*Biomedical & Pharmacology Journal*, December 2018. Vol. **11**(4), p. 1841-1849

# Comparative Analysis of Heart Rate Variability Parameters for Arrhythmia and Atrial Fibrillation using ANOVA

**Kirti, Harsh Sohal and Shruti Jain**

Electronics and Communication Engineering Department,
JaypeeUniversity Information Technology, Solani, H.P, India.
*Coressponding author E-mail: kirtitripathifzk@gmail.com

http://dx.doi.org/10.13005/bpj/1556

(Received: 13 September 2018; accepted: 06 December 2018)

**Heart Rate Variability (HRV) is an important criterion to check the cardiac health. Sudden HRV signifies the unhealthy condition of the heart, particularly when the person is suffering from a cardiac disease. HRV parameters on different patients of different ages, gender and health conditions are observed using time domain, geometrical domain and frequency domain. Statistical comparison is done on three different databases MIT/BIH Normal Sinus Rhythm (NSR), MIT/BIH Arrhythmia (AR) and MIT/BIH Atrial Fibrillation (AF) using Analysis of Variance (ANOVA) technique. Features of all the three domains show weak, moderate or strong significant changes as per the relation during comparison with respective databases. The best features were selected among the various HRV features which will be used for classification in future.**

**Keywords:** Heart Rate, Heart Rate Variability, Finite Impulse Response, Analysis of Variance.

Electrocardiography (ECG) is a non-invasive method to extract electrical signals produced by Sino-Atrial (SA) node in the heart. By monitoring the ECG of a patient, prediction can be made which helps the doctor to take effective steps clinically[1,2]. A significant decrease in the mortality rate of sudden heart attack cases can be achieved by regularly monitoring the ECG. ECG signal consists of several features delineating from *P*, *QRS* and *T* wave. Among all these features Heart Rate (HR) is an important feature to diagnose the cardiac health[2]. HR is calculated using the *RR* interval (the difference between two consecutive *R*-peaks). RR interval is the most prominent feature of the HRV, while other features are directly or indirectly dependent on it. HRV components can be calculated in the time domain, geometrical domain and frequency domain[3].

Various platforms are available nowadays for the collection of the database (offline and online)[4]. Firstly, pre-processing is done on the database to eliminate the undesired signal from the useful signal[5]. Generally, two types of unwanted signals are present which are Power Line Interference (PLI) and Baseline Wander (BLW). To remove these noises, Finite Impulse Response (FIR) or Infinite Impulse Response (IIR) digital filters are used. IIR Butterworth filter is used to diminish these noises[6]. The same filter of 4th order is used for pre-processing[7].

Butterworth filter is an IIR filter and it is not hardware implementable due to its unstable nature. FIR filter to is used to remove these unexpected signals due to ease of its hardware implementable nature[8]. Secondly, feature extraction is doneto extract the parameter of HRV for different duration

and domains[2,9]. There are various statistical tests which are used for the comparison in between and within groups like *t*-test, *f*-test, ANOVA etc. If the comparison is in between two groups then the *t*-test or *f*-test is used, while if more than two groups are there then ANOVA test is preferred. If we use the *t*-test or *f*-test on more than two groups, Type-I error will be introduced, which is unacceptable during statistical test performance. ANOVA test is a statistical test used for comparison or finding a relationship between or within features. Extraction of time and frequency domain features of HRV for the stress level evaluation is also used to executethe ANOVA test for the statistical comparison among different level of stress[10]. ANOVA test on HRV features for the evaluation between meditative state and non-meditative state of a person is used [11].

In this paper, we have used three online databases of MIT/BIH Normal Sinus Rhythm (NSR), MIT/BIH Arrhythmia (AR) and MIT/BIH Atrial Fibrillation (AF). From these databases of ECGs, removal of artifact is done by using digital FIR Kaiser Window filter because it has better frequency response than other conventional windows. Later from the filtered ECG signal, we have extracted time domain, geometrical domain and frequency domain HRV features. In the end, comparison of different databases on the basis of best features is done by using ANOVA test.

