Heartbeat Classification based on Combinational Feature Selection Method for Analysing Cardiac Disorders

Hemant Kasturiwale*, Sujata N. Kale** and Sunil K. Nair***
* Associate Professor
hemant1273@rediffmail.com
** Assistant Professor
sujatankale@rediffmail.com
PG student
snair545@gmail.com

Abstract: Electrocardiogram (ECG) gives the information of electrical activity of the heart. ECG and heart rate gives the condition of cardiac health. Analysis of non linear features of ECG signal for arrhythmia characterization is considered. Time domain and frequency domain analysis is done on the ECG signal which can be useful in arrhythmia detection. The statistical parameters in time domain which have been considered are the standard deviation of the NN intervals (SDNN) and the root mean square of successive difference intervals which are taken from heart rate signals (RMSSD). The frequency domain parameters which are considered is low frequency (LF), high frequency (HF) and LF/HF ratio and the analysis of normal to normal interval (NN interval) data gives information of significant difference in very low frequency power, low frequency power and high frequency power. The indexes based on HRV prove a strong predictor of increased all-cause cardiac and/or arrhythmic mortality, particularly in patients at risk after MI or with CHF. Algorithm was implemented in such a way that it will surpass the limitations of existing algorithms. The paper also reveals the role physiological situations, various pathological settings & its role along with HRV indexes for interpretation of subject health.

Keywords: ECG, heart rate, SDNN, RMSSD, LF/HF, cardiac arrhythmia.

Introduction

In recent years, in the field of HRV analysis has found more interest in these time phase and new statistics are found from the physics and have been recommended as complimentary to traditional measures of time and frequency domain [1]. At the same instant, older algorithms are constantly being refined, and advanced methods are being tested in addition to further improve the estimation of required parameters in health and disease [2]. The electrocardiogram of a person gives the total electrical and muscular functions of a human heart and information about cardiac health [3]. The waveform with a peak is usually referred as P-wave [4]. It represents the atrial depolarization that is the requisite time for an electrical impulse generated from the sinoatrial node to proliferate all over the atrial musculature. The QRS complex represents ventricular depolarization and comprises of three waves - the Q-wave, the R-wave and the S-wave. Q wave is present at the beginning of QRS complex [5]. It is the first negative deflection. The first positive deflection is depicted by R-wave, irrespective of the fact that it is surpassed by a Q-wave or not S wave is the next negative aberration which is superseded by the R-wave [6]. Processing of cardiac signal and recognizing the cardiac disorders is challenging task in biomedical signal processing. The state of cardiac heart is generally studied in the shape of ECG waveform. Computer based ECG analysis can provide information regarding various Cardiac diseases.

Subjects and Methods

The database used in this study is acquired from physio-Bank which is available on-line. The source of the ECGs included in the Dataset is a set of over 4000 long-term Holter recordings. Around 60% of these recordings were obtained from inpatients. It consists of a variety of rare but clinically important phenomena that would not be well-depicted by a small random sample of Holter recordings. The first group behaves as a representative sample of the variety of waveforms and shows that an arrhythmia detector might be used in routine clinical use. The second groups were chosen to include complex ventricular, junctional, and supraventricular arrhythmias and conduction aberrancies. By placing the electrodes on the chest, we can get a modified limb lead II (MLII) that is the upper signal in most of the records. The lower signal is usually a modified lead V1. Nine Del Mar Avionics model 445 two-channel recorders were used for original analog recordings [7]. In order to restrict analog-to-digital converter (ADC) saturation firstly the analog outputs of the playback unit were filtered and then antialiasing is done by utilizing a pass band from 0.1 to 100 Hz relative to real time. The sampling frequency was chosen in such a way that it makes easier the implementations of 60 Hz (mains frequency) digital notch filters in arrhythmia detectors [8]. Figure 1 represents a standard routine of ECG signal processing. The digitized ECG signal is processed in MATLAB software. The Time domain and frequency domain analysis is done onto the series of successive R-R interval values.

