American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-6, Issue-10, pp-28-35
www.ajer.org
Research Paper Open Access

# Cardiac Abnormalities Detection using Wavelet Transform and Support Vector Machine

[1] Ojo, J. A., [2]Amoran, A. E.
*Department of Electronic and Electrical Engineering,Ladoke Akintola University of Technology, P. M. B. 4000, Ogbomoso, Oyo State, Nigeria.*

***ABSTRACT****: An Electrocardiogram is a medical equipment that records the electrical activity of the human heart. Data obtained from the Electrocardiogram (ECG) have been used to assess the state and diseases associated with the heart. Detection of cardiac abnormalities is dependent on the correct identification of individual waves (Q, R, S, T) in a complete cardiac cycle. QRS detection has been reported in earlier works with less emphasis on T waves. In this paper, we developed a hybrid method comprising of Hilbert transform with autocorrelation and modified Analysis of Amplitude, Slopes and Width (MAASW) for QRS complex detection and Support Vector Machine (SVM) with radial based function were used for feature and classification respectively. Accuracy, Sensitivity and Specificity were used as Performance Evaluation methods. The developed algorithm was implemented in MATLAB and validated by ECG database from Massachusetts Institute Technology/Beth Israel Hospital (MIT-BIH) arrhythmia database. The proposed method has an accuracy of 82.33%, Sensitivity of 82.33% and Specificity of 96.47% while Five cardiac abnormalities namely Normal Sinus Rhythm, Left Bundle Branch Block Beat (LBBB), Right Bundle Branch Block (RBBB), Paced Beat (PACE), Premature Ventricular Contraction (PVC) and Atrial Premature Contraction (APC) were classified.*

Date of Submission: 21-09-2017 Date of acceptance: 06-10-2017

## I. INTRODUCTION

Cardiovascular diseases are diseases associated with the heart and it is a major cause of death globally as more people die annually from cardiovascular diseases than from any other causes and according to the World Health Organization report(Organisation, 2007) approximately 17.5 million people died from cardiovascular diseases in 2005, representing at least 30% of all global deaths, and according to GBD 2015 Mortality and Causes of Death, Collaborators 17.9million deaths in 2015 representing 32% of global death according (Murray, Lopex, Mohsen, & Haidong, 2015). However, early detection is critical to the treatment of these diseases and because of the large volume of patient's data, a computer-aided diagnostic system was developed to assist medical professionals. This computer-aided medical diagnostic system for monitoring the heart is called Electrocardiogram (ECG).

ECG is a diagnostic tool that measures and records the electrical activity of the heart in details and the state of cardiac health is generally reflected in the shape of the ECG waveform and heart rate (Sokolow, Mcllory M, & Cheithin, 1990). Figure 1 shows the ECG waveform over a single cardiac cycle comprising P wave, a QRS complex and a T wave respectively. An ECG is a graph of amplitude against time and it contains a critical pointer to the nature of diseases affecting the heart. The correct interpretation of these details allows the diagnosis of a wide range of heart problems. One cycle of the normal ECG is composed of a P wave, a QRS complex and a T wave, corresponding to the atrial depolarization, the ventricular depolarization and the rapid repolarization of the ventricles respectively

