DWT. In the lower figure complex coefficient is *(*0, 8*) which* is computed by dual-tree CWT. Any test signal is basically a step edge at *n* , where n is the number, *x(n)* = *u(n − no)*. The value of the wavelet coefficient *d (*0, 8*) is represented in figure 6,* (the eighth coefficient at stage 3) in the "Real-Valued Discrete Wavelet Transform" with Filter Banks also as a function of *no*.

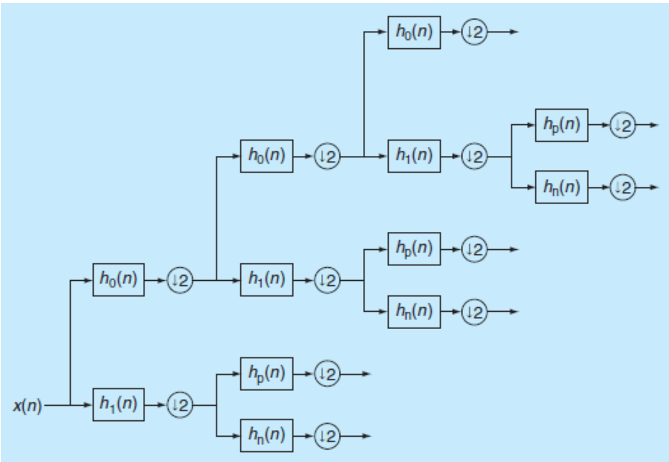

Fig 9: **Analysis FB for the DWT with invertible complex post-filtering**

It is calculated using GLCM are mean, correlation, dissimilarity, energy .The number of gray levels in the image determines the size of the CCM. Texture feature are extracted using gray level co-occurrence matrix (GLCM). Energy is treated as an homogeneousness image, it can be calculated from the normalized complex. Any disorder in texture image can be calculated as shown in figure 9. Entropy gives a measure of complexity of the image. High entropy is a part of complex textures. p(i , j) is the co occurrence matrix

3.2 FEATURE EXTRACTION AND CLASSIFICATION

RR interval is the defined as the interval between any two fiducial points of heartbeats, where R point is the fiducial points of heartbeats. We need to calculatethe RR interval first and then next normalized features of anterior RR interval. QRS wave is the 100 milliseconds after and before any R point, the fields are sampled by 10 points in the plane. Features of morphology are obtained using neighborhood field points of R points. After R points in-between 200 and 450 milliseconds a feature of certain disease is shown in T neighborhood.

Application of wavelet packet decomposition happens on signals up to 6 decomposition level, one of the calculated features is the wavelet packet entropy. A defined feature of statistical parameters is called as kurtosis, skewers and minimum and maximum value of heartbeat.

The main goal of choosing cross validation is for defining a dataset to "test" the model. In the training phase the model is used to check how a independent data set is generated from a model. Model is feed with known data datasets as input, in a prediction problem. On this model training is run generating a dataset of unknown data vs the model that is tested. Instead of using the conventional validation, a fair way to estimate is model prediction. For improving the performance of model prediction we cross validate the average measures, this corrects optimistic nature of training error.

4. IMPLEMENTATION

MATLAB is the software used for model prediction. It has features for application-specific solutions namely toolboxes. It is used as a high-performance language in technical computing. This is used for research and performing analysis with higer productivity and lesser errors. For designing and prediction of the output of the model, toll boxes are highly used. A GUIDE is used to simplify the creation of GUI applications; this is automated by generating a M-file frame work. M-File is coded with help of this frame work. Global data can be stored and handled in M-file. M-File can be used to efficiently store any object handling and to execute any function calls.

For ECG arrhythmia the popular method used is Artificial Neural Networks (ANNs). Neurons are used in this structure; they are simple, nonlinear processing elements. Every Neuron calculates the weighted sum of inputs passed through the synapse input and the subtracted sum from predefined bias and the pass its result through a function. The output value of function ranges between 0 and 1. One popular model of ANN is Multilayer Perceptron (MLP) as shown in figure 10, where each layer is connected to next layer in the model.