The paper is structured as: Section 2 describes the different processing steps to be applied on different databases like MIT/BIH NSR, MIT/BIH AR and MIT/BIH, Section 3 discusses the results of the different processing steps and conclusion are presented in Section 4.

## METHODOLOGY

The process for the statistical analysis of feature processing consists primarily of pre-processing, feature extraction, feature selection and classification phase[12, 13]. In this paper, we are concentrating on the first three steps[14, 15]. Figure 1 illustrates the block diagram of HRV feature processing.

**Database Collection**

The online platform of Physionet is used for database collection. These databases are MIT/BIH Normal Sinus Rhythm (NSR), MIT/BIH Arrhythmia (AR) and MIT/BIH Atrial Fibrillation (AF) with 18 records each. The sampling frequency of NSR, AR and AF database are 128 Hz, 360 Hz, and 250 Hz respectively. The database contains measurable data of the interval scale so, all the statistical parameters can be applied on it.

**ECG Pre-processing**

Several noises in ECG signals are introduced during the acquisition of the signal from the patient body. Among them, Power Line Interference (PLI) and Base Line Wander (BLW) are of main concern[16]. These two unwanted signals create hindrance while extracting the useful information from the ECG wave. Pre-processing is done using a digital filter to diminish these undesirable signals[17]. In this paper, FIR digital filters are used due to its simplicity in hardware implementation. Among various FIR filters, FIR Kaiser Window is used because it gives better frequency response as compared to other windowing technique. Useful bandwidth of ECG signal lies in between 0.5 Hz to 50 Hz. PLI is a high-frequency noise introduced due to the biomedical instrument used for the acquisition of ECG signal. PLI is eliminated using FIR Kaiser Low Pass Filter (LPF) having cut off frequency ('$f_c$') of 50 Hz. Respiration during signal acquisition is the main cause of BLW containing low-frequency noise. Removal of BLW can be done be using FIR Kaiser High Pass Filter (HPF), having '$f_c$' of 0.5 Hz. Later, we have obtained a denoised/filtered signal which is used for HRV parameter extraction.

***R*-peak detection**

ECG signal consists of mainly three waves *P*, *QRS* and *T*. Among these *QRS* wave is the prime focus, which can be further divided into three sub-waves *Q*, *R* and *S*. *R* wave is the most prominent feature of *QRS* due to its high positive peak. In this paper, the Pan-Tompkins Algorithm for *QRS* and *R* peak detection is used[18]. For *QRS* detection, squaring and moving window integration is used. Since *R* wave is a positive peak, so squaringis done to eliminate all the negative peaks. Adaptive Thresholding is used for extraction of the *R*-peak from *QRS* wave after eliminating abnormal beats.

**HRV Feature Extraction**

Feature extraction from the *R*-peak of a filtered signal is the essential part for HRV feature processing[17]. HRV features are divided into three domains :time domain, geometric domain and frequency domain. Spectrum for *RR* interval in time

and frequency domain is demonstrated in Figure 2. In Table 1 explanation of different domain features with their description and formulas for calculating them is explained.

**HRV Feature Selection**

A number of HRV features are present in various domains as defined but only some of them shows a noteworthy effect. That's why correlation among different HRV features of various domains expressed at two significance levels (95% and 99%) is shown in this paper. HRV parameters which show high correlation factor get selected while others are discarded. For the comparison of three groups of ECG databases, ANOVA test is used in different domains. ANOVA test demonstrates the variation among the mean of HRV parameters for different databases.

## RESULT AND DISCUSSION

The input signal of ECG from different databases is used in this work which is demonstrated in figure 3. Fig. 3a represents the NSR ECG wave, Fig. 3b shows the AR ECG wave and Fig. 3c illustrates the AF ECG wave.

In Pre-processing stage FIR Kaiser Window is used to denoise the signal from unwanted noises like PLI and BLW using LPF and HPF respectively. Fig. 4a shows the removal of PLI using LPF and Fig. 4b represents the removal of BLW using HPF.