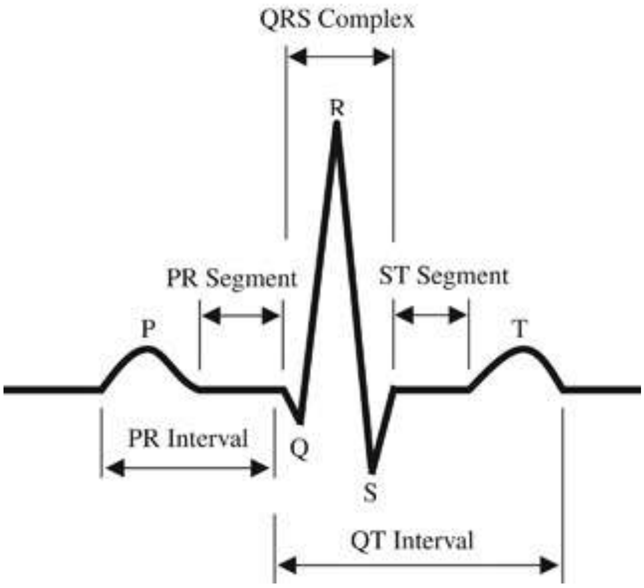


**Figure 1:** Normal ECG and its components(Ghorbanian, Jalali, Ghaffari, & Nataraj, 2012)

In literature, Different methods and algorithms from signal processing theory have been used to extract features from ECG signals. (Pan & Tompkins, 1985) proposed an algorithm to recognize QRS complex in which they analyzed the positions and magnitudes of sharp waves and used a special digital band pass filter (BPF). (Xu & Liu, 2004) described an algorithm using Slope Vector Waveform (SVW) for ECG QRS complex detection and RR interval evaluation. The work proposed a variable stage differentiation method which is used to achieve the desired slope vectors for feature extraction, and the non-linear amplification is used to get better of the signal-to-noise ratio. The method allows for a fast and accurate search of the R location, QRS complex duration, and RR interval and yields excellent ECG feature extraction results but T waves were not identified nor detected with this method.

(Mahmoodabadi, Ahmadian, & Abolhasani, 2005) described an approach for ECG feature extraction which utilizes Daubechies Wavelets transform. The work developed and evaluated an electrocardiogram (ECG) feature extraction system based on the multi-resolution wavelet transform. The ECG signals from Modified Lead II (MLII) were chosen for processing. The wavelet filter with scaling function further intimately like the shape of the ECG signal achieved better detection. The first step was to de-noise the ECG signal by removing the equivalent wavelet coefficients at higher scales. Then, QRS complexes are detected and each one complex is used to trace the peaks of the individual waves, including onsets and offsets of the P and T waves which are present in one cardiac cycle. The results revealed that the proposed approach for ECG feature extraction achieved the sensitivity of 99.18% and a positive predictivity of 98%.

(Melgani & Pasolli, 2010) presented active learning strategies for ECG signal classification. Starting from a small and suboptimal training set, these learning strategies select additional beat samples from a large set of unlabeled data. These samples are labelled manually and then added to the training set. The entire procedure is iterated until the construction of a final training set representative of the considered classification problem. All the proposed strategies are based on iterative procedures and use SVM to classify the signals. Three different strategies are described and compared: margin sampling (MS) in which the samples of the learning set closer to the hyperplanes between the different classes are chosen, posterior probability sampling (PPS) in which posterior probabilities are estimated for each class, then samples that maximize the entropy between the posterior probabilities are selected and query by committee (QBC) in which a pool of classifiers is trained on different features to label the set of learning samples. Then, the algorithm chooses the samples with the maximum disagreement between classifiers.

(Nasiri, Naghibzadeh, Yazdi, & Naghibzadeh, 2009)presented a new approach for cardiac arrhythmia disease classification using Support Vector Machine (SVM) and Genetic Algorithm approaches. Twenty-two features from electrocardiogram signal are extracted which represented the across the one cardiac cycle. Genetic Algorithm was used to improve the generalization performance of the SVM classifier. In order this, the design of the SVM classifier is optimized by searching for the best value of the parameters that tune its discriminate function and looking for the best subset of features that optimizes the classification fitness function. Experimental results demonstrate that the approach adopted better classifies ECG signals. Four types of arrhythmias were distinguished with 93% accuracy**.**