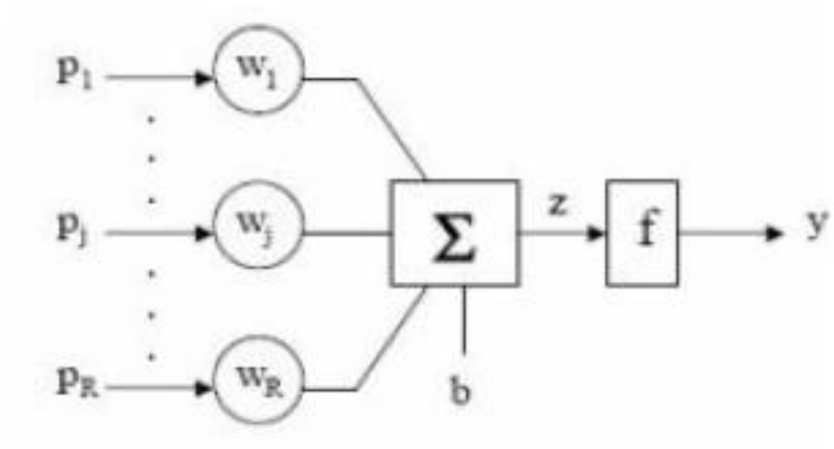

Fig 10: Functional description of single neuron

5. RESULTS AND CONCLUSION

Normally the sample data is equally spited into two components which do not overlap with each other, where one is used for training the model another is used for testing the model. While training the model the test data is hidden and not tried. The validation results in proper data without any overlaps, this will yield in better results and performance of the model based on the algorithm defined. For example if test data is too difficult or too simple then the results can be skew. If the datasets are not splinted again the results will be in skew manner. If the data are improper the results will not be achieved correctly, problems can arise by considering partially addressed results by repeating any hold-out validation multiple times with the averaging of the obtained results. Another drawback is the test data can be included in test data set multiple time, on the contradiction few data does not fall into test set in any of the testing.

Most commonly data is stratified and then it is spitted into k folds. Stratification is denotes as the process of rearranging the data in dataset so as to ensure that in each fold it is a good representative of the whole data set considered for the model. Later iteration for training are performed k times and each of the iteration is validated, finally the k-1 folds pending is used for learning. Let us take example of a classification problem using binary where each class needs to have 50% of test data, in order to obtain best results each fold has around half instances.

After each of the iteration the remaining data is used for making the model learn. One special case is Leave-one-out cross-validation (LOOCV), having number of instances in the data. LOOCV estimate is unbiased as compared to the others, it has a high variance which can be misleading to non reliable results. But still this model is widely used in data samples having very rare data, one application of this model is in the field of bioinformatics that has a few dozens of sample data set.