After getting filtered output, the Pan-Tompkins algorithm is used for real-time *QRS* detection. This algorithm is used for extracting the *R*-peak of *QRS* wave. *R*-peak is the most important

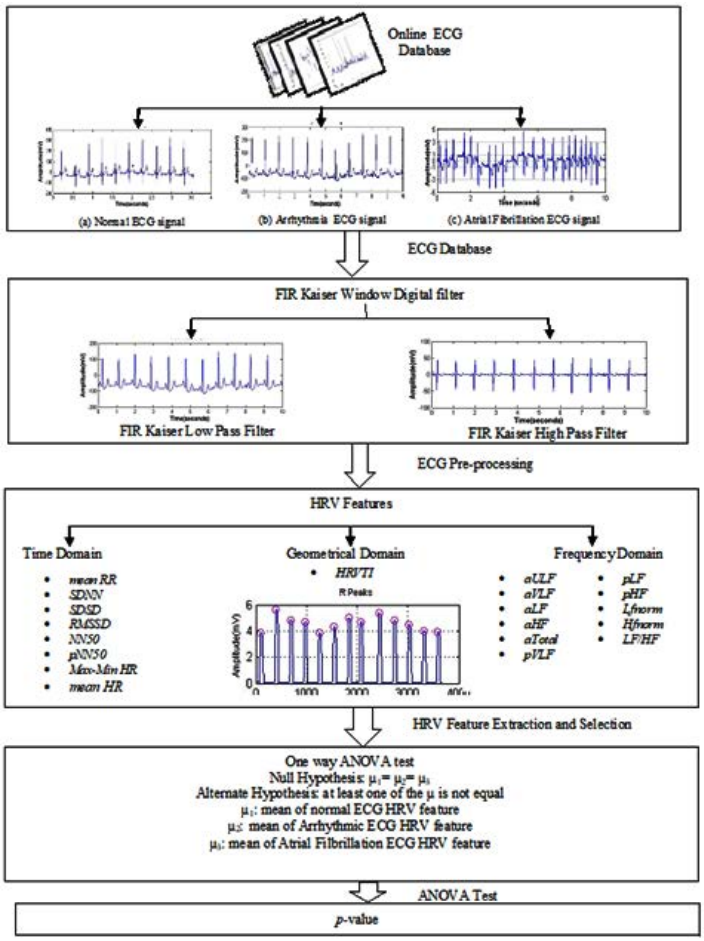


**Fig. 1.** Block diagram of methodology of HRV Feature processing

feature to calculate various HRV parameters. Figure 5 demonstrates the *R*-peak from different *QRS* waves

Detection of *R*-peak helps in calculating various HRV features for different domains as described in Table 1. Time domain HRV feature describes the temporal variability of heart rate