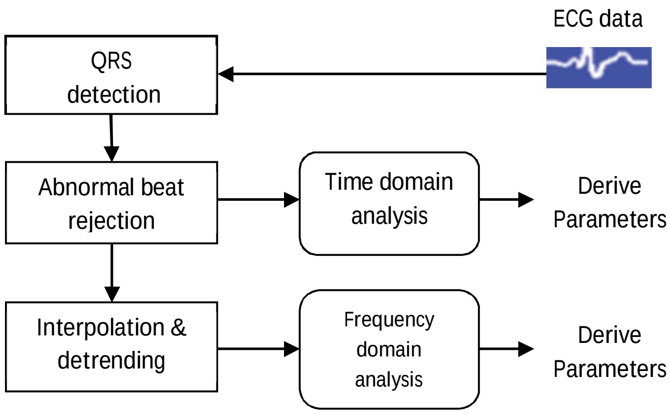

Figure 1: A standard routine of ECG signal processing

Techniques Used

Time Domain Analysis

Variation in heart rate may be assessed by a number of methods. Perhaps the simplest to perform are the time domain measures. With these techniques either the heart rate at any point in time or the intervals between successive normal complexes are determined [9]. In a continuous ECG record, each QRS complex is detected, and the so-called normal-to-normal (NN) intervals, or the instantaneous heart rate is found,. The parameters the mean and the variance of R-R interval signal plays an important role and can be utilized for the classification along with the power content in low and high frequency bands [10]. From the original R-R intervals, the standard measures parameters used in this work are:
1. The standard deviation of the NN intervals (SDNN).
SDNN is the simplest statistical HRV feature to calculate. NN stands for time interval between consecutive normal sinus heart beats. The standard deviation reflects all the cyclic elements responsible for variability in the period of recording. It can be calculated for 24 hours long-term recordings and for short-term, five minutes recordings. SDNN can be calculated for short periods and long periods as a measure of long-term variation [11].
2. The root mean square successive difference of intervals (RMSSD).
Root Mean Square of the Successive Differences (RMSSD) is one of a few time-domain tools used to assess heart rate variability, the successive differences being neighbouring RR intervals. Thus, it can be calculated using the expression:

$$RMSSD = \frac{\sqrt{\sum_{i=1}^{N-1}(x_{i+1} - x_i)^2}}{N}$$

Where x_i denotes the length of an R-R interval with index i.

Frequency Domain Analysis

Frequency domain is concerned with the analysis of mathematical functions or signals with respect to frequency. The two main frequency components are the low frequency (LF) components and the high frequency (HF) components. High frequency component (HF) i.e. spectral component around the respiratory frequency and mainly related to vagal affair, the second one a low frequency component (LF) having power variations related to sympathetic activity [9]. Frequency domain analysis was done by non-parametric based method. We have also evaluated and analyzed the LF/HF ratio.
LF/HF: HRV signal is transformed into frequency domain and the ratio of spectral power in lower bound to spectral power in upper bound is to be calculated. The lower bound frequency power is correlated to controlling temperature and cardiovascular mechanism and the upper frequency is related to the cardiac vagal activity.

Non Linear Method

There is an increasing interest to analyze HRV using methods other than the standard linear methods. These techniques are termed as nonlinear HRV analysis. The new approach Scale Invariant Structure (SIS) method is based on Detrended Fluctuation Analysis (DFA) [10]. A new approach has been introduced to characterise the variability of the RR series. This approach, known under the different names of Poincare´ plot, Lorenz plot and return map, gives a graphic representation of the behaviour of a dynamic system observed stroboscopically at given, constant time intervals. In the specific case of HRV analysis, the observation is usually performed in the bidimensional space of pairs of consecutive beats (RR*i*, RR*i*q1) [11].

Implementation

The flowchart of the implementation is as shown below in the following figure 2. The start of the process is by taking the signal of the cardiac patient. The ECG signal of the cardiac patient is extracted providing the necessary information of the signal and the cardiac disorder. Using the process, the geometrical and statistical parameters like time domain and frequency domain parameters are calculated. The parameters which are necessary to identify the patient is suffering from cardiac disorder of cardiac arrhythmia is selected. The signal is further processed through the classifier, here it is neural network which gives more better results of the parameters to confirm the person having any cardiac disorder or arrthythmia.

