VOL. 10, NO. 21, NOVEMBER 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

DEVELOPMENT OF CARDIOID BASED GRAPH ECG HEART ABNORMALITIES CLASSIFICATION TECHNIQUE

Siti Nurfarah Ain Mohd Azam, Nur Izzati Zainal and Khairul Azami Sidek
Department of Electrical and Computer Engineering, Faculty of Engineering, International Islamic University Malaysia, Jalan Gombak, Kuala Lumpur, Malaysia
E-Mail: azami@iium.edu.my

ABSTRACT

In this study, the development of Cardioid based graph electrocardiogram heart abnormalities classification technique is presented. ECG signals in this work were acquired from a public online database UCD Sleep Apnea database (UCDB) with sampling rate of 250 Hz. Each recording has 60 seconds of electrocardiogram signals. Unique features were extracted using the Pan Tompkins algorithm, later Cardioid based graph was formed as the result of the differentiation process. The various shapes of closed-loop created were then observed. From the Cardioid loop, we evaluated the area and standard deviation to differentiate between normal and abnormal heartbeats. As a result, the area and standard deviation values of abnormal heartbeat were twice the value of a normal heartbeat thus indicating the differences between two types of heart morphologies. In order to justify the results, the signal is then classified by using Bayes Network classifier. Classification outcomes suggests that the proposed technique gives heart abnormality identification with a classification accuracy of as low as 12.5% when normal and abnormal heartbeat are matched (two different conditions). Thus, the output of the study suggests the proof-of-concept of our proposed mechanisms to detect heart abnormalities and has the potential to act as an alternative to the current techniques.

Keywords: ECG signal, cardioid, bayes network, pan tompkins, heart abnormalities, sleep apnea.

1. INTRODUCTION

Sleep apnea is a common and potentially serious disorder in which many of us are not aware of. According to the National Sleep Foundation, sleep apnea affects more than 18 million Americans [1]. Sleep apnea is seen more frequently among men than among women. Left untreated, sleep apnea can have serious and life-shortening consequence such as high blood pressure, heart disease and stroke. This is caused by the reduction of oxygen in blood due to blocked airway while sleeping as shown in Figure-1. Sleep disturbances and repeated reductions in blood oxygen levels result in excessive daytime sleepiness, reduced quality of life, and impaired cognitive function such as memory loss and poor concentration.

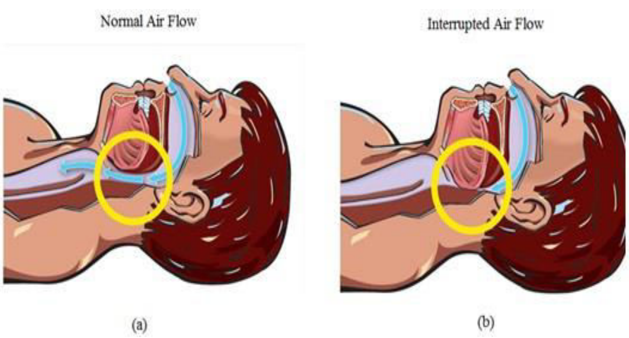

Figure-1. (a) Normal air flow while sleeping, (b) Interrupted air flow while sleeping [1].

Apnea is adapted from the Greek word which means without breath and sleep apnea refers to pauses in breathing that occurs during sleep. A major symptom of sleep apnea is extremely loud snoring. Common sleep apnea symptoms include waking up with a very sore or dry throat, occasionally waking up with a choking or gasping sensation, sleepiness or lack of energy during the day, sleepiness while driving, morning headaches, restless sleep, forgetfulness, mood changes, decreased interest in sex and recurrent awakenings or insomnia [2].

Figure-2 shows the ECG for the heart of a person who suffers from sleep apnea. The ECG waveform of a normal and healthy person is different with an ECG pattern of a person with heart disease.