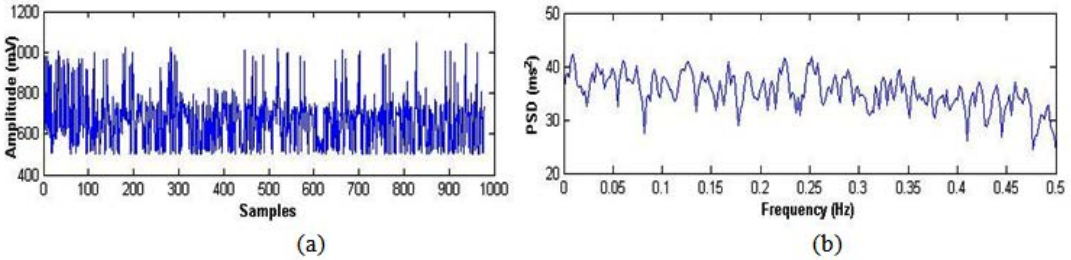


**Fig. 2.** Spectrum of *RR* interval (*a*)Time Domain(*b*) Frequency Domain

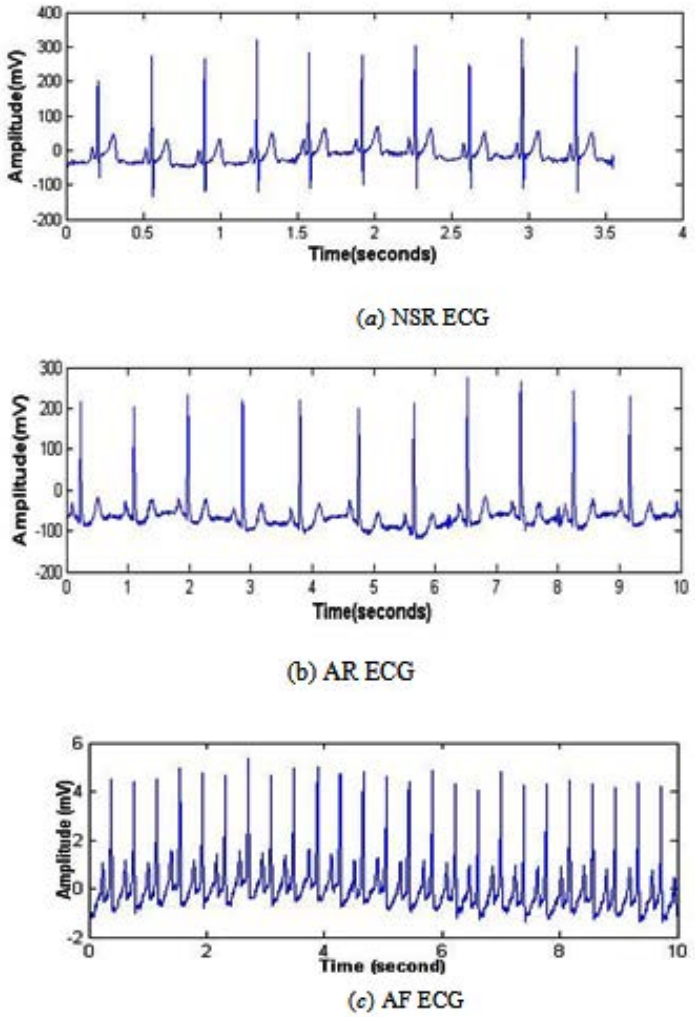


**Fig. 3.** ECG signal of (*a*) Normal Sinus Rhythm ECG (*b*) Arrhythmia ECG (*c*) Atrial Fibrillation

**Table 1.** HRV measures of time, geometrical and frequency domain2

| Domain | Features | Description | Formula | Unit |
|---|---|---|---|---|
| Time domain | mean RR | Mean of all RR interval | $mean\ RR=1/N\sum_{i=1}^{N}(RR_i)$ | ms |
| | SDNN | Standard deviation of NN interval | $SDNN=\sqrt{\frac{\sum_{i=1}^{N}(RR_i-mean(RR))^2}{N}}$ | ms |
| | SDSD | Standard deviation of SD | $SDSD=\sqrt{\frac{\sum_{i=1}^{N}(SD_i-mean(SD))^2}{N}}$ | ms |
| | RMSSD | Root Mean Square of SD | $RMSSD=\sqrt{\frac{\sum_{i=1}^{N-1}(RR_i-mean(RR))^2}{N-1}}$ | ms |
| | NN50 | Number of SD >50ms | NN50= number of SD > 50ms | nu |
| | pNN50 | Percentage of NN50 >50ms | NN50*100 | % |
| | Max-Min HR | Mean difference between maximum and minimum HR | Max-Min HR=Max(HR)-Min(HR | bpm |
| | mean HR | Mean of total HR | $Mean\ HR=1/N\sum_{i=1}^{N}HR_i$ | bpm |
| Geometrical domain | HRVTI | Ratio of integral of density distribution of RR interval to the maximum intensity distribution. | HRVTI=((Total number(NN))/max(NN) ) | nu |
| Frequency domain | aULF | Absolute power of ULF (? 0.003 Hz) | $aULF=\sum_{0}^{0.003}p(k)$ | ms2 |
| | aVLF | Absolute power of VLF (0.033-0.04 Hz) | $aVLF=\sum_{0.033}^{0.04}p(k)$ | ms2 |
| | aLF | Absolute power of LF (0.04-0.15 Hz) | $aLF=\sum_{0.04}^{0.15}p(k)$ | ms2 |