A modified combined wavelet transforms technique was developed by (Sexena, Kumar, & Hamde, 2002). The technique was developed to analyze multi-lead electrocardiogram signals for cardiac disease diagnostics. Two wavelets have been used, i.e. a quadratic spline wavelet (QSWT) for QRS detection and the Daubechies six coefficient (DU6) wavelet for P and T detection. A procedure evolved using electrocardiogram parameters with a point scoring system for diagnosis of various cardiac diseases. The consistency and reliability of the identified and measured parameters were confirmed when both the diagnostic criteria gave the same results. (Hadadi, Abdulmounim, & Belauid, 2014) developed for detection of QRS complexities. They presented an algorithm which utilized a modified definition of slope, of ECG signal, as the feature for detection of QRS. A succession of transformations of the filtered and baseline drift corrected ECG signal is used for mining of a newly modified slope-feature. In the presented algorithm, filtering procedure based on moving averages (Owis, Abou-Zied, Youssef, & Kadah, 2001) provides smooth spike-free ECG signal, which is appropriate for slope feature extraction. But the major drawback of the techniques mentioned methods is that they put more emphasis on the detection of QRS complex with less emphasis on T waves. This paper presents a hybrid method for QRS complex detection comprising Hilbert transform with autocorrelation; analysis of slopes, amplitude and widths. Daubechies wavelet as a feature extraction method. The extracted features served as input to the Support Vector Classifier (SVM) which employed a one-to-one multiclass SVM with radial based function.

## II. DISCRETE WAVELET TRANSFORM

The implementation of wavelet transform using a set of wavelet scales and translations obeying some defined rules is called Discrete Wavelet transform. It is an efficient algorithm for calculating the wavelet coefficients of a discrete series. The DWT analyzes the signal at different frequency bands with different resolutions by decomposing the signal into a coarse approximation and detail information. It employs two sets of functions, called scaling functions and wavelet functions, which are associated with low pass and highpass filters, respectively. Signals are decomposed into different frequency band by successive lowpass and highpass filtering of the time domain signal. The original signal a(n) is first passed through a highpass filter b[n] and a lowpass filter c[n]. After the filtering, half of the samples can be eliminated according to the Nyquist's rule, since the signal now has the highest frequency of $\pi/2$radians instead of $\pi$. The signal can, therefore, be subsampled by 2, simply by discarding every other sample(Polikar, 2003).

In CWT, the signals are analyzed using a set of basic functions which relate to each other by simple scaling and translation. In the case of DWT, a time-scale representation of the digital signal is obtained using digital filtering techniques. The signal to be analyzed is passed through filters with different cutoff frequencies at different scales. To receive a set of transform equations for the time discrete case, it is assumed that the samples are equally spaced. Hence, they form a sequence x (n) where n is an integer. The scaling parameter is changed to j, it selects a point on a logarithmic divided axis. j is also considered to be the selected octave(Rioul & Vetterli, 1991). The DWT produces a set of coefficients/features estimated from equation (1)

$$DWT_x(k,j) = (x, \psi_{k,j}) = 2^{\frac{-j}{2}} \sum_{n=-\infty}^{\infty} x(n)\psi(2^{-j}\, n - k) \tag{1}$$

.

These coefficients measure the signal energy distribution in each frequency channel corresponding to the scaling parameter j at the time k. If the wavelets $\psi_{k,j}$ form a set of orthonormal basis functions, the inverse DWT can be expressed as:

$$X(n) = \sum_{j=0}^{\infty} \sum_{K=-\infty}^{\infty} DWT_x(k,j)\psi(2^{-j}\, n - k) \tag{2}$$

If the wavelet transform of a continuous time signal, *x (t)*, is defined as:

$$T(a,b) = \int_{-\infty}^{+\infty} x(t)\psi^*\left(\frac{t-b}{a}\right) dt \tag{3}$$

where $\psi^*(t)$ is the complex conjugate of the analyzing wavelet function $\psi(t)$, $a$ is the dilation parameter of the wavelet and $b$ is the location parameter of the wavelet, then the signal x (t) is filtered by$\psi$. This filter changes its length according to the scale. When j is increased by 1, the wavelet is dilated by 2. In discrete time, this also means that its frequency is halved and the obtained coefficients could be subsampled according to the sampling theorem. Starting with j = 0, the DWT will first compute the coefficients with the highest frequency content (or the highest detail level) and will continue with a filter half the frequency and so on. This can be used to build a filter tree according to this principle and it links the Wavelet Theory to Multiresolution Analysis (MRA) and Subband Coding techniques (Hoffmann, 1998).

