International Journal of Advanced Science and Technology
Vol. 29, No. 9s, (2020), pp. 6803-6815

Wavelet Transform Application for Biomedical Data Mining – A Review

T.Selvapriya (Ph.D Scholar), Dr.V.R.Kavitha (Associate Professor)
Department of Computer Applications, PSG College of Arts and Science Coimbatore, Tamil Nadu India

Abstract

Data mining plays a vital role in several biomedical applications. In medical systems several data in the form of text, image, signal and video are present. These data are to be mined to understand the normal or abnormal activities of several organs. In past manual methods were used by the physician to investigate the disease, but nowadays autonomous methods in data mining does the work in an efficient way. For analysis the information has to be extracted. Several methods like time, frequency and time scale analysis were present. Among them the Wavelet transform is efficient. The paper presents the detail investigation of the wavelet transform in Data Mining for investigation biomedical data is considered. The analysis has widened in De noising, extraction of features and classification. From the review its been found that the wavelet transform plays a vital role in Data mining of Biomedical data. The wavelet method for De noising, Extraction, Classification were analyzed.

Keywords: *ECG, Data mining, Wavelet transform, feature extraction, De noising, classifier*

1. Introduction

The objective of data mining is to identify valid novel, potentially useful, and understandable correlations and patterns in existing data [Chung and Gray 1999]. Finding useful patterns in data is known by different names (including data mining) in different communities (e.g., knowledge extraction, information discovery, information harvesting, data archeology, and data pattern processing) [Fayyad, et al, 1996]. The term ―data mining‖ is primarily used by statisticians, database researchers, and the MIS and business communities. The term Knowledge Discovery in Databases (KDD) is generally used to refer to the overall process of discovering useful knowledge from data, where data mining is a particular step in this process.[Fayyad, et al, 1996; Peacock, 1998a; Han and Kamber, 2000] The additional steps in the KDD process, such as data preparation, data selection, data cleaning, and proper interpretation of the results of the data mining process, ensure that useful knowledge is derived from the data. Data mining is an extension of traditional data analysis and statistical approaches in that it incorporates analytical techniques drawn from a range of disciplines including, but not limited to Communications of the Association for Information Systems (Volume 8, 2002) 267-296 Data Mining: A Conceptual Overview by J. Jackson

- numerical analysis,
- pattern matching and areas of artificial intelligence such as machine learning,
- neural networks and genetic algorithms.

While many data mining tasks follow a traditional, hypothesis-driven data analysis approach, it is common to employ an opportunistic, data driven approach that encourages the pattern detection algorithms to find useful trends, patterns, and relationships.

Essentially, the two types of data mining approaches differ in whether they seek to build models or to find patterns. The first approach, concerned with building models is, apart from the problems inherent from the large sizes of the data sets, similar to conventional exploratory statistical methods. The objective is to produce an overall summary of a set of data to identify and describe the main features of the shape of the distribution [Hand 1998]. Examples of such models include a cluster analysis partition of a set of data, a regression model for prediction, and a tree-based classification rule. In model building, a distinction is sometimes made between empirical and mechanistic models [Box and Hunter 1965; Cox 1990; Hand 1995]. The former (also sometimes called operational) seeks to model relationships without basing them on any underlying theory. The latter (sometimes called substantive or phenomenological) are based on some theory or mechanism for the underlying data generating process. Data mining, almost by definition, is primarily concerned with the operational.The second type of data mining approach, pattern detection, seeks to identify small (but nonetheless possibly important) departures from the norm, to detect unusual patterns of behavior. Examples include unusual spending patterns in credit card usage (for fraud detection), sporadic waveforms in EEG traces, and objects with patterns of characteristics unlike others. It is this class of strategies that led to the notion of data mining as seeking ―nuggets‖ of information among the mass of data. In general, business databases pose a unique problem for pattern extraction because of their complexity.

Complexity arises from anomalies such as discontinuity, noise, ambiguity, and incompleteness [Fayyad, Piatetsky-Shapiro, and Smyth, 1996]. And while most data mining algorithms are able to separate the effects of such irrelevant attributes in determining the actual pattern, the predictive power of the mining algorithms may decrease as the number of these anomalies increase [Rajagopalan and Krovi, 2002].

Biomedical signal processing. Bio signal processing has been rapidly developing, increasing the understanding of complex biological processes in a wide variety of areas. Wavelet transform (Daubechies, 1991) is a powerful time frequency approach which has been applied to multiple domains of bio signal processing, such as EMG (e.g. Englehart, Hudgins, & Parker, 2001), EEG (e.g. Kurt, Sezgin, Akin, Kirbas, & Bayram, 2009; Rosso et al., 2001; Subasi, 2005; Ting, Guo-zheng, Bang-hua, & Hong, 2008), ECG (e.g. Engin, Fedakar, Engin, & Korurek, 2007; Manikandan & Dandapat, 2008; Singh & Tiwari, 2006), VPA (e.g. Rafiee, Rafiee, & Michaelsen, 2009). A significant focus on the application of wavelet transforms (e.g. Englehart et al., 2001; Farina, Lucas, & Doncarli, 2008) has permitted rapid development in the field. However, the selection of the most appropriate mother wavelet to characterize commonalities amongst signals within a given domain is still lacking in bio signal processing. The main contributions to find the optimum basis function can be found in several papers (e.g. Brechet, Lucas, Doncarli, & Farina, 2007; Farina, do Nascimento, Lucas, & Doncarli, 2007; Flanders, 2002; Landolsi,2006; Lucas, Gaufriau, Pascual, Doncarli, & Farina, 2008; Rafiee & Tse, 2009; Singh & Tiwari, 2006; Tse, Yang, & Tam, 2004). The mother wavelet function is the main base of wavelet transforms that would permit identification of correlated coefficients across multiple signals. The more similar the mother wavelet function is to the wavelet coefficients across signals, the more precisely the signal of interest can be identified and isolated; hence, identification of a mother wavelet function is of paramount significance.