| | | | |
|---|---|---|---|
| aHF | Absolute power of HF (0.15-0.4 Hz) | $aHF=\sum_{0.15}^{0.4} p(k)$ | ms2 |
| aTotal | Absolute power of total frequency spectrum | $aTotal=\sum_{0}^{0.4} p(k)$ | ms2 |
| pVLF | Relative power of VLF | pVLF=(aVLF/aTotal)×100 | % |
| pLF | Relative power of LF | pLF=(aLF/aTotal)×100 | % |
| pHF | Relative power of HF | pHF=(aHF/aTotal)×100 | % |
| Lfnorm | Normalized change of Total power on LF | Lfnorm= (aLF/(aTotal-aVLF))×100 | nu |
| Hfnorm | Normalized change of Total power on HF | Hfnorm= (aHF/(aTotal-aVLF))×100 | nu |
| LF/HF | Ratio of LF power to HF power | LF/HF=(aLF/aHF)×100 | % |

NN: Successive Difference between normal to normal R to R peak (exclude abnormal and ectopic beats)
N: Total number of RR or NN intervals
p(k): Power Spectral Density of frequency spectrum
SD: Successive Difference between consecutive RR intervals.
bpm: beats per minute
nu: No units
ULF: Ultra Low Frequency
LF: Low Frequency
VLF: Very Low Frequency
HF: High Frequency

whereas the frequency domain of HRV feature defines the frequency distribution in a different frequency range. The geometric distribution gives the basic measurement to calculate the width of the *RR* interval histogram. Before doing any statistical test for comparing the data must be normalized( the '*p*' value should be greater than 0.05). After a normality check, ANOVA test for comparison is used.

In the ANOVA test, hypothesis testing is done on HRV features for different databases as mentioned earlier. In this paper, we have considered two hypothesis, null hypothesis ($H_o$) i.e. no variation or alternative hypothesis ($H_1$) i.e. at least one of the mean is unequal(Ellis, Zhu, Koenig, Thayer, & Wang, 2015). ForANOVAaverage case
$H_o$:$\mu_1$=$\mu_2$=$\mu_3$
$H_1$=at least one of the mean is unequal
where, $\mu_1$ = Mean of NSR HRV feature
$\mu_2$ = Mean of AR HRV feature
$\mu_3$ = Mean of AF HRV feature

Table 2 depicts the comparative results of the ANOVA test, which represents the significant variation among parameters of different datasets. The table illustrates the value in Mean ± Standard Deviation (SD) form. On the basis *p*-value significance level can be illustrated as

*p*> 0.05 = Weak Significance (WS)
0.05<*p* d" 0.10 = Moderately Significance (MS)
*p* d" 0.05 = Strongly Significance (SS)

The feature values can be affected by the sampling frequency and waveform duration of the database taken while processing. In this paper, we have generated the results at 1000 Hz sampling frequency for a long-term database. After observing the results of ANOVA test for these specifications, it is interpreted that some HRV features like *SDSD, RMSSD, NN*50*, pNN*50*, aULF, aVLF, aHF, and aTotal* gives a strong indication to reject $H_o$. It represents that these features show a significant variation in their mean value. While, all other parameters show weak evidence against to reject $H_o$, except mean HR which represents moderate significant change. Mostly HRV feature represents a strong significant change which denotes that there is a substantial difference occurring among different databases. After analyzing the *p*-value of ANOVA which signifies the noticeable variation among different databases, we have selected the best HRV features. Best selected features are *SDSD, RMSSD, NN*50*, pNN*50*, mean HR aULF, aVLF, aHF* and aTotal which we will be used for further classification in the future.