The major principle behind DWT is sub-band coding which is the upsampling and downsampling (also known as subsampling) operation of a signal which obeys the Nyquist theorem for sampling. Downsampling a signal reduces the sampling rate or removing some sampling of the signals. For example, Downsampling by two refers to dropping every other sample of the signal. Downsampling by a factor *n* reduces the number of samples in the signal *n* times. Upsampling, on the other hand, corresponds to increasing the sampling rate of a signal by adding new samples to the signal. For example, upsampling by two refers to adding a new sample, usually a zero or an interpolated value, between every two samples of the signal. Upsampling a signal by a factor of *n* increases the number of samples in the signal by a factor of *n*. The resolution of the signal, which is a measure of the amount of detail information in the signal, is changed by the filtering operations, and the scale is changed by upsampling and downsampling operations (Polikar, 2003). The above procedure can be repeated for further decomposition. At every level, the filtering and subsampling will result in half the number of samples (and hence half the time resolution) and half the frequency bands spanned (and hence double the frequency resolution)

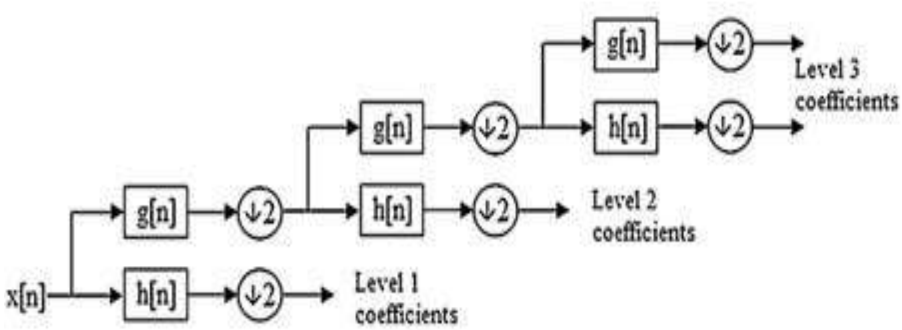


**Figure 2:** Third level filter bank block diagram representation of DWT (Suraj, Mathew Francis, & Niral, 2014)

## III. RESEARCH METHODOLOGY

The block diagram of the developed system for detection of cardiac abnormalities is shown in Figure 2**.** A passband frequency of 5-15 Hz containing a bandpass Finite Impulse Response (FIR) Butterworth filter was used to remove the power-line interference and high-frequency noises from the original ECG signal to improve the signal to noise ratio of the signal. The method combines modified form Analysis of Amplitude, Slopes and Width (Pan & Tompkins, 1985) algorithm and QRS detection complex detection using Hilbert transform and Autocorrelation (Saboo, Behera, & Ari, 2011) to detect the QRS complex and T waves. This combination allows for a more accurate detection of the QRS peaks and the T-waves in an ECG signal.

Modified Analysis of Amplitude, Slopes and Width (MAASW) uses a real-time QRS detection algorithm based on analysis of the slope, amplitude, and width of the QRS complexes of a typical cardiac signal. The algorithm includes a series of filters and operators that perform differentiation, squaring function and integration on the ECG signal, while the slope information was obtained through differentiation. Instead of squaring function an absolute value was used, this makes all data points positive and the QRS detector less gain-sensitive, thus improving its performance. It also eliminates the non-linear amplification caused by squaring the function. For the QRS detection using Hilbert transform and Autocorrelation, the first derivative of the ECG signal with respect to time was taken and its Hilbert transform was used to locate the R peak in the ECG waveform. Then the Autocorrelation was used to estimate the period of one cardiac cycle. Discrete Wavelet Transform (DWT) was used to extract useful features from the QRS complex and T waves detected. Principal Components Analysis (PCA) was used to select 28 Approximate components of the DWT representing a single beat of the ECG.

SVM is usually used for classification tasks introduced by (Vapnik, 1995) and it is used to solve the problem of empirical data modelling which is germane to many engineering applications. In empirical data modelling, a process of induction is used to build up a model of the system, from which it is hoped to deduce responses of the system that have yet to be observed. Ultimately the quantity and quality of the observations govern the performance of this empirical model. These empirical data problems are usually called generalization errors which include approximation error, estimation error and optimization error.