The Daubechies (db) wavelet functions (Daubechies, 1988) have been applied in several areas with the lower orders (db1 to db20) used most often (Rafiee & Tse, 2009).

There are two types of Wavelet Transform.

1.2. Continuous Wavelet Transform

The continuous wavelet transform was developed as an alternative approach to the short time Fourier transform to overcome the resolution problem[3]. The continuous wavelet transform is defined as follows

$$CWT_x^{\psi}(\tau, s) = \Psi_x^{\psi}(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t)\psi^* \left(\frac{t-\tau}{s}\right) dt$$

The above equation , the transformed signal is a function of two variables, tau and s , the translation and scale parameters, respectively. psi(t) is the transforming function, and it is called the mother wavelet . The term mother wavelet gets its name due to two important properties of the wavelet analysis as explained below:

The term wavelet means a small wave. The smallness refers to the condition that this window function is of finite length. The wave refers to the condition that this function is oscillatory. The term mother implies that the functions with different region of support that are used in the transformation process are derived from one main function, or the mother wavelet. The mother wavelet is defined as a prototype for generating the other window functions.

1.3. Computation of the CWT:

The signal to be analyzed is taken.
The mother wavelet is chosen and the computation is begun with s = 1. The CWT is computed for all values of s. the wavelet will dilate is s increases and compresses when s is decreased.
The wavelet is placed in the beginning of the signal at the point which corresponds to time = 0.

The wavelet is multiplied with the signal and integrated over all times. The result is then multiplied by the constant 1/sqrts.

The above step normalizes the energy so that the transformed signal has same energyat every scale.

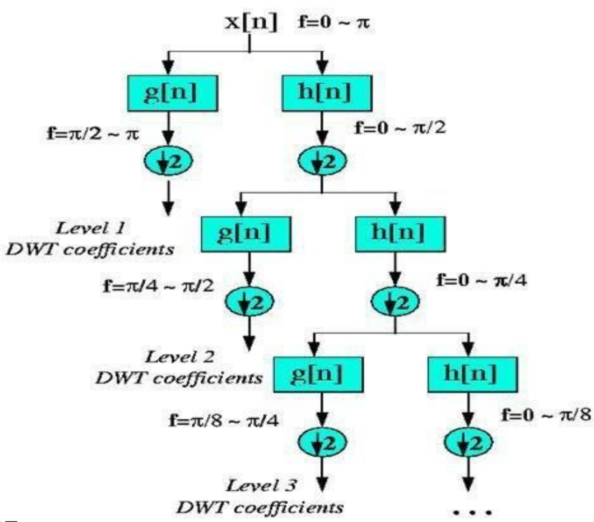
The wavelet at scale s =1 is then shifted to the right by t and the above steps arerepeated until the wavelet reaches the end of the signal.

1.4. The Discrete Wavelet Transform

Discretized continuous wavelet transform enables the computation of the continuous wavelet transform by computers, but it is not a true discrete transform. In fact the wavelet series is simply a sampled version of the CWT, and the information provided by it is highly redundant as far as the reconstruction of the signal is concerned. This redundancy requires a significant amount of computation time and resources. Whereas the discrete wavelet transform (DWT), provides sufficient information both for analysis and synthesis of the original signal, with a significant reduction in thecomputation time. The DWT is considerably easier to implement when compared to the CWT. The basic concepts of the DWT are introduced in this paper along with its properties and the algorithms used to compute it.

1.5. Wavelet Decomposition and Construction:

Wavelet analysis is the breaking up or sub sampling of a signal into shifted and scaled versions of the original (or mother) wavelet. It conserves the time-frequency components of a signal at different resolution and scales. Noise components from the contaminated ECG signal can be removed at a good resolution of wavelet analysis. The discrete wavelet transform DWT can be achieved by successive low pass and high pass filter at discrete time domain. x[n] is the input signal which pass through a high pass filter where the impulse response is h[n]. The same signal x[n] simultaneously pass through the low pass filter with an impulse response g[n].The output of the high pass filter gives the detail coefficient and the output of the low pass filter gives the approximate coefficient. The filter output is given in Y_{high} and Y_{low} where kvaries from -∞ to ∞ .



The signal is then sub sampled by 2 since half of the number of samples are redundant. This doubles the scale. This procedure can mathematically be expressed as where Yhigh[k] and Ylow[k] are the outputs of the high pass and low pass filters, respectively, after sub sampling by 2

.

The reconstruction in this case is very easy since half band filters form orthonormal bases. The above procedure is followed in reverse order for the reconstruction.

$$y_{high}[k] = \sum_n x[n] \cdot g[-n + 2k]$$

$$y_{low}[k] = \sum_n x[n] \cdot h[-n + 2k]$$

$$y[n] = \sum_{k=-\infty}^{\infty} h[k] \cdot x[2n - k]$$

$$y_{high}[k] = \sum_n x[n] \cdot g[2k - n]$$

$$y_{low}[k] = \sum_n x[n] \cdot h[2k - n]$$

Therefore, the reconstruction formula becomes (for each layer).

$$x[n] = \sum_{k=-\infty}^{\infty} \left(y_{high}[k] \cdot g[-n + 2k]\right) + \left(y_{low}[k] \cdot h[-n + 2k]\right)$$

Mother Wavelets Used For ECG Signal:

Daubechies Wavelet Haar Wavelet Symlet
Cubic B-Spline Wavelet

2. Noises in biomedical signal

2.1 Baseline Wander

It is a low-frequency noise which occurs in the range 0.15 Hz to 0.3 Hz in ECG. Due to poor contact of electrodes, respiration, and perspiration or body movements this kind of noise occurs when the patient breathes rapidly or coughs while ECG is being recorded.