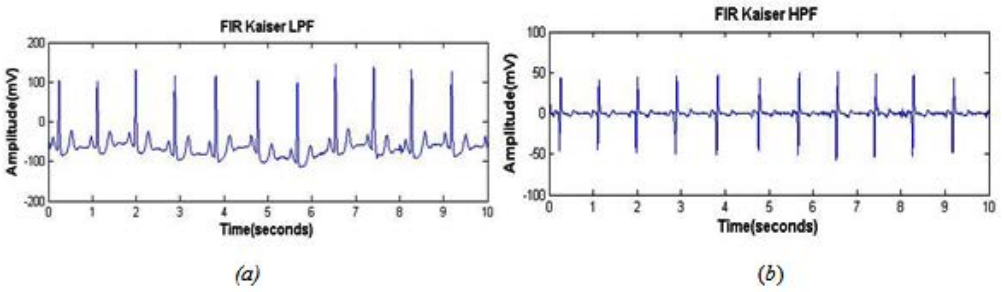


**Fig. 4.** Filtered ECG signal using FIR Kaiser Window(a) LPF for PLI removal (b) HPF for BLW removal

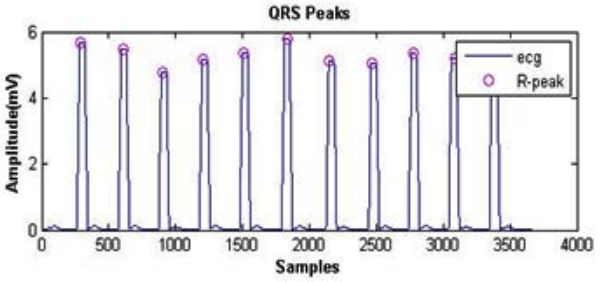


**Fig. 5.** R-peak detection

**Table 2.** Comparison of HRV features using ANOVA

| Domain | HRV Features | MIT/BIH NSR Database | MIT/BIH AR Database | MIT/BIH AF Database | p-value |
|---|---|---|---|---|---|
| | mean RR | 6.63E2±6.72 | 6.71E2±14.40 | 6.77E2±25.96 | p >0.05 |
| Time | SDNN | 1.44E2±3.89 | 1.42E2±17.20 | 1.46E2±9.17 | p >0.05 |
| Domain | SDSD | 1.98E2±5.53 | 1.84E2±20.76 | 1.97E2±13.47 | p <0.05 |
| | RMSSD | 1.98E2±5.64 | 1.84E2±20.75 | 1.97E2±13.47 | p <0.05 |
| | NN50 | 7.88E2±19.23 | 5.37E2±28.54 | 8.09E2±50.45 | p<0.05 |
| | pNN50 | 52.28±0.89 | 55.51±2.65 | 54.80±3.12 | p <0.05 |
| | Max-Min HR | 61.11±0.76 | 61.78±1.00 | 61.28±1.18 | p >0.05 |
| | mean HR | 97.05±1.83 | 93.78±4.05 | 94.89±5.32 | 0.01< p ≤0.05 |
| Geometrical Domain | HRVTI | 22.24±3.63 | 26.92±5.23 | 21.49±6.50 | p <0.01 |
| Frequency | aULF | 0.07±0.008 | 0.073±0.007 | 0.76±0.008 | p <0.05 |
| Domain | aVLF | 1.34±0.03 | 1.33±0.058 | 1.37±0.041 | p <0.05 |
| | aLF | 4.08±0.60 | 4.01±0.11 | 3.95±0.66 | p >0.05 |
| | aHF | 9.34±0.082 | 9.13±0.25 | 8.89±1.90 | p <0.05 |
| | aTotal | 14.76±0.14 | 14.47±0.37 | 14.81±0.29 | p <0.05 |
| | pVLF | 9.07±0.18 | 9.17±0.28 | 9.27±0.30 | p >0.05 |
| | pLF | 27.64±0.27 | 27.72±0.47 | 27.72±0.36 | p >0.05 |
| | pHF | 63.28±0.21 | 63.04±0.58 | 63.06±0.57 | p >0.05 |
| | Lfnorm | 0.30±0.002 | 0.30±0.005 | 0.30±0.004 | p >0.05 |
| | Hfnorm | 0.70±0.002 | 0.69±0.005 | 0.69±0.004 | p >0.05 |
| | LF/HF | 0.44±0.005 | 0.44±0.01 | 0.44±0.009 | p >0.05 |

## CONCLUSION

This paper presents the statistical comparison among HRV features of different domain and different databases. FIR Kaiser digital filter is used to eliminate noise from NSR, AR and AF database. HRV feature extraction is implemented in time, geometrical and frequency domain, further for feature selection correlation is used. For the comparison between the three databases, ANOVA test has been used which yields the best features which will be used for classificationin future. The hardware implementation of these statistical tests will be done on FPGA using VIVADO tool as extension of this work in future.