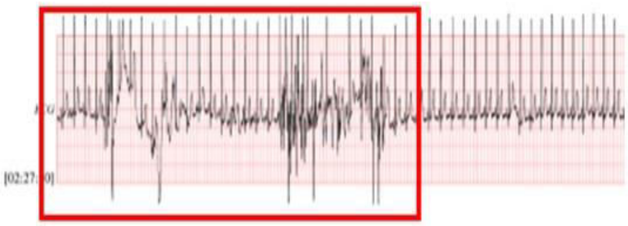

Figure-2. ECG signals for a person with sleep apnea.

As can be seen in the figure, the differences between a normal and abnormal ECG signal is visible where a normal heart gives a constant shape of ECG wave whereas an abnormal heart has different amplitude reading at certain levels. Due to the differences, this study suggests a simpler and efficient method to identify and classify heart irregularities implementing our propose feature extraction technique.

The remaining sections in this paper are structured as follows; the next section will review the related works on heart abnormalities classification technique. Later, Section III, elaborates more on the method of the study which includes the data collection procedure, pre-processing, feature extraction and the classification mechanism. After that, in Section IV, the performance of our proposed system is discussed. Last but not least, in Section V, the study is concluded based on the experimentation and results in the previous section.

2. LITERATURE REVIEW

In the past decades, there exists a variety of methods introduced to determine the heart's activity. The most widely used technique is analyzing the ECG signal. From the ECG readings, a medical practitioner is able to identify abnormalities in the heart waveform. In the study, we divide the related literature into two categories; i) non-cardioid method and ii) cardioid-method. These methods will be briefly explained sequentially.

Wang *et al.* in [3] portrays the outline and acceptance of a compelling sleep stage characterization technique for patients with sleep apnea. The method comprises of a sequential forward selection (SFS) highlight choice strategy and a decision-tree-based support vector machines (DTB-SVM) classifier for segregating three sorts of sleep in view of ECG signals. Every 5-minutes epoch of ECG signal information gathered amid sleep was utilized to create 24 highlights utilizing heart rate variability (HRV) investigation. A SFS highlight determination strategy was then utilized to figure out which noteworthy highlights ought to be chosen to enhance order precision. A DTB-SVM was then prepared utilizing chose includes as a part of request to segregate three sleep stages, including pre-sleep attentiveness, NREM sleep and REM sleep. The accuracy of the proposed technique was 73.51 %.

Besides that, Mendez *et al.* in [4] also proposes an option assessment of OSA taking into account ECG signal during sleep time which due to respiratory disturbance produces a specific pattern on ECG. Extraction of ECG characteristics, as Heart Rate Variability (HRV) and peak R region, offers option measures for a sleep apnea pre-diagnosis. 50 recordings originating from the apnea Physionet database were utilized as a part of the examination. A bivariate autoregressive model was utilized to assess beat-by-beat power spectral density of HRV and R peak zone. k-Nearest Neighbor (kNN) supervised learning classifier was utilized for classifying apnea occasions from ordinary ones, on a minute-by-minute premise for every recording. Information was part into two sets, preparing and testing set, each data with 25 recordings. The characterization results demonstrated a precision higher than 85% in both training and testing.

As explained in the previous section, sleep apnea is the occasion when one either has stops of taking in their sleep, or has low breath while asleep. This delay in breathing can go in recurrence and span. Obstructive sleep apnea (OSA) is the normal manifestation of sleep apnea, which is presently tried through polysomnography (PSG) at sleep labs. PSG is both lavish and badly arranged as a specialist human eyewitness is obliged to work over night. Almazaydeh *et al.* in [5] proposed a mechanized characterization calculation which procedures brief time ages of the ECG information. The exhibited arrangement procedure is taking into account support vector machines (SVM) and has been prepared and tried on sleep apnea recordings from subjects with and without OSA. The outcomes demonstrate that our computerized grouping framework can perceive ages of sleep issue with a high precision of 96.5% or higher.

These are among the research studies performed in the past using non-cardioid techniques. However, identifying heart abnormalities by using the technique of Cardioid based graph for ECG biometric is recognized as a faster and more convenience method as compared to ECG recordings from the Holter readings [6, 7]. There are previous works on Cardioid based graph have been proposed and will briefly discuss about these approaches.