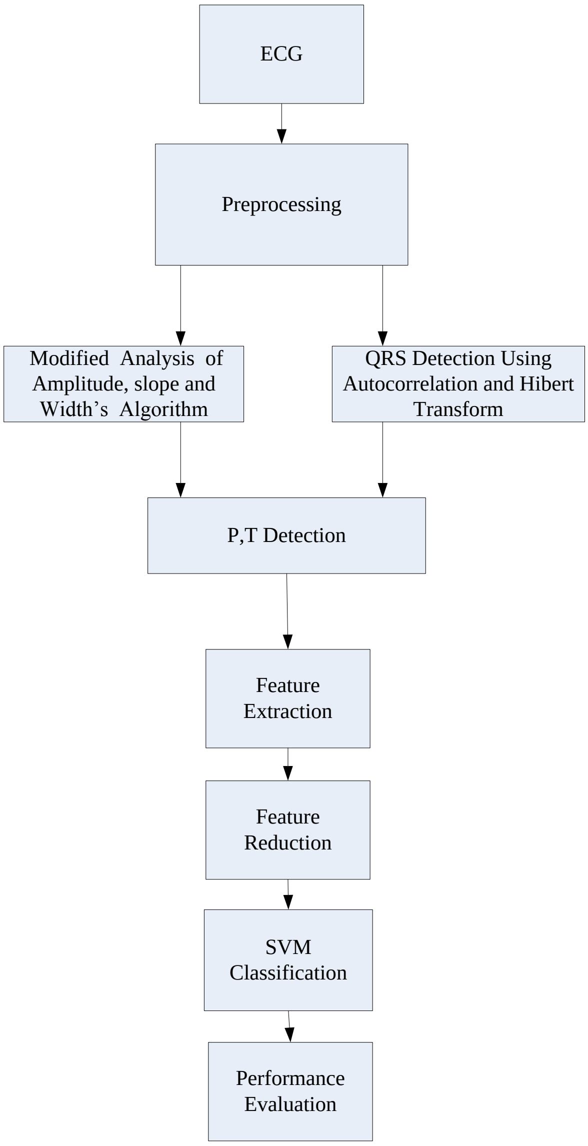


**Figure 3:** Block diagram of the Proposed Cardiac Abnormalities Detection System

A pattern x is given a class by first transforming the pattern into a feature vector and taking the sign of a linear discriminant function. The hyperplane defines a decision boundary in the feature space. The problem specific feature (x) is usually chosen by hand. The parameters w and b are determined by running a learning procedure on a training set. In this paper, the data set from ECG database from Massachusetts Institute Technology/Beth Israel Hospital (MIT-BIH) arrhythmia database were used in training and testing the classifier. The 28-sample sequence of wavelet coefficients was reduced in dimensional using the principal component analysis (PCA) generating a feature of 3-dimension sequence. To generate the 3-dimensional feature for the training process a pre-classified arrhythmia set in the same class were selected. The PCA then generates a transform matrix of from the selected set. For each type of the Arrhythmia beat, the new principal basis was produced. The wavelet coefficients vector was projected into the new basis to obtain the 3-dimensional points. The extracted features were fed as a support vector to classify into different cardiac abnormalities, the SVM receives features as the input and classification were done. For the training set for this research a total of 1200 cases comprising of each case of each cardia cabnormality from the database was used to train the classifier such. The new support vectors were labelled according to the cardiac abnormalities they represented. For the test set consists 6000 cases comprising of 1000 cases of each of cardiac abnormality for testing or validation of the algorithm. One against one SVM with radial based function was used.