2.2 Power Line Interference

This is a high frequency noise whose value is 50 Hz/60 Hz. Due to loose contact or dust in the electrodes this type of noise occurs and it can also be produced by improper grounding of adjacent devices.

2.3 Instrumentation Noise

The electrical equipment which is used in ECG measurements also contributes noise. Electrode probes, cables, signal processor/amplifier, and the Analog-to-Digital converter are the major sources of this form of noise. Unfortunately instrumentation noise cannot be eliminated, but it can be reduced through higher quality equipment and careful circuit design. One type of electrical noise is resistor thermal noise (also known as Johnson noise). Random fluctuations of the electrons due to thermal agitation produce this noise.

2.4 Motion Artifact

Motion artifacts are basically caused by muscle tremor/noise, which will result in minute electrode motion and change the electrode-skin impedance. When one‘s skeletal muscles undergo tremors, the ECG is interfered by seemingly random activities. Motion artifacts do not manifest under a fixed pattern and instead has a tendency to suddenly change depending on the users movements and actions. Several motion artifacts were recorded during ECG monitoring with our in-house hardware board. They have various shapes, but they are similar to displays motion artifacts detected with our board. In addition, the bandwidth of motion artifacts changes over time, but usually manifest in the form of low-frequency elements. The useful ECG bandwidth for patient monitoring or healthcare purpose is between 0.05 Hz and 35 Hz. Therefore, low frequency noise elements that resulting from respiration and motion artifacts are known to overlap with the ST segment (0.8 Hz or below) of the ECG signal, and using a high pass filter to eliminate low frequency noise elements like these can lead to distortion of the ST segment, which contains clinical data on heart ischemia and myocardial infarction. Thus, simple filter sets can‘t effectively remove motion artifacts while maintaining signal integrity.

2.5 DC Noise

The three main noise sources in instrumentation amplifiers are the flicker and thermal noise of the amplifier circuit itself, electromagnetic and electrostatic signals coupling through the cables and the human body, and the electrode noise. Using shorter length of cables of the leads greatly minimizes the problem of noise coupling through cables.

The electrode impedance by itself will generate some noise which corrupts the signal being detected. The type of electrode used, its chemical makeup and the surface of 19 the human body determine the value of this impedance and the noise voltage it produces

The circuit noise consists of the thermal noise and flicker noise

3. Review of Denoising methods

Pinjala N Malleswari (et..al) in the paper Volume-8, Issue-1, May 2019 Signal processing is an efficient mathematical tool in biomedical engineering, especially to identify and eliminate noises in ECG. A corrupted ECG signal may lead to wrong diagnosis of the heart condition, which indications to erroneous detection of cardiac disorders. Preprocessing is widely used in the heart rate, QR interval, and extraction of features and distribution of QRS complexes. It eliminates the contamination in ECG signal caused by many factors such as body activities, ECG leads, loose contact between skin and electrode, etc. Electrocardiogram signal is degraded by the following types of noises. Many number of techniques are available to denoise ECG signals.The proposed technique includes denoising ECG waveforms by wavelet family such as Daubechies, Haar, Symletand, BiorSplines, digital filters like IIR notch and window based FIR Filters using Hanning, Hamming, and Blackman and Rectangular windows.

Step 1: Decompose the noisy data by choosing particular wavelet from its family for _7‘ levels, which gives many number of coefficients of different lengths.

Step 2: Make the uppermost level coefficients a7 and d7 as zeroes.

Step 3: Apply soft thresholding to the remaining coefficients [3] (d6, d5, d4, d3, d2 and d1).

Step 4: Reconstruct the signal using modified coefficients, with the same wavelet family up to _7‘ levels.

Step 5: Apply noisy ecg signal to the IIR notch filter to get the denoised signal.

Step 6: Apply noisy signal to the window based FIR filter to get the noise free signal.

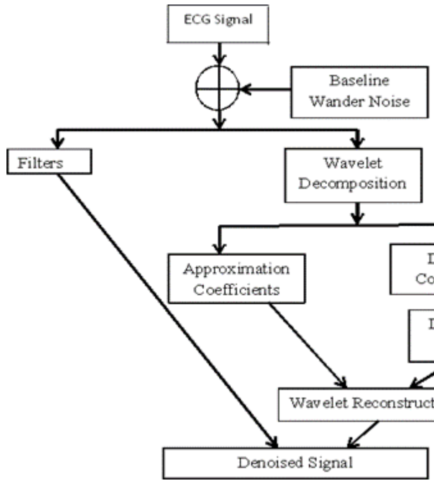


Fig. 1 Flow Chart of Proposed Method

The work flow of proposed method is shown in Fig.1. It clearly describes the step by step procedure.