In Sidek *et al.* [6], an analytical approach was used in order to gain the QRS complex as this portion of the ECG waves is mostly affected by cardiac irregularities. These QRS complexes were used to form the Cardioid based graph by using the differentiated ECG signals. Due to this procedure, the time series representation will be lost and being replaced by set of points. These points are important to determine the centroid (centre coordinate) and the extrema points (distance from the centroid to a certain point). By using this information, the Euclidean distances are calculated. Then MLP was applied to classify the individuals with early and severe Cardiac Autonomic Neuropathy (CAN) obtaining an average classification accuracy of early and severe CAN patients of 99.6% and 99.1%.

In another paper, Sidek *et al.* agreed that the demand for a faster diagnosis system is crucial due to CVD mortalities. Thus, related works in [7, 8] emphasized that Cardioid is one of the most effective technique to reduce the delay in monitoring CVD patients. ECG recordings were collected from online public databases. The Cardioid based graph was created using the same method in [6]. As classification technique, Bayes Network was implemented in data analysis and pattern recognition with an average classification of 98.4%.

Sidek and Khalil in [9] shared their concern of the ECG overwhelming ECG data in hospital servers from multiple locations which might affect the overall performance of the system. In [10] ECG biometric using Cardioid based graph gives an alternative approach of identification and classification for this scenario. The centroid is used as the reference point of the Cardioid and the Mahalanobis distances are then computed which differ from the previous technique. As compared to using Euclidean distance the classification rate was enhance from 97.15% to 99.8% for NSRDB and 96.5% to 99.4% for MITDB. Radial Basis Function network classifier was used to class the diseases.

A different paper written by Sufi, Khalil and Tari suggested different and faster technique of identifying cardioid by focusing using only 5 points [7] which are the centroid and four extremes (upper extreme, lower extreme, left extreme, and right extreme). The extremes points were obtained by first equating the equation of two samples and

VOL. 10, NO. 21, NOVEMBER 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

then calculating the intersecting point between the straight lines. For a normal heart, the 5 points of regular beats will result in the same value of measurement but for abnormal heart will shows the 5 points of different beats. Therefore, this thresholding approach could effectively identify abnormal beats from the normal beats [7]. The weakness of this method is that by measuring by only 5 points might not be applicable for all types of diseases.

However, based on our knowledge, the previous works were incapable of identifying heart diseases using Cardioid based graph technique. The function of this approach is to identify individual can be further expanded in the domain of classifying heart diseases due to the consistency of abnormal heart morphology and visible abnormal cardiac irregularities detection using Cardioid based graph method. Thus, in this study, we will propose of using Cardioid based graph to detect and identify heart abnormalities.

3. METHODOLOGY

Figure-3 summarizes the proposed identification system which consists of the Data Collection, Pre-processing, Feature Extraction and Classification stages. Each stage will be elaborated further in the next sub-sections.

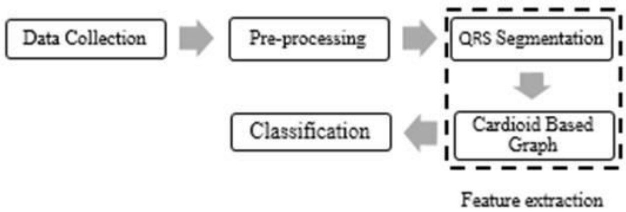

Figure-3. The proposed identification system.

a) Data collection

A total of 4 ECG signals in this work were taken from a public online database called PhysioNet. The data were taken from UCD Sleep Apnea database (UCDB) with sampling rate of 250 Hz. Each recording has 60 seconds of ECG signals. Furthermore, these subjects suffer from sleep apnea.

b) Pre-processing

The purpose of pre-processing stage is the removal of high frequency noise and disturbance such as baseline wanderings, muscle noise, etc. Besides that, it is also to enhance the accuracy of the quality of ECG signals. In this stage, the signals are filtered by using low pass filter which is Butterworth Filter. This technique is used to isolate the ECG waveform from different noises and to ensure other interfering signals are inactive. We apply Butterworth Filter with normalized cut-off frequency of 0.5 Hz to remove high frequency base line wanderings. The order of the filter taken is at 4th order.

c) Feature extraction

This stage is divided into two main steps which are QRS segmentation and Cardioid based graph.

i. QRS segmentation

The QRS waveform is segmented using Pan Tompkins algorithm. As can be observed from Figure-4, the R wave corresponds to the highest peak in the ECG signals as a result from the activity of the ventricular. The main factor QRS segment was chosen in our analysis because it is less affected by cardiac irregularities, noise and artefacts as shown in previous works as in [8].