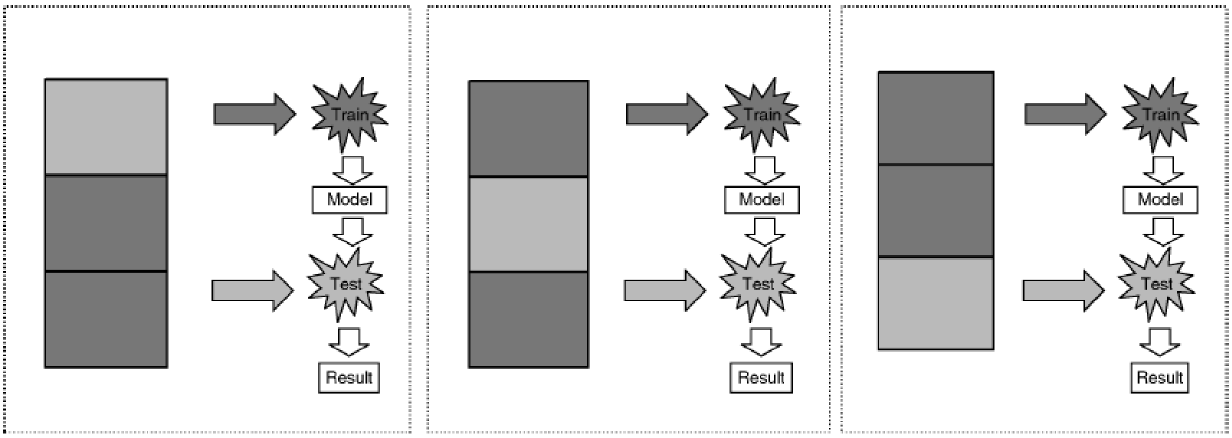

Fig11: Our Testing and Training Validation Result

The application of neural network experiments to automate the diagnosis of Arrhythmia using ECG was explained, after applying the neural network classification also post after performance validation of the combination of ECG morphology and the segment feature extraction on the ECG signals of a patient. The feasibility of the approach has been confirmed with the experiments performed with sample dataset. The approach has consistency, and has an advantage of forming desirable diagnosis accuracy.

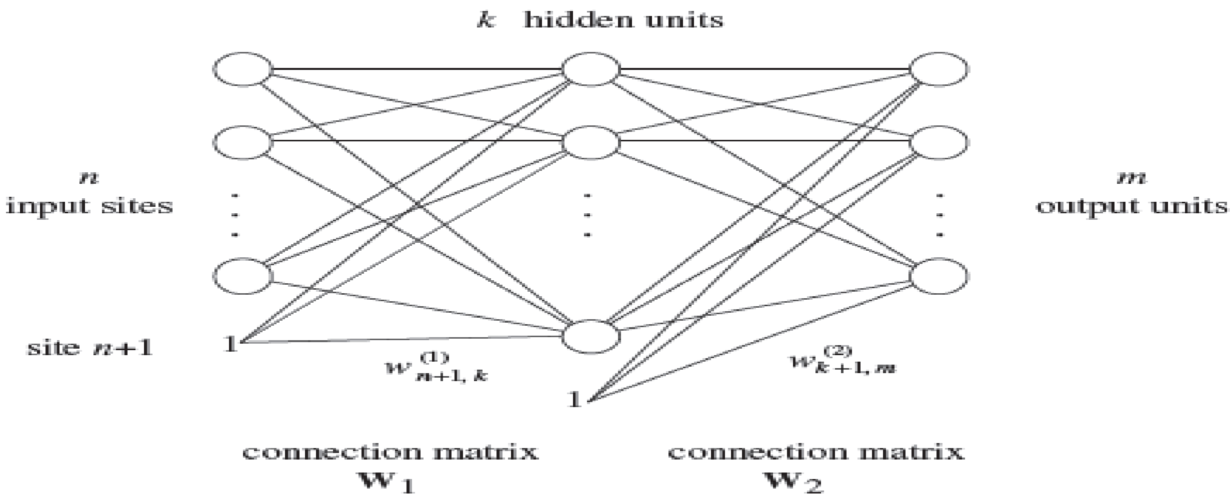

Fig 12: Neural Network Structure

Two phases of ANN are training and the testing phase. When the model is trained the various data set used for training are feed into the model and the output is given by the MLP. As the training algorithm has the desired outputs predefined, it is expected the correct the bias and weights of the model. Training phase is continued until the performance of the model is as desired, and the output matches the sample output dataset. During testing the model, samples data set that are not used during the training phase are used on trained MLP model. MLP classifiers can used to check if the unknown features are classified are not. One of the applications of neural network is automatic detection of the Arrhythmia disorder using ECG signals. High accuracy of desirable diagnosis is a popular advantage of neural network. The output data set confirms the feasibility of the neural network approaches.