## REFERENCES


1. N. Parashar, S. Jain, M. Sood, and J. Dogra, "Review of biomedical system for high performance applications,"*4th International Conference on*in *Signal Processing, Computing and Control (ISPCC)*, 300-304 (2017).
2. F. Shaffer and J. Ginsberg, "An overview of heart rate variability metrics and norms," *Frontiers in public health,* **5**: 258 (2017).
3. A. Dhiman, S. Dubey, S. Jain, "Design of Lead II ECG Waveform and Classification Performance for Morphological features using Different Classifiers on Lead II," *Research Journal of Pharmaceutical, Biological and Chemical Sciences (RJPBCS),***7**: 1226- 1231(2016).
4. A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark*, et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation,* **101**: e215-e220(2000).
5. U. R. Acharya, H. Fujita, V. K. Sudarshan, S. L. Oh, M. Adam, J. H. Tan*, et al.*, "Automated characterization of coronary artery disease, myocardial infarction, and congestive heart failure using contourlet and shearlet transforms of electrocardiogram signal," *Knowledge-Based Systems,***132**: 156-166, (2017).
6. R. A. Tanya Kambo, "De-noising and statistical feature extraction of the ecg signal using wavelet analysis," *International Journal Of Electrical, Electronics And Data Communication,* **4**:(2016).
7. R. J. Ellis, B. Zhu, J. Koenig, J. F. Thayer, and Y. Wang, "A careful look at ECG sampling frequency and R-peak interpolation on short-term measures of heart rate variability," *Physiological measurement,* **36**: 1827 (2015).

8. A. M. A. Hassan, "FPGA Realization for Baseline Wander Noise Cancellation of ECG Signals using Wavelet Transform," *International Journal of Computer Applications,* **168**: (2017).
9. M. Urooj, K. Pillai, M. Tandon, S. Venkateshan, and N. Saha, "Reference ranges for time domain parameters of heart rate variability in indian population and validation in hypertensive subjects and smokers," *Int J Pharm Pharm Sci,* **3**: 36-39(2011).
10. Z. Qin, M. Li, L. Huang, and Y. Zhao, "Stress level evaluation using BP Neural network based on time-frequency analysis of HRV," *IEEE International Conference on Mechatronics and Automation (ICMA)*, 1798-1803(2017) .
11. S. Agarwal and A. Wadhwani, "Analysis of Heart Rate Variability During Meditative and Non-Meditative State Using Analysis of Variance," *International Journal of Advanced Biological and Biomedical Research,* **1**: 728-736(2013).
12. S. Bhusri, S. Jain, and J. Virmani, "Classification of breast lesions using the difference of statistical features," *Research Journal of Pharmaceutical Biological and Chemical Sciences,* **7**: 1365-1372, (2016).
13. S. Jain, S. Rana , J. Virmani "SVM-Based Characterization of Focal Kidney Lesions from B-Mode Ultrasound Images," *Research Journal of Pharmaceutical, Biological and Chemical Sciences (RJPBCS),* **7**: 837- 846(2016).
14. U. R. Acharya, H. Fujita, M. Adam, O. S. Lih, V. K. Sudarshan, T. J. Hong, *et al.*, "Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ECG signals: A comparative study," *Information Sciences,* **377**: 17-29(2017).
15. S. Sharma, S. Jain, and S. Bhusri, " Two Class Classification of Breast Lesions using Statistical and Transform Domain features," *Journal Of Global Pharma Technology Methodology,* **9**: pp. 18-24, (2017).
16. N. Prashar, "Removal of electromyography noise from ECG for high performance biomedical systems," *Network Biology,* **8**: 12 (2018).
17. G. Singh and N. R. Prakash, "FPGA Implementation of Higher Order FIR Filter," *International Journal of Electrical and Computer Engineering (IJECE),* **7**: 1874-1881(2017).
18. J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng,* **32**: 230-236 (1985).