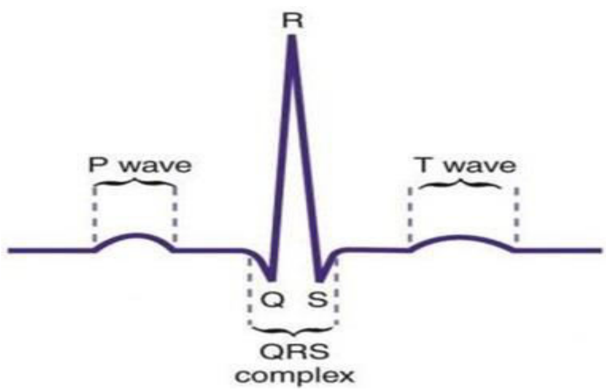

Figure-4. A Typical ECG signal [9]

ii. Cardioid based graph

After obtaining the QRS segments, the Cardioid based graph is created. This step is very important in order to ensure the accuracy of the classification by using this technique. The ECG signal can be represented by $x(t)$ as in Equation (1).

$$\mathbf{x}(t) = \{x(1), x(2), x(3), \ldots, x(N)\} \tag{1}$$

where, x(t) = ECG waveforms and,
N = the total number of QRS complexes for a given period

The QRS complexes are then differentiated as in Equation (2) in order to obtain the points to form the Cardioid.

$$y(t) = x(n) - x(n-1) \tag{2}$$

where, t = 1, 2, 3. . . (N − 1) and,

$\boldsymbol{y}(t)$ = The differentiated ECG dataset.

A closed loop graph is then generated based on a scattered XY graph after obtaining the vectors of x and y. The ECG amplitudes of the QRS signals are the x-axis and the differentiated ECG values of x are the y-axis. The time series representation is lost once the Cardioid has been generated and closed loop are formed as shown in Figure-5. and then the time series ECG signals is converted to a two dimensional loop and from the closed loop pattern, new features are extracted which are the centre coordinate of the graph called centroid and the distance of the

ARPN Journal of Engineering and Applied Sciences

centroid to a given point on the cardioid called extrema points.

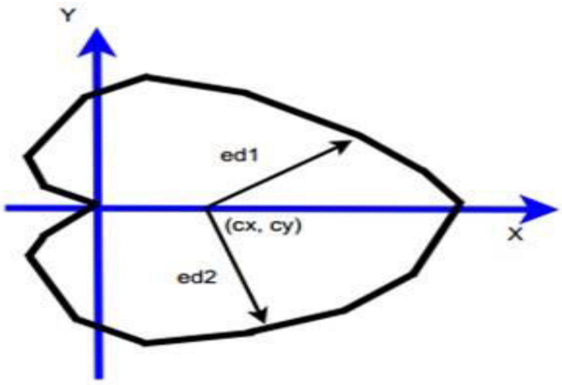

Figure-5. Cardioid based graph for normal heartbeat [6]

Centroid of the Cardioid graph is obtained by using Equation (3) and it is represented as *cx* and *cy*.

$$= \left[\frac{\sum_{i=1}^{N} x(i)}{N}, \frac{\sum_{i=1}^{N} y(i)}{N}\right] \qquad (3)$$

where cx and cy are the coordinate position of the centroid in the Cardioid graph. Using the centroid, the Euclidean distances, *ed(i)* are then computed using Equation (4).

$$ed(i) = \sqrt{(cy - y(i))^2 + (cx - x(i))^2} \qquad (4)$$

d) Classification

For classification of the proposed system, two methods are implemented which are statistical and non-statistical approaches. As for statistical method, the results are evaluated by calculating the area and standard deviation of the Cardioid graph formed. The area of the Cardioid is obtained by applying the formula of a polygon. The standard deviation is then derived after calculating the area of the Cardioid. Both calculations are executed by using MATLAB. The purpose of this approach is to validate the accuracy of the propose technique by determining the differences between normal and abnormal heartbeats.