The following Performance Evaluation methods were used to validate the proposed algorithm

1. Sensitivity (Se) measures the proportion of actual position which is correctly identified. It is defined by

$$Se(\%) = \frac{TP}{TP + FN} \times 100 \qquad (4)$$

2. Specificity measures the proportion of negatives which are correctly identified. It is defined by

$$S(\%) = \frac{FN}{TP + FN} \times 100 \qquad (5)$$

3. The measurement of delineation accuracy, Acc, is given by

$$Acc(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \qquad (6)$$

Where TP represents for True Positive, TN represents for True Negative, FP represents for False Positive and FN represents for False Negative

## IV RESULTS

The developed algorithm (MAASW) was implemented in MATLAB 8.1.0.64 and performance of the algorithm was validated using Accuracy, Sensitivity and Specificity. Experimental Result shows an accuracy of 82.33%, a sensitivity of 82.33% and specificity of 96.47%. Tables 1 and 2 shows the comparison of the developed method with existing methods. There is an increase in the accuracy of the system when compared with earlier methods of Hilbert with autocorrelation and Analysis of Amplitude, slopes and width's Algorithm (AASWA) that have 81.68% and 81.67% respectively. MAASW also gave higher sensitivity values, while it retains the same specificityvalue with the AASWA

Figures 4 and 5 represents the raw ECG signal and QRS, T waves on Filtered Signal for Analysis of Amplitude, slopes and width's Algorithm and Figure 6 represents the output of Hilbert with autocorrelation method. The circled part of figure 5 shows where T waves have been detected.

**Table 1:** Comparison of the methods adopted by this work

| **Methods** | **Performance** | | |
|---|---|---|---|
| | Accuracy (%) | Sensitivity (%) | Specificity (%) |
| Hilbert with autocorrelation | 81.68 | 81.68 | 96.34 |
| Analysis of Amplitude, slopes and width's Algorithm | 81.67 | 76.90 | 96.47 |
| MAASW | 82.33 | 82.33 | 96.47 |

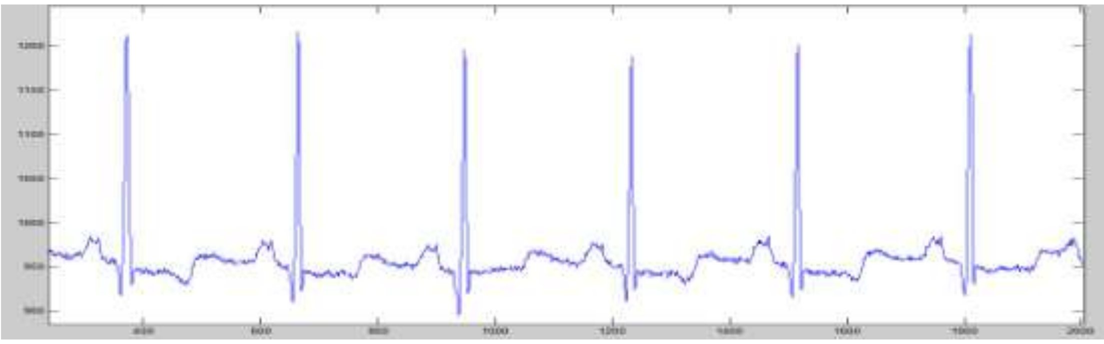

**Figure 4:** Raw ECG Signal

**Table 2:** Comparison with existing works

| Classifier model | Characteristics | | Performance |
|---|---|---|---|
| SVM Classifier (Melgani & Pasolli, 2010) | 1) SVM with Margin sampling | | 79% |
| | 2)SVM with Posterior probability sampling | | 79% |
| | 3) SVM with Query by committee | | 79% |
| ECG Arrhythmia Classification with SVM and Genetic Algorithm (Nasiri, Naghibzadeh, Yazdi, & Naghibzadeh, 2009) | With PCA | Linear kernel | 80.00% |
| | | Polynomial kernel | 79.25% |
| | Without PCA | Linear kernel | 79.23% |
| | | Polynomial kernel | 78.08% |
| **MAASW** | DWT and SVM | SVM with Radial Based Function | **82.33%** |