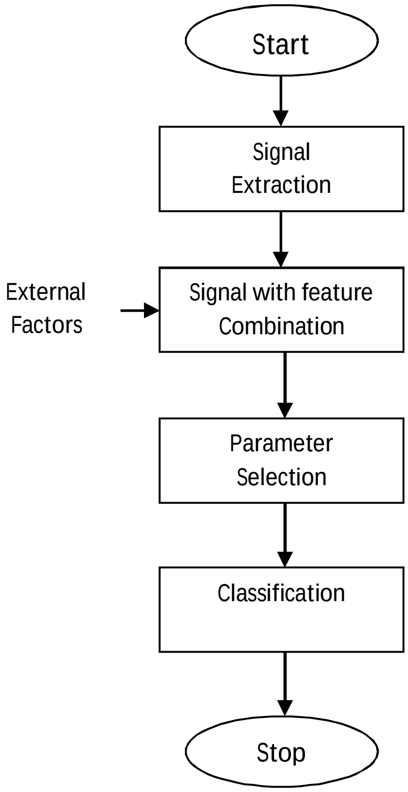

Figure 2: Flowchart of the implementation

Results

Heart rate variability shows significant dissimilarities for both time domain and frequency domain parameters between disorder group (A) and healthy group (B). Time domain and frequency domain measures of heart rate variability are depicted in table-1 and table-2 respectively. In Time domain, group (A) has reduced parameters like SDNN (ms) and RMSSD (ms), compared to the control group (B). Control group have comparatively higher values than disorder group i.e. group A.

Table 2 summarizes the results of frequency domain parameters. In frequency domain analysis in the disorder group, HF power, indicating parasympathetic activity and LF power, specifying mainly sympathetic activity, both were decreased. LF/HF ratio is calculated as shown in the table. The value of LF/HF ratio is significantly higher in group-A than group-B, which indicates dominance of sympathetic activity. Depressed HRV after MI may reflect a decrease in vagal activity directed to the heart. HRV in patients surviving an acute MI reveal a reduction in total and in the individual power of spectral components.

Table 1: Comparison between Time domain parameters of both groups

Time Domain parameters	(A) Disorder Group	(B) Healthy control Group
SDNN (ms)	34.39 ± 15.85	45.78 ± 9.96
RMSSD (ms)	27.19 ± 17.70	36.18 ± 11.69

Table 2: Comparison between frequency domain parameter of both groups

Frequency Domain parameters	(A) Disorder Group	(B) Healthy controls Group
LF power (ms^2/Hz)	612.57 ± 652.47	1080.72 ± 645.50
HF power (ms^2/Hz)	484.53 ± 497.20	889.57 ± 600.29
LF/HF Ratio	10.81 ± 4.23	1.39 ± 0.78

The database of male and female is taken into consideration with age group between 40 to 90. The database which is been taken are 6 male subject and 6 female subject and calculated the time domain and frequency domain parameters of all these subjects as per the process. We have taken the average values of all these subjects and separated the male and female values. The calculated parameters are SDNN, RMSSD and LF/HF as shown in the figure.

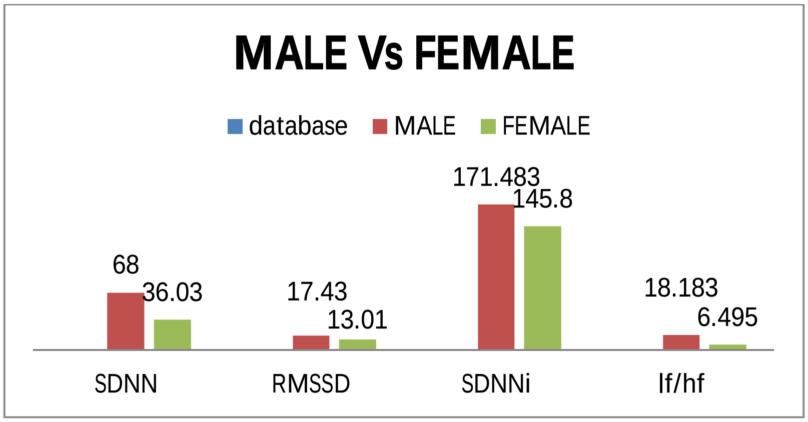

Figure 3: Comparison of MALE Vs FEMALE

In the figure above, it can be seen that the values of time domain parameters like SDNN and RMSSD is comparatively higher in male then female subjects. Also the ratio of LF/HF is higher in male then compared to female which indicates the dominance of sympathetic activity more in male. Even if heart rate value increases, LF/HF ratio was largely unaffected by either acute myocar-dial ischemia, exercise Thus, the LF component of HRV does provide an index of cardiac sympathetic drive but also reflects a complex sympathetic, parasympathetic, and other unidentified factors accounting for the largest portion of the variability in this frequency range.

In the next section, the bifurcation of age wise comparison is shown. In the first group age below 60 are taken and in the second group age above 60 are taken which are the senior citizens. The total subject taken is 10, out of which 5 are below 60 and rest 5 are above 60 i.e. senior citizens.