For non-statistical method, a commonly used classification algorithm is implemented. In this study, Bayesian Network is used as the classification method to identify the class labels for instances, each typically described by a set of features for the ECG signals because of various reasons. The model encodes dependencies among all variables, it readily handles situation where some data entries are missing. Besides that, a Bayesian network can be used to learn causal relationships, and hence can be applied to gain understanding about a problem domain and to predict the consequences of intervention. In addition, the model has both a causal and probabilistic semantics, thus it is an ideal representation for combining prior and data. Last but not least, Bayesian statistical methods in conjunction with Bayesian networks offers an efficient and principled approach for avoiding the over fitting of data [9].

4. EXPERIMENTATION AND RESULTS

In this section, the experimentation and result using the proposed identification system as shown in Figure-3 is described in detail. To briefly recap, the stages involved are Data Collection, Pre-processing, Feature Extraction and Classification stages.

Figure-6 shows the raw ECG signals of subject *ucddb010_recm* sleep apnea patient as an example of the Data Collection stage.

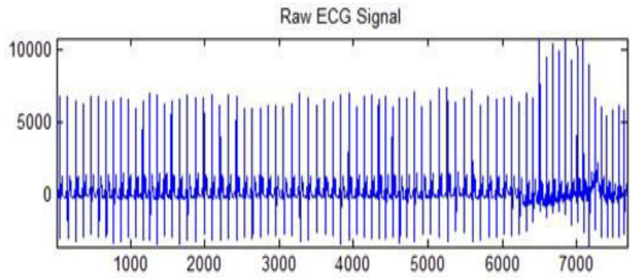

Figure-6. Raw ECG signal of subject ucddb010_recm with sleep Apnea

The ECG signal is then filtered using derivative filter as shown in Figure-7. The purpose of filtering is to remove the unwanted noise such as baseline wandering produced from the movement of the body. Besides that, it is also to enhance the accuracy of the quality of ECG signals. This step fulfils the second stage which is Pre-processing.

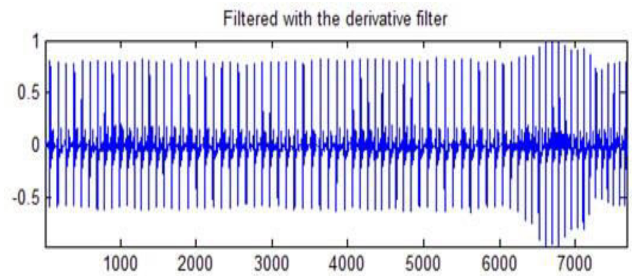

Figure-7. ECG signal after filtering using derivative filter.

Pan Tompkins algorithm was implemented in order to segment the QRS complexes after filtering process as shown in Figure-8. This is the first part of the Feature Extraction stage whereas the second part involves the formation of the Cardioid graph.

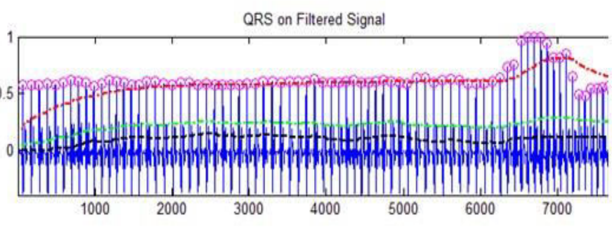

Figure-8. QRS segmentation of ECG filtered signal.

ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

After the removal of the noise and segmentation, a Cardioid graph is formed as shown in Figure-9. In order to produce the Cardioid graph, the R peak points are identified, and the coordinates from Q to S waves are taken.