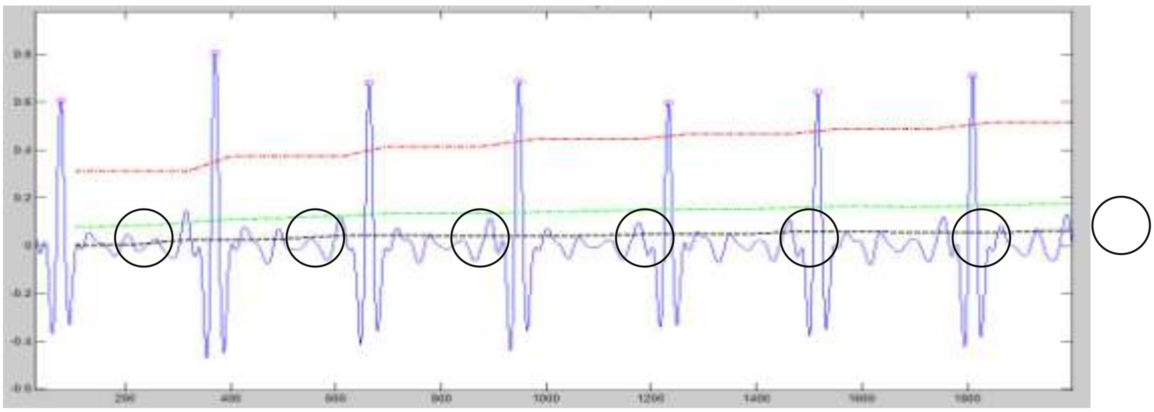

**Figure 5:** QRS on Filtered Signal

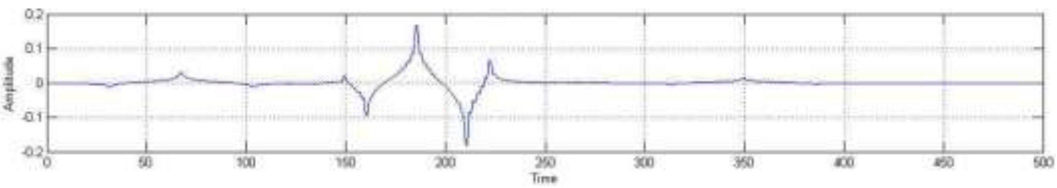


**Figure 6:** Output of the Hilbert Transform

## IV CONCLUSION

An algorithm for detection of cardiac abnormalities using wavelet transform and SVM has been developed. The input ECG signal was preprocessed to remove both high frequency and low-frequency noise using Finite Impulse Response Filter with a pass band of frequencies 5-15Hz. QRS complex in the ECG was then detected from the preprocessed ECG. A hybrid method comprising Hilbert transform with autocorrelation and analysis of slopes, amplitude, widths was used for QRS complex detection and extraction of Q, R, S, T waves. The extracted features served as input to the Support Vector Classifier (SVM) which employed a one-to-one multiclass SVM with radial based function. The algorithm was implemented in MATLAB 8.1.0.64 version. Performance evaluation was done using the metrics of Accuracy, Sensitivity and Specificity.

## REFERENCES


[1]. Chouhan, V. S., & Mehta, S. S. (2007). Total Removal of Baseline Drift from ECG Signal. *International Conference on Computing: Theory and Applications*, (pp. 512-515).

[2]. Chouhan, V. S., & Mehta, S. S. (2008). Detection of QRS Complexes in 12-lead ECG using Adaptive Quantized Threshold. *International Journal of Computer Science and Network Sercurity, 8*(1).

[3]. Distribution, M.-B. D. (1998). *Massachusett Institute of Technology, 77 Massachusett Avenue Cambridge MA02139.* Retrieved September 2014, from www.physionet.org/physiobank/database/mitdb.

[4]. Ghorbanian, P., Jalali, A., Ghaffari, A., & Nataraj, C. (2012). An improved procedure for detection of heart arrhythmias with novel pre-processing techniques. *Expert Systems, 29*, 478-491.

[5]. Hadadi, R. A. (2009). Discrete Wavelet Transform based Algorithm for Recognition of QRS Complex. *World of Computer Science and Information Technology Journal*, 127-132.