The figure above shows that the values of SDNN and RMSSD i.e. time domain parameters is low in senior citizens as compared to others. The value of RMSSD is almost alike with a very less variation as seen in the figure. The LF/HF ratio is also calculated. The ratio value is higher in patients with age less than 60 as compared to senior citizens. The value is also identical with a very less difference as seen from the figure; still the value is higher in patients with age less than 60, indicating the supremacy of sympathetic activity more. The presence of an alteration in neural control is also reflected in a blunting of day-night variations of RR interval.

The process of classification is shown and as a universal estimator in figure 5. The panel gives information related to Performance, Training State and Regression. The training function used here is Levenberg - Marquardt function and the Performance analysis function chosen here is the Mean Square Error (MSE) as shown in Table 3. The table below shows the values of absolute mean square error, relative mean square error and root mean square values of the signal in male and female processed through neural networks. The database is taken of 10 subjects i.e. 5 males and 5 females and the average values of these five subject is taken as shown. Every result goes fine without considering variation in respiratory rate. It does have

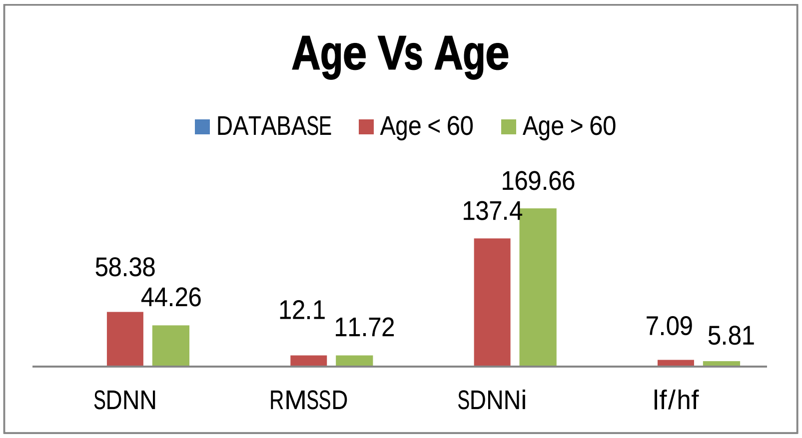

Figure 4: Age-wise Comparison

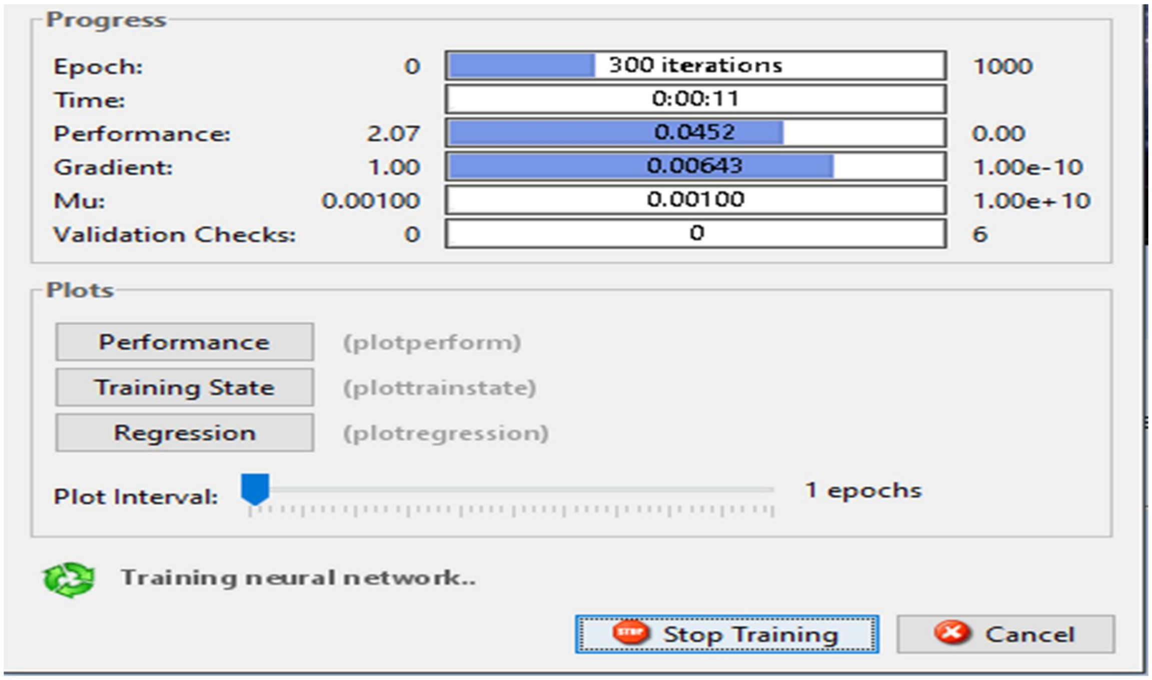

Figure 5: Training process as part of classification

variations which will introduce possible non-linear effects of varying cardiac sympathetic and cardiac parasympathetic nerve activity on LF/HF, MSE parameters.