As can be observed from the figure, the shape of the graph is inconsistent where some are small and some are big. This is due to the presence of normal and abnormal heartbeats. The purpose of this step is to calculate the differences between normal and abnormal heartbeat. The looping of the graph for the normal ECG is consistent in shape as can be seen from the figure. The graph that is formed outside the constant shaped loop is the abnormal heart beat which indicates the occurrence of sleep apnea.

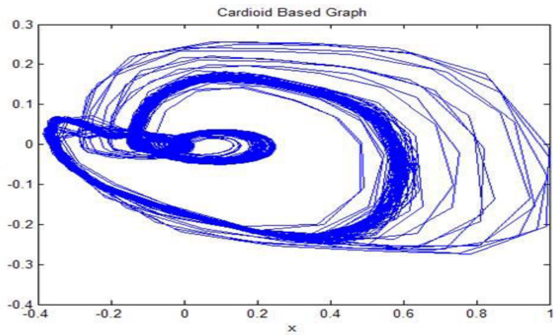

Figure-9. Cardioid based graph of sleep Apnea patient.

Two segmented QRS were taken for normal heartbeat and another two segmented QRS for abnormal heartbeat. Then, the Cardioid graph is formed to calculate its area and the standard deviation.

Figures-10, 11, 12 and 13 show the results from the segmented ECG signals of normal and abnormal heartbeat. As can be seen from the figures, the area of the abnormal ECG signals is larger than the normal signals. The value of the standard deviation also differs where the abnormal ECG signals are twice higher than the value of standard deviation of the normal ECG signals. The two values of standard deviation indicate which axis they originate where the right side is for the x-axis and the left side is the y-axis.

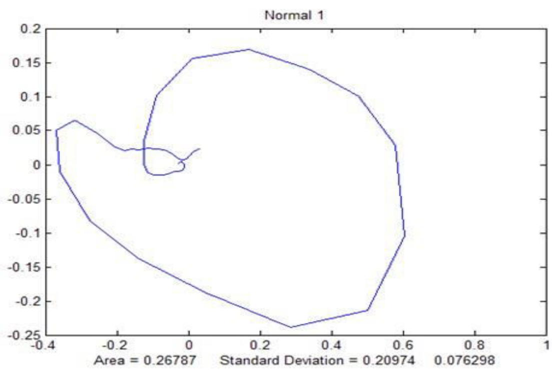

Figure-10. Result of the segmented ECG normal 1 heartbeat.

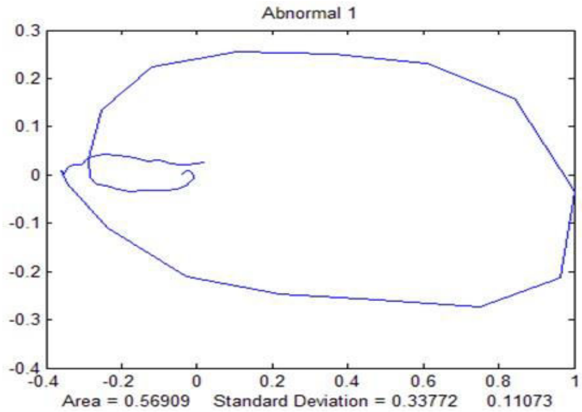

Figure-11. Result of the segmented ECG abnormal 1 heartbeat.

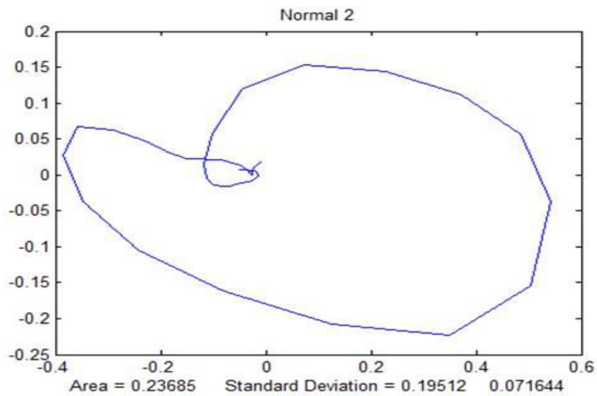

Figure-12. Result of the segmented ECG normal 2 heartbeat.