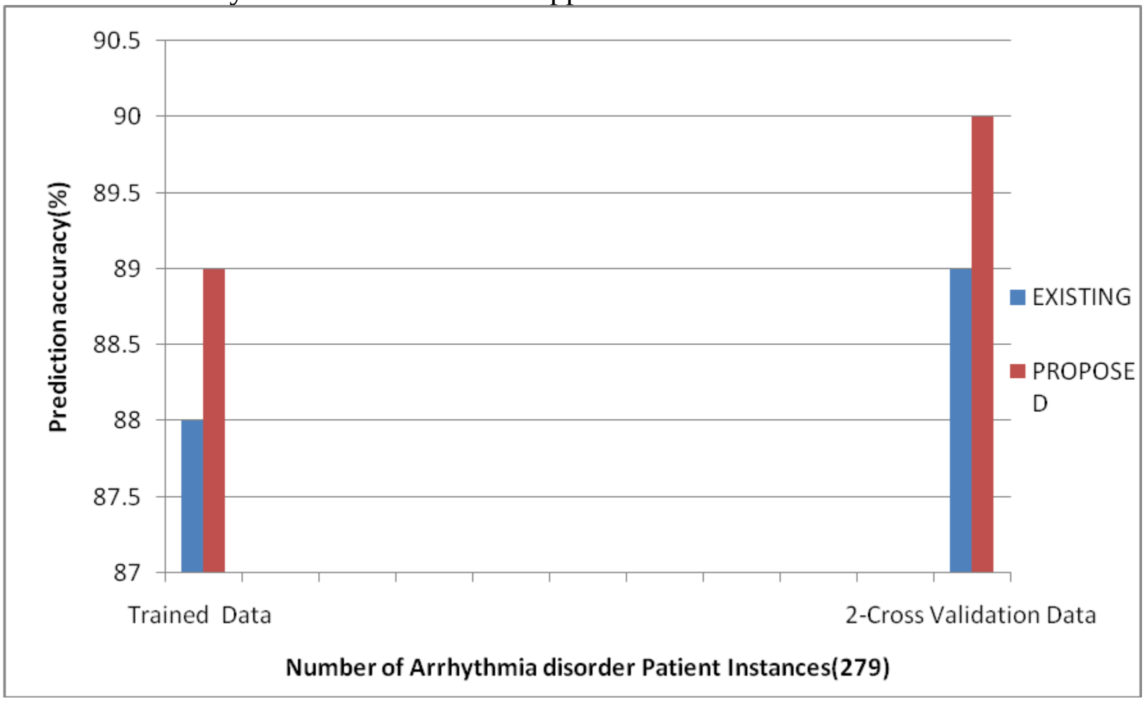

Fig 13: Results of Prediction Accuracy

The result shown in figure 13 depicts the prediction accuracy of Arrhythmia disorder where the datasets have been deployed in the CNN based model. The datasets consisted of 279 instances [8]with the parameters of Age, sex, height ,duration of QRS ,Weight and heartbeat rate . The conclusion provides a depiction where the proposed system acuuracy of prediction is high than the existing one.

6. REFERENCES

[1] K.PolatandS.Günes,“ClassificationofepileptiformECGusingahybrid system based on decision tree classifier and fast Fourier transform,”Appl. Math. Comput., vol. 187, no. 2, pp. 1017–1026, 2007

[2] A. Tzallas, M. Tsipouras, and D. Fotiadis, “Epileptic seizure detection in ECGs using time-frequency analysis,”IEEE Trans. Inf. Technol.Biomed., vol. 13, no. 5, pp. 703–710, Sep. 2009.

[3] N. Robinson, A. Vinod, K. K. Ang, K. P. Tee, and C. Guan,“ECG-based classification of fast and slow hand movements usingwavelet-CSP algorithm,”IEEE Trans. Biomed. Eng., vol. 60, no. 8,pp. 2123–2132, Aug. 2013.

[4] N. Mammone, F. La Foresta, and F. C. Morabito, “Automatic artifactrejection from multichannel scalp ECG by wavelet ICA,”IEEE SensorsJ., vol. 12, no. 3, pp. 533–542, Mar. 2012.

[5] R. B. Pachori and P. Sircar, “ECG signal analysis using FB expansionand second-order linear TVAR process,”Signal Process., vol. 88, no.2, pp. 415–420, 2008.

[6] R.J.OweisandE.W.Abdulhay,“SeizureclassificationinECGsignalsutilizing Hilbert-Huang transform,”Biomed. Eng. Online, vol. 10, no.1, p. 38, 2011

ISSN 2515-8260 Volume 7, Issue 4, 2020

[7] H. Altay Guvenir, Burak Acar, Gulsen Demiroz, Ayhan Cekin "A Supervised Machine Learning Algorithm for Arrhythmia Analysis." Proceedings of the Computers in Cardiology Conference, Lund, Sweden, 1997