3.1 Selecting the Required De-Noising Points

Dengyong Zhang(et..al) in the paper the coefficient energy of each layer calculated above, the wavelet energy curve can be obtained, and then the position of the maximum and minimum points of the energy curve can be calculated, that is the decomposition level where they are. According to the positions of the maximum and minimum points of the calculated wavelet energy curve, the variation points of the number of detail layers that need to be de- noised can be judged. The previous level of the variation points is considered to be de-noised, while the level of the variation point and the subsequent level are not considered to de-noise.In practical processing, we only locate the initial several detail coefficients that required processing, without considering the low-frequency components behind. Therefore the first position of the energy change point is particularly important.

S. Banerjee(et..al) Measurement 45 (2012) 474–487 the paper Denoising the signal In this stage the noise frequency identification and their elimination is carried out in two basic steps, viz,
(A) high frequency noise removal,
(B) baseline wander correction and power line interference removal. High frequency noise removal Electrosurgical noise and muscle contraction noise are high frequency noise. Electrosurgical noise completely destroys the ECG and without its removal an accurate feature extraction cannot be accomplished. The frequency content of this noise is 100 kHz–10 MHz. Since the sampling frequency of the data is 1 kHz an aliased version gets added to the ECG. The muscle contraction noise has a frequency range from dc-10 kHz. These two noises are eliminated by discarding the detail coefficients D1, D2. After removal of these noises the remaining ECG signal ranges from 0 to 125 Hz. Baseline correction and power line interference removal the drift of the baseline is caused due to respiration and is likely to be as a nearly sinusoidal component and the frequency of respiration gets added to the ECG during its acquisition. The baseline variation frequency is 0.15– 0.8 Hz. Motion artifacts are transient baseline changes caused by change in the electrode skin impedance with electrode motion. The baseline disturbances caused by motion artifact can be assumed as a signal resembling one cycle of a sine wave and is within the frequency range of baseline drift. These two type of noises stated above can be eliminated by removal of the lowest frequency component, after decomposition of the ECG signal coefficient A10 contains this frequency along with the DC component of the ECG. Discarding A10 frequency band and reconstructing the signal eliminates these two noises. eliminated by the dyadic scale DWT technique by identification of the frequency band containing the frequency range of 59.5–60.5 Hz. The signal output from stage B is in the frequency band 0–125 Hz in Fig. 3. On further decomposing the signal, the low pass component (A5) will have 0–62.5 Hz and high pass component (D5) will have 62.5–125 Hz. As shown in Fig. 2 the

noise is in the low pass component, so it is decomposed. The next level decomposition yields 0–32.25 Hz signal component (say D5a) and 32.5–62.5 Hz signal (say D5b). Now the high pass component D5b (containing the 60 Hz noise) is decomposed. The signal component D5b containing the noise frequency is continuously decomposed until the frequency band of 59.5–60.5 Hz is obtained. This component is rejected and the signal from the rest of the decomposed components is reconstructed. This reconstructed signal is the noise free signal, taken in an array named ECG_DENO. Fig. 5 shows a typical representation of original noisy signal, extracted high and low frequency noises and noisefree signal. This is shown in stage B of Fig. 3.

For example the steps in the ECG signal analysis is given below

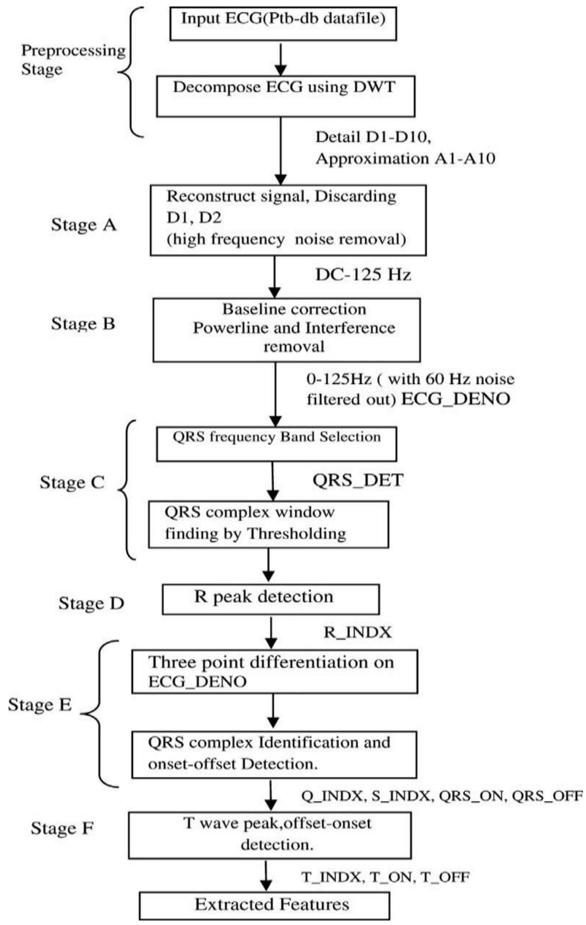


Fig. 3. Block schematic of the processing stages.