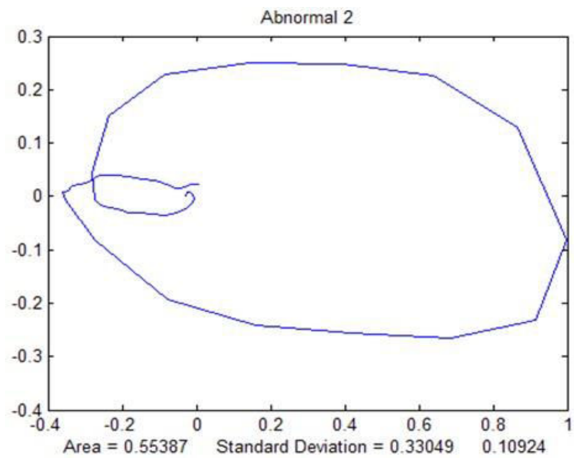

Figure-13. Result of the segmented ECG abnormal 2 heartbeat.

In order to confirm the result, the Cardioid graph was again plotted. From Figure-14, the red lines indicate the occurrence of sleep apnea and the blue line shows the normal heartbeat.

ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

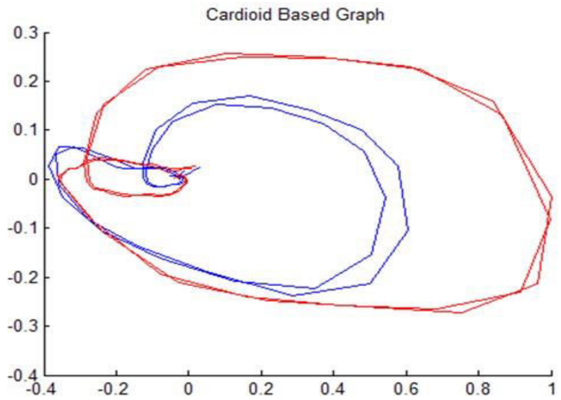

Figure-14. Cardioid based graph of segmented ECG signals.

Tables-1 and 2 reflect the results of the area and the standard deviation of Cardioid based graph for all the subject tested. As can be observed, the values of the area for normal are smaller than the abnormal heartbeat. Similarly, the same results are true for standard deviation. The value for normal heartbeat are smaller as compared to the abnormal heartbeat. While sleeping, minimal activity occur which causes the heartbeat to remains calm. However, when the sleep apnea occurs, the heartbeat amplitude experience a sudden increase which leads to bigger Cardioid graphs. Thus, the value of the area and its standard deviation will be higher.

Table-1. Area of the cardioid based graph.

Subject	Normal 1	Normal 2	Abnormal 2	Abnormal 2
ucddb002_recm	0.17	0.19	0.56	0.78
ucddb006_recm	0.06	0.07	0.14	0.16
ucddb010_recm	0.27	0.24	0.57	0.55
ucddb013_recm	0.12	0.14	0.33	0.30

Table-2. Standard deviation of the cardioid based graph.

Subject	Normal 1 (σ)	Normal 2 (σ)	Abnormal 2 (σ)	Abnormal 2 (σ)
ucddb002_recm	X= 0.20 Y= 0.06	X= 0.19 Y= 0.06	X= 0.38 Y= 0.11	X= 0.39 Y= 0.13
ucddb006_recm	X= 0.13 Y= 0.04	X= 0.13 Y= 0.04	X= 0.19 Y= 0.05	X= 0.21 Y= 0.06
ucddb010_recm	X= 0.21 Y= 0.08	X= 0.20 Y= 0.07	X= 0.34 Y= 0.11	X= 0.33 Y= 0.11
ucddb013_recm	X= 0.16 Y= 0.05	X= 0.17 Y= 0.05	X= 0.32 Y= 0.07	X= 0.30 Y= 0.08

Furthermore, Bayes Network classifier is used to evaluate the accuracy of the proposed technique when compared with normal-to-normal signals and normal-to-abnormal signals. A total of 2 QRS complexes from each subject were taken from the database where 8 instances were used. Half of the samples were used as training data and the remaining acts as the testing data.