Table 3: Comparison of various parameters in males and females after features selection and classification

Parameter	Male	Female
Absolute MSE	6.20E-02	4.60E-02
Relative MSE	3.40E-02	2.60E-02
RMSSD	7.18E-02	1.18E-01

The values of absolute mean square error and relative mean square error are almost identical in male and female. It gives indications that method with even predefined indexes worked well giving variations as per database which will help researcher to choose long HRV or short HRV database.

Conclusion

In the present study, we have evaluated HRV in Indian patients with cardiac disorder and tried to correlate the HRV with disease activity and other clinical parameters. Time domain and frequency domain analysis of the RR interval variability of cardiac disorder and normal subjects shows that there is significant difference in these measures for disorder patients with respect to normal subjects. HRV parameter like Low Frequency (LF) components was higher in males, which reflect sympathetic dominance in males. High Frequency (HF) component of HRV was significantly higher in females, which indicates parasympathetic dominance in females. Our study gives good results of the parameters irrespective of the age and gender of the patient having cardiac disorder or arrhythmia. So we can expect that these measures allow early detection and treatment/subsequent management of patients and thus can avoid complications. The study also reveals individual parameters or indexes will extract information without physiological interventions which is to be investigated further for more accurate interpretations.

References

[1] P. Mansier, J. Clairambault, N. Charlotte, C. Medigue, C. Vermeiran, G. LePape, F. Carre, A. Gounaropoulou, B. Swynghehauw, *Linear and nonlinear analyses of heart rate variability: a minireview, Cardiovascular Research 31*, 371, 1996.

[2] A. Camm, M. Malik, J. Bigger Jr., Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Task force of the European Society of Cardiology and the North American Society of Pacing and Electrophsiology, Circulation 93* pages-1043-1065, 1993.

[3] H.P.Kasturiwale, C.N. Deshmukh, “Quality assessment of ICA algorithms for ECG signal analysis” *International Conference on Emerging Trends in Engineering & Technology (ICETET),*page no. 73 – 75 , 16-18 Dec. 2009.

[4] N. Jafarnia-Dabanlooa, D. C. McLernona, H. Zhangb, A. Ayatollahic, V. Johari-Majd “*A modified Zeeman model for producing HRV signals and its application to ECG signal generation”, Journal of Theoretical Biology 244, 2007*, pp 180–189.

[5] V. Pichot, J. M. Gaspoz, S. Molliex, A. Antoniadis, T. Busso, F. Roche F. Costes, L. Quintin, J. R. Lacor, J. Barthelemy, “*Wavelet transform to quantify heart rate variability and to assess its instantaneous changes” J. Appl. Physiol. 86*, pp 1081–1091, 1999.

[6] N. Y. Belova, S. V. Mihaylov, G. Piryova, “*Wavelet transform: A better approach for the evaluation of instantaneous changes in heart rate variability”, Autonomic Neuroscience: Basic and Clinical 131*, pp 107-122, 2007.

[7] E. Toledo, O. Gurevitz, H. Hod, M. Eldar, S; Akselro, “*Wavelet analysis of instantaneous heart rate: a study of autonomic control during thrombolysis”. Am J Physiol Regulatory Integrative Comp Physiol 284, 2003*, pp 1079-1091.

[8] G. Kheder, R. Taleb, A. Kachouri, M. BenMassoued, M. Samet *“Feature extraction by wavelet transforms to analyze the heart rate variability during two meditation technique”*, Chapter Book, Springer-Verlag2008

[9] N. Y. Belova, S. V. Mihaylov, G. Piryova, *“Wavelet transform: A better approach for the evaluation of instantaneous changes in heart rate variability”, Autonomic Neuroscience: Basic and Clinical 131*, pp 107-122, 2007.

[10] H.P.Kasturiwale, Sujata N. Kale , “*Performance of detrended fluctuation analysis on biomedical time series”, International Conference on Electronics and Communication Systems(ICECS)*, 25-26 ,2016, in press.

[11] P. Stoica, R. Moses, *Introduction to Spectral Analysis*, Prentice Hall, Upper Saddle River, NJ, 1997.