Power line interference contains a 60 Hz noise intertwined with the ECG. In India the power line frequency is 50 Hz. In this work, ptb-db data files under Physionet are used for evaluation of the proposed method, where the power line frequency is 60 Hz. Electrode contact noise is transient interference caused by loss of contact caused by the electrode and skin. During this time short duration power line interference corrupts the ECG. These two noises can be eliminated by the dyadic scale DWT technique by identification of the frequency band containing the frequency range of 59.5–60.5 Hz. The signal output from stage B is in the frequency band 0–125 Hz in Fig. 3. On further decomposing the signal, the low pass component (A5) will have 0–62.5 Hz and high pass component (D5) will have 62.5–125 Hz. As shown in Fig. 2 the noise is in the low pass component, so it is decomposed. The next level decomposition yields 0–32.25 Hz signal component (say D5a) and 32.5–62.5 Hz signal (say D5b). Now the high pass component D5b (containing the 60 Hz noise) is decomposed. The signal component D5b containing the noise frequency is continuously decomposed until the frequency band of 59.5–60.5 Hz is obtained. This component is rejected and the signal from the rest of the decomposed components is reconstructed. This reconstructed signal is the noise free signal, taken in an array named ECG_DENO. Fig. 5 shows a typical representation of original noisy signal, extracted high and low frequency noises and noisefree signal. This is shown in stage B of Fig. 3.

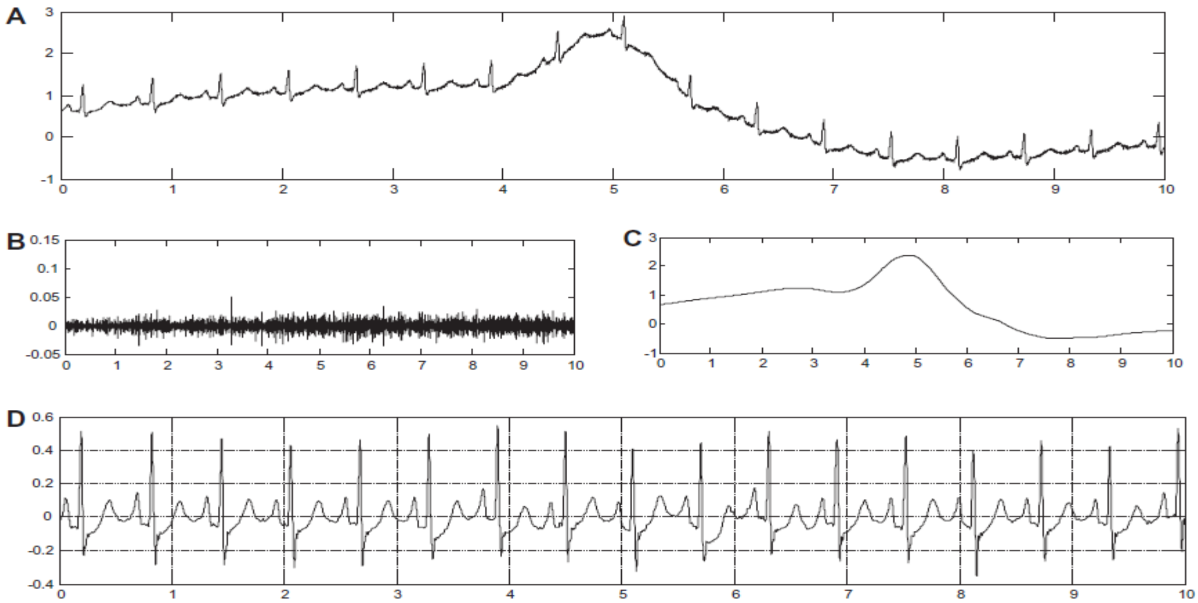


Fig. 5. A. Original ECG signal, B. High Frequency Noise, C. Low Frequency noise, D. Denoised and Baseline Corrected ECG signal

4. Feature Extraction of wavelet transforms:

Dean Cvetkovic (et..al) in his paper he discussed the extraction of wavelet transforms feature is a distinctive or characteristic measurement, transform, structural component extracted from a segment of a pattern. Features are used to represent patterns with the goal of minimizing the loss of important information. The feature vector, which is composed of the set of all features used to describe a pattern, reduces the dimensional space needed to represent that pattern.This, in effect, means that the set of all features that could be used to describe a given pattern (a large and in theory an infinite number as very small changes in some parameters are allowed to separate different features) is limited to those actually represented in the feature vector. One purpose of the dimension reduction is satisfying engineering constraints in software and hardware complexity, the cost of data processing, and the (why else?) desirability of compressing pattern information. In addition, the classification often becomes more accurate when the pattern is simplified by including important features or properties only .A feature extraction is the determination of a feature or a feature vector from a pattern vector. In order to make pattern processing problems solvable one needs to convert patterns into features, which become condensed representations of patterns, ideally containing only salient information. Feature extraction methods could be based on either calculating statistical characteristics or producing syntactic descriptions. The feature selection process usually is designed to provide a means for choosing the features which are best for classification optimized against on various criteria. The feature selection process performed on a set of predetermined features. Features are selected based on either best representation of a given class of signals, or best distinction between classes. Therefore, feature selection plays an important role in classifying systems such as neural networks For the purpose of classification problems, the classifying system has usually been implemented with rules using if then clauses, which state the conditions of certain attributes and resulting rules. However, it has proven to be a difficult and time consuming method. From the viewpoint of managing large quantities of data, it would still be most useful if irrelevant or redundant attributes could be segregated from relevant and important ones, although the exact governing rules may not be known. In this case, the process of extracting useful information from a large data set can begreatly facilitated. Various methodologies of automated diagnosis have been adopted, however the entire process can generally be subdivided into a number of disjoint processing modules: preprocessing, feature extraction/selection, and classification. The signal acquisition, artefact removing, averaging, thresholding, signal enhancement, and edge detection are the main operations in the course of preprocessing. The accuracy of signal/image acquisition is of great importance since it contributes significantly to the overall classification result. The markers are subsequently processed by the feature extraction module. The module of feature selection is an optional stage, whereby the feature vector is reduced in size including only, from the classification viewpoint, what may be considered as the most relevant features required for discrimination. The classification module is the final stage in automated diagnosis. It examines the input feature vector and based on its algorithmic nature, produces a suggestive hypothesis.In the feature extraction stage, numerous different methods can be used so that several diverse features can be extracted from the same raw data. The wavelet transform (WT) provides very general techniques which can be applied to many tasks in signal processing. Wavelets are ideally suited for the analysis of sudden short-duration signal changes. One very important application is the ability to compute and manipulate data in compressed parameters which are often called features. Thus, the time- varying biomedical signal, consisting of many data points, can be compressed into a few parameters by the usage of the WT. These parameters characterize the behavior of the time-varying biomedical signal. This feature of using a smaller number of parameters to represent the time-varying biomedical signal is particularly important for recognition and diagnostic purposes. The objective of the present study in the field of automated detection of changes in time-varying biomedical signals is to extract the representative features of the signals under study in order to obtain the accurate classification models.