[6]. Hadadi, R., Abdulmounim, E., & Belauid, A. (2014). Discerete Wavelet Transform based Algorithm for Recognition of QRS Complex. *World of Computer Science and Information Technology Journal, 4*, 127-132.

[7]. Hoffmann, R. (1998). *Signal Analyse-und Erkennung.* Springer Verlag.

[8]. Hoffmann, R. (1998). *Signal Analyse-und Erkennung.* Springer Verlag.

[9]. Mahmoodabadi, S., Ahmadian, A., & Abolhasani, M. D. (2005). ECG Featute Extraction using Daubechies Wavelets. *5th IASTED International conference on Visualization, Imaging and Image Procesing*, (pp. 343-348).

[10]. Melgani, E., & Pasolli, F. (2010). Active Learning Methods for Electrocardiographic Signal Classification. *IEEE Transaction on information Technology in Biomedicine, 14*(6).

[11]. Murray, C., Lopex, A., Mohsen, N., & Haidong, W. (2015). *GBD 2015 Mortality and cause of Death Collaborators*. Retrieved September 11th, 2016, from www.thelancet.com/journal/lancet/article

[12]. Nasiri, J. A., Naghibzadeh, M., Yazdi, M., & Naghibzadeh, B. (2009). ECG Arrhythmia Classification with Support Vector Machines and Genetic Algorithm. *Third UK Sim European Symposium on Computer Modeling and Simulation.*

[13]. Organisation, W. H. (2007). *Cardiovascular Diseases Fact sheet No 317*. Retrieved March 12th, 2013, from www.who.int/cardiovascular_diseases

[14]. Owis, M. A.-z. (2018). Robust Feature Extraction for ECG Signal based Nonlinear Dynamical Modelling. *23rd Annual Conference IEEE Engineering in Medicine and Biology Society*, (pp. 1585-1588).

[15]. Owis, M., Abou-Zied, A., Youssef, A. B., & Kadah, Y. (2001). Robust Feature Extraction for ECG Signals based Nonlinear Dynamical Modeling. *23rd Annual Conference IEEE Engineering in Medicine and Biology Society*, (pp. 1585-1588).

[16]. Pan, J., & Tompkins, W. J. (1985). A real-time QRS Detection Algorithm. *IEEE Transaction on Information Technology in Biomedicine Engineering*, 230-236.

[17]. Polikar, R. (2003). *Wavelet Tutorial.* Retrieved August 20th, 2017, from web.iitd.ac.in/sumeet/wavelet tutorial.

[18]. Rioul, R., & Vetterli, M. (1991). Wavelets and Signal Processing. *IEEE SP Magazine*, 14-38.

[19]. Saboo, J. P., Behera, S., & Ari, S. (2011). A Novel Technique for QRS Complex Detection in ECG Signal based on Hilbert Transform and Autocorrelation. *National Institute of Rourkela India*.

[20]. Sexena, S., Kumar, V., & Hamde, S. (2002). Feature Extraction from ECG Signals using Wavelet Transforms for Disease Diagnostics. *International Journal of Systems Science, 33*(13), 1073-1085.

[21]. Sokolow, M., Mcllory M, B., & Cheithin, M. D. (1990). *Clinical Cardiology* (5th ed.). Vlange Medical Books.

[22]. Suraj, A., Mathew Francis, T. S., & Niral, T. (2014). Discrete Wavelet Transform based image fusion and denoising in FPGA. *Journal of Electrical System and Infromation Technology*, 72-81.

[23]. Vapnik, V. (1995). *The Nature of Statistical Learning Theory.* New York: Springer Verlag.

[24]. Xu, X., & Liu, Y. (2004). ECG QRS Complex Detection using Slpoe Vector Waveform (SVM) Algorithm. *26th Annual International Conference of the IEEE EMBS*, (pp. 3597-3600).


Ojo, J. A. "Cardiac Abnormalities Detection using Wavelet Transform and Support Vector Machine." American Journal of Engineering Research (AJER), vol. 6, no. 10, 2017, pp. 28-35