Table-3 shows the classification accuracies between normal with normal heartbeats. The purpose is to verify the pattern of the normal graph are consistent which will be used as training set to compare with abnormal heartbeat. From Bayes Network, the accuracies for x and y axes are 100% and 87.5%. Based on these results which gives high accuracies, thus proves the reliability of the datasets.

Table-3. Normal ECG signals classification accuracy.

Parameter	Bayes Net (%)
x-axis	100.0
y-axis	87.5

Table-4 is the results of the training and testing set for normal and abnormal segments. As can be observed, the accuracy is low. This indicates that the pattern of abnormal signal is different from the normal signal. Therefore, it can be conclude that, when the abnormal heartbeat being tested with a normal heartbeat, the results of accuracy is low. Thus, by using Bayes Network, we are able to differentiate between normal and abnormal wave shape based on the classification results.

Table-4. Classification accuracy of comparison between normal and abnormal heartbeat.

Parameter	Bayes Net (%)
x-axis	12.5
y-axis	25.0

5. CONCLUSIONS

As a conclusion, in this study, we have demonstrated an efficient and accurate Cardioid based graph ECG heart abnormality classification technique for sleep apnea patients using statistical methods and a classification algorithm. In the statistical approach, the area and standard deviation values of abnormal heartbeat double the value of a normal heartbeat. All of these statistical parameters indicate the reliability of the proposed system to classify the heart abnormality for sleep apnea. Furthermore, the classification outcomes approach suggests that the proposed method gives significant heart abnormality identification with a classification accuracy of as low as 12.5%. Therefore, this output indicates that Cardioid based graph ECG heart abnormality classification technique has the ability to differentiate between normal and abnormal heartbeat as an alternative to the current techniques.

REFERENCES

[1] American Sleep Apnea Association. Sleep Apnea. (Accessed: 7th April 2015) http://www.sleepapnea.org/learn/sleep-apnea.html

[2] Blahd, W, Sleep Apnea Symptoms. (Accessed: 7th April 2015) http://www.webmd.com/sleep-disorders/sleep-apnea/symptoms-of-sleep-apnea

[3] Wang J. S., Shih G. R., Chiang W. C. 2012. "Sleep stage classification of sleep apnea patients using decision-tree-based support vector machines based on ECG parameters," Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on , pp. 285-288, 5-7 January.

[4] Mendez M.O., Ruini D.D., Villantieri O.P., Matteucci M., Penzel T., Cerutti S. and Bianchi A. M.. 2007. "Detection of Sleep Apnea from surface ECG based on features extracted by an Autoregressive Model," Engineering in Medicine and Biology Society. EMBS 2007. 29th Annual International Conference of the IEEE, pp. 6105-6108, 22-26 August.

[5] Almazaydeh L., Elleithy K. and Faezipour M. 2012. "Obstructive sleep apnea detection using SVM-based classification of ECG signal features," Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, pp. 4938-4941, 28 August – 1 September.

[6] Sidek K. A., Jelinek H.F. and Khalil I. 2011. "Identification of Cardiac Autonomic Neuropathy patients using Cardioid based graph for ECG biometric," Computing in Cardiology, pp. 517-520, 18-21 September.

[7] Sufi F. Khalil I. and Tari Z. 2010. "A Cardiod based technique to identify Cardiovascular Diseases using mobile phones and body sensors," Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, pp. 5500-5503, 31 August – 4 September.

[8] Sufi F., Khalil I. and Habib I. 2011."Cardioids-based faster authentication and diagnosis of remote cardiovascular patients". Security Comm. Networks, Vol. 4, pp. 1351–1368.

[9] Sidek K. and Khali I. 2012. "Biometric sample extraction using Mahalanobis distance in Cardioid based graph using electrocardiogram signals," Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, pp. 3396-3399, 28 August - 1 September.

[10] Sidek K.A., Jelinek H.F. and Khalil I. 2011. "Identification of Cardiac Autonomic Neuropathy patients using Cardioid based graph for ECG biometric," Computing in Cardiology, 2011, pp. 517-520, 18-21 September.