Jaysree (et..al) discussed Feature extraction involves reducing the amount of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which over fits the training sample and generalizes poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. Cross wavelet transform generates WCS and WCOH, which are matrices containing the WCS and WCOH between two signals.The features such as maximum and mean value are extracted from Wavelet Cross Spectrum and it will be Feature extraction involves reducing the amount of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which overfits the training sample and generalizes poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. Cross wavelet transform generates WCS and WCOH, which are matrices containing the WCS and WCOH between two signals. The features such as maximum and mean value are extracted from Wavelet Cross Spectrum and it will be suitable and should be taken for classification of ECG patterns. So, the data obtained after the extraction of features may contain considerable information. The features are used as input to the nearest neighborhood classifier for classifying normal and abnormal patterns.

Hari Mohan Rai(et..al) in his paper he discussed the feature extraction of Wavelet transforms After the noise elimination, baseline wanders removal and peak detection it is necessary to extract the feature of the ECG waveform in order to use it in the next stage of ECG signal analysis. The ability to manipulate and compute the data in compressed parameters form is one of the most important application of wavelet transform, are often known as features. Feature extraction is the most important step in pattern recognition. There are several ways to extract the feature of ECG signal.

In this work, there are two types of features are extracted of ECG waveforms. (i) Morphological feature of ECG signal.
(ii) Wavelet co-efficient based features.

4.1 Meyer Wavelet Transform

Pirambu preethika (et..al) discussed Meyer wavelet transform (MVT) is applied to solve the issue of non-stationary feature of ECG signals. The Meyer wavelet is an orthogonal wavelet introduced by Yves Meyer. It is a kind of

$$\gamma(x) = \begin{cases} 0 & if\ x < 0, \\ x & if\ 0 < x < 1 \\ 1 & if\ x > 1 \end{cases}$$

continuous wavelet and mainly used in multi-fault classification, fractal random fields and adaptive filters. This MVT is substantially differentiable with unlimited provision and well-defined in domain of frequency domain in terms as The reason to uses MVT wavelet is that it has additional beneficial features compared to the standard DWT and FFT

$$\psi(w) = \begin{cases} \frac{1}{\sqrt{2\pi}} \sin\left(\frac{\pi}{2} \gamma\left(\frac{3|w|}{2\pi} - 1\right)\right) e^{\frac{jw}{2}} & if\ \frac{2\pi}{3} < |w| < \frac{4\pi}{3} \\ \frac{1}{\sqrt{2\pi}} \cos\left(\frac{\pi}{2} \gamma\left(\frac{3|w|}{4\pi} - 1\right)\right) e^{\frac{jw}{2}} & if\ \frac{4\pi}{3} < |w| < \frac{8\pi}{3} \\ 0 & otherwise \end{cases}$$

$$\gamma(x) = \begin{cases} 0 & if\ x < 0, \\ x & if\ 0 < x < 1 \\ 1 & if\ x > 1 \end{cases}$$

The reason to uses MVT wavelet is that it has additional beneficial features compared to the standard DWT and FFT wavelet transforms. Because it could automatically adjust to the different ECG signal features, by adapting large size window to look for long lived low frequency signals aspects and it uses small window to analyse the short time high frequency components.

4.2 Classification methods

Taiyong Li discussed his paper As for the classifiers of ECG, in theory, any multi-class classifier can be used in ECG classification. In practice, the most commonly used classifiers include support vector machine (SVM) artificial neural network (ANN), K-nearest neighbours (KNN) and decision tree (DT) SVM is one of the most popular ECG classifiers. used a multiclass SVM with the error output codes to build a ECG classifier based on the features calculated from the wavelet coefficients.

Osowski et al. presented a new approach for ECG classification by combining SVM with the features extracted by two preprocessing methods, and the results on recognizing the 13 heart rhythm types showed that the proposed method was reliable and advantageous.

Mohammadzadeh et al. used SVM and the generalized discriminant analysis (GDA) feature reduction scheme to classify cardiac arrhythmia from the heart rate variability (HRV) signa]. Some variations of SVM have also been applied in ECG classification, such as least square SVM , hierarchical SVM, weighted SVM and SVM combined with particle swarm optimization(PSO) Multi-layer perception (MLP) and probabilistic neural network (PNN) are the most popular ECG classifiers associated with ANN. The authors in used the sequential forward floating search to get a feature subset and then MLP was applied to do the classification. The experimental results showed that the proposed methods exceed some previous work under the same constraints.

Luz et al. compared MLP with some other classifiers on different feature sets. Alickovic and Subasi used autoregressive modeling to extract features from the de-noised signals by multiscale PCA, and several classifiers including MLP were used to train models. Yu et al. used PNN to build classifiers on the combined features of RR- interval and the features by Wavelet and ICA respectively.

Wang and Chiang et al. pointed out that the integration of PNN with the proposed PCA and LDA feature reduction can achieve satisfactory results. Some other researchers investigated the performance of fuzzy NN and combined NN on ECG classification . Owing to the simplicity, KNN and DT were also widely applied to ECG classification. Besides the above-mentioned classifiers, some scholars also use Linear Discriminants, Extreme Leaning Machine , Optimum-path forest ,Active Learning, and so on to build classification models.

5. Compression of Wavelet transforms

Ranjit kumar P (et..al), the storage has limitation which has made ECG data compression as an important issue of research in biomedical signal processing. In addition, the transmission speed of real-time ECG signal is also enhanced and economical due to ECG signal compression. An ECG signal contains steep slopes QRS complexes and smoother P and T waves. It is recorded by applying electrodes to various locations on the body surface and connecting them to a recording apparatus. There are certain amounts of sample points in ECG signal which are redundant and replaceable. ECG data compression is achieved by elimination of such redundant data sample points. During the past few decades, many schemes for ECG signal compression have been proposed. Most of them are lossy compression techniques in which the reconstruction signal is not exact replica of the original input signal. Generally, these techniques are classified two categories: direct techniques and general techniques. In early stage of research, several methods such as the amplitude zone time epoch coding (AZTEC), and coordinate reduction time encoding system (CORTES) were developed based on direct scheme in which compression is achieved by eliminating redundancy between different ECG samples in the time domain. Some other examples are the turning-point (TP) data reduction algorithm, the scan-along polygonal approximation (SAPA) and differential pulse code modulation (DPCM).All such techniques involve simple signal processing and yield minimum distortion with good compression. A detailed review on these techniques is presented in [4-6] and the references there in the last two decades, a substantial progress has been made in the field of data compression. So far, several efficient ECG compression techniques have been reported in literature such as Linear Predictive Coding (LPC), Waveform coding and Subband coding. In these techniques, more sophisticated signal processing techniques are employed. Linear predictive coding is robust tool widely used for analyzing speech and ECG signal in various aspects such as spectral estimation, adaptive filtering and data compression. Several efficient methods have reported in literature based on linear prediction. While in sub band decomposition, spectral information is divided into a set of signals that can then be encoded by using a variety of techniques. Based on sub band decomposition, various techniques have been devised for the ECG signal compression. In the past, marked researches have made in the many transformation methods such as Discrete Cosine Transform (DCT), Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT) which are extensively used in data compression. Here, compression is achieved by transforming original signal into another domain to compact much of the signal energy into a small number of transform coefficients. In this way, many small transform coefficients can be discarded in the hope of achieving better compression. The fast Fourier transform is a discrete Fourier transform (DFT) algorithm which reduces the number of computations required and is exploited for analyzing signal in frequency domain. Discrete cosine transform gives nearly optimal performance in the typical signal having high correlations in adjacent samples. The detailed discussion on different techniques based on FFT and DCT is given in [14-17]. During the last decade, the Wavelet Transform, more particularly Discrete Wavelet Transform has emerged as powerful and robust tool for analyzing and extracting information from non-stationary signal such as speech signal and ECG signal due to the time varying nature of these signals. Non-stationary signals are characterized by numerous abrupt changes, transitory drifts, and trends. Wavelet has localization feature along with its time-frequency resolution properties which makes it suitable for analyzing non-stationary signals such as speech and electrocardiogram (ECG) signals. Recently, several other methods have been developed based on wavelet or wavelet packets for compressing ECG signal.

Conclusion: The Review of several methods and algorithms were presented. The analysis reported shows that a vast area of research and development is carried out for Data Mining of using wavelet transforms. In future investigation will be extended for other methods and so new method will propose for Data mining.

References

1. Berry, M. J., Linoff, G. S. (2000), ―Mastering Data Mining: The Art and Science of Customer Relationship Management‖. Wiley Computer Publishing, New York.
2. Chung, H. M., Gray, P. (1999), ―Special Section: Data Mining‖. *Journal of Management Information Systems*, (16:1),11-17.
3. Colin, S. (2000), ―The CRISP-DM Model: The New Blueprint for Data Mining‖, *Journal of Data Warehousing*, (5:4), Fall, 13-22.
4. Fayyad, U., Piatetsky-Shapiro, G., and Smyth, R (1996). "The KDD Process for Extracting Useful Knowledge from Volumes of Data," *Communications of the ACM*, (39:11), pp. 27-34.
5. Fayyad, U., (2001), ―The Digital Physics of Data Mining‖, *Communications of the ACM*, March, (44:3), 62-65
6. Antonino-Daviu, J. A., Riera-Guasp, M., Folch, J. R., & Palomares, M. P. M. (2006). Validation of a new method for the diagnosis of rotor bar failures via wavelet transform in industrial induction machines. IEEE Transactions on Industry Applications, 42(4).
7. Asghari Oskoei, M., & Hu, H. (2007). Myoelectric control systems—A survey. Biomedical Signal Processing and Control, 2, 275–294.
8. 1M. Goldman (Ed.), Principle of Clinical Electrocardiography, eleventh ed., Lange Medical Publication, Drawer L., Los Altos, California 94022,1982.
9. 2E. Pietka, ―Feature extraction in computerized approach to the ECG analysis*," Pattern Recog.*, vol. 24, pp. 139-146, 1991.
10. TAN Yun-fu,Du Lei, ―Study on Wavelet Transform in the Processing for ECG Signals‖, *World congress on Software Engineering, IEEE p515-518, 2009.*
11. THE WAVELET TUTORIAL by ROBI POLIKAR
    Nagendra H, S.Mukharji, Vinod Kumar, ―Application of Wavelet Techniques in ECG Signal Processing:An Overview‖, International Journal of Engineering Science and Technology (IJSET), ISSN: 0975-5462, Oct 2011.
12. Rishi Raj Sharma, et al., ―Baseline wander and power line interference removal from ECG signals using eigen value decomposition,‖ *IJBSPC*, vol. 45, pp. 33-49, August 2018.
13. N. Rashmi, et al., ―ECG Denoising using Wavelet Transform and Filters,‖ in *IEEE International Conference on WiSPNET*, 2017.
14. Savita Chandel and Kuldeep Singh, ―ECG denoising using wavelet transform and soft thresholding,‖ *IJARCSSE*, vol. 6, no. 9,Sep. 2016.
15. Changhsi Wu, Yanrong Zhang, et al., ―Implementation of ECG signal processing algorithms for removing baseline wander and electromyography interference,‖ in *8th IEEE International Conference on Communication Software and Networks*, 2016.
16. Dengyong Zhang , Shanshan Wang, Feng Li , Jin Wang , Arun Kumar Sangaiah 3,Victor S. Sheng and Xiangling Din An ECG Signal De-Noising Approach Basedon Wavelet Energy and Sub-Band Smoothing Filter MDPI applied sciences 18 November 2019.
17. Sambhu D, Umesh A. C. ―Automatic Classification of ECG Signals with Features Extracted Using Wavelet Transform and Support Vector Machines‖ International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 2, Special Issue 1,

December 2013.

18. 18 .Dean Cvetkovic , Elif Derya Übeyli , Irena Cosic –Wavelet transform feature extraction from human PPG, ECG, andEEG signal responses to ELF PEMF exposures: A pilot study‖ Elsevier Digital Signal Processing 18 (2008) 861–874
19. Hari Mohan Rai ⇑, Anurag Trivedi, Shailja Shukla ― ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier‖ Measurement 46 (2013) 3238– 3246.
20. SK. Piramu Preethika, R. Gobinath "Feature Extraction of ECG signal using Meyer Wavelet Transform‖ International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue-10S, August 2019.
21. Jayasree T, Afsin Samina, Chella Bama, Jefrin Infant Sindhya,Latha Jenifer –Cross Wavelet Transform based ECG Pattern Analysis and Classification‖ International Journal of Engineering and Management Research Page Number: 627-632, Volume-6, Issue-2, March-April 2016.
22. Taiyong Li,_ and Min Zhou ―ECG Classification UsingWavelet Packet Entropy and Random Forests‖Entropy 2016, 18, 285; doi:10.3390/e18080285.
23. K. Ranjeet, A. Kumar, and R.K. Pandey –ECG Signal Compression using optimum wavelet Filter Bank based on Kaiser Window‖ Procedia Engineering 38 (2012) 2889 – 2902.
24. Afonso VX, Tompkins WJ, Nguyen TQ, Luo S. ECG beat detection using filter banks. IEEE Trans. Biomed. Eng. 1999;46: 192–202.
25. Wang X, Meng J. A 2-D ECG compression algorithms based on wavelet transform and vector quantization. Digital Signal Processing, 2008; 18:179-188.
26. Blanco-Velasco M, Cruz-Roldan F, Lopez-Ferreras F, Bravo-Santos A, Martinez-Murioz D. A low computational complexity algorithm for ECG signal compression. Medical Engineering & Physics, 2004; 26: 553-568.
27. Addison PS. Wavelet transform and the ECG: a review. Institute of Physics Publishing-Physiological Measurement, 2005; 26:155-199.
28. Jalaleddine SMS, Hutchens CG, Strattan RD, Coberly WA. ECG data compression techniques: a unified approach.. IEEE Trans. Biomed. Eng. 1990; 37(4): 329-343.
29. Ole-Aase S, Nygaard R, Husoy JH. A comparative study of some novel ECG data compression technique. IEEE Nordic Signal Processing Symposium (NORSIG'98),1998; 3: 273-276.
30. Nave G, Cohen A. ECG compression using long-term prediction. IEEE Trans. Biomed. Eng. 1993; 40: 877–885.
31. Ramakrishna AG, Saha S. ECG coding by wavelet-based linear prediction. IEEE Trans. Biomed. Eng.1997;44(12): 1253–1261.
32. Al-Shroud A, Abo-Zahhad M, Ahmed SM. A novel compression algorithm for electrocardiogram signals based on the linear prediction of the wavelet coefficients. Digital Signal Processing, 2003; 13:604–622.
33. Aydin MC. ECG data compression by sub-band coding. IEEE electronics letters, 1991; **27**:359-360.
34. Husoy JH, Gjerde T. Computationally efficient sub-band coding for ECG signals. Medical Engineering and Physics, 1996; 18(2):132–142.
35. Zigel Y, Cohen A, Katz A. Signal Compression Using Analysis by Synthesis Coding. IEEE Trans. on Biomed. Eng. 2000; 47(10):1308-1316.
36. Blanco-Velasco M, Cruz-Roldán F, Moreno-Martínez E, Godino JI, Barner KE. Embedded filter bank-based algorithm for ECG compression. Signal Processing, 2008; 88(6):1402–1412.