

RESEARCH ARTICLE



Analysis and classification of ECG beat based on wavelet decomposition and SVM

Sundari Tribhuvanam^{1*}, H C Nagaraj², V P S Naidu³

¹ Research Scholar, Department of Electronics, University of Mysore, Mysore, 570005, India. Tel.: +91-973-912-7272

² Department of Electronics, Nitte Research and Education Academy, NMIT Campus, Bengaluru, 560064, India

³ MSDF, FMCD, CSIR-NAL, Bengaluru, 560078, India



Received: 17-05-2020

Accepted: 13-06-2020

Published: 08-07-2020

Editor: Dr. Natarajan Gajendran

Citation: Tribhuvanam S, Nagaraj HC, Naidu VPS (2020) Analysis and classification of ECG beat based on wavelet decomposition and SVM. Indian Journal of Science and Technology 13(24): 2404-2417. <https://doi.org/10.17485/IJST/v13i24.452>

*Corresponding author. Sundari Tribhuvanam

Research Scholar, Department of Electronics, University of Mysore, Mysore, 570005, India. Tel.: +91-973-912-7272 stribhuvanam@yahoo.co.in

Funding: None

Competing Interests: None

Copyright: © 2020 Tribhuvanam, Nagaraj, Naidu. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment ([iSee](https://www.indjst.org/))

Abstract

Objectives : To extract the features of single arrhythmia ECG beat. To develop efficient algorithms for automated detection of arrhythmia based on ECG.

Methods/Statistical analysis: The methodology includes pre-processing and segmentation of ECG. Extraction of ECG features are to support the ECG beat classification and analysis of cardiac abnormalities using machine learning techniques. Wavelet decomposition is considered for feature extraction and classification with multiclass support vector machine. **Findings:** This work evaluates the suitability of the wavelet features of ECG for classifier. The proposed arrhythmia classifier results in an accuracy up to 98% for various classes of arrhythmia considered in this work. **Novelty/Applications:** This work is an assistive tool for medical practitioners to examine ECG in a limited time with their expertise to make the accurate abnormality diagnosis of the arrhythmia.

Keywords: Arrhythmia; classification; feature extraction; support vector machine; wavelet decomposition

1 Introduction

Cardiac diseases are the most common cause of death around the globe. The design of health monitoring systems is always a topic of active research to support the cardiac patient. Electrocardiogram (ECG) provides detailed information of the condition of the heart^(1,2). Cardiologists can infer heart conditions from ECG wave patterns and inter wave intervals. To assist the medical doctors, researchers have proposed many algorithms for segmentation and classification of ECG signals more precisely and correctly in real-time⁽³⁾. An arrhythmia classification system includes the pre-processing of ECG signal, abnormal beat segmentation, extraction of wavelet domain features and beat classification⁽⁴⁻⁶⁾. The objective is to identify the various ECG arrhythmias as per AAMI standard thereby assisting the cardiologist for early diagnosis of heart disease. Arrhythmia detection procedure differs in selecting the size of ECG signal window, ECG feature extraction and classification approaches⁽⁷⁾. The heart performance and prediction of future complications are done using Linear prediction method, Grid partitioning and Fuzzy C-means clustering for ECG classification⁽⁸⁾. ECG feature selection is implemented by Bacterial Forging Optimization (BFO) and Particle Swarm

Optimization (PSO), classification with Levenberg–Marquardt Neural Network classifier⁽⁹⁾. QRS detection is done with KNN algorithm and recurrent neural network for classification⁽³⁾. This work aims at identifying the suitability of ECG features to recognise different arrhythmic ECG beats by the classifier using machine learning techniques.

Table 1 indicates the characteristics of the significant waves and their origin of a normal human heart⁽¹⁰⁾.

Figure 1 shows the typical normal sinus rhythm for a healthy adult with a typical heart rate of 82 bpm⁽¹¹⁾.

Table 1. Characteristics ECG waves and its origin for a healthy adult

ECG Labels	Description	Origin
P wave	Upright, smooth, not notched or peaked	depolarization and contraction of the atria
QRS complex	positive deflection with a large, upright R , negative deflection with a large deep S wave	depolarization and contraction of the ventricles
ST segment	isoelectric, slanting upwards to the T wave, never normally depressed greater than 0.5 mm	Brief Resting period for ventricles before ventricular repolarisation
T wave	T wave deflection should be in the same direction as the QRS complex	repolarisation of the ventricles
U Wave	U wave amplitude < 1/3 T wave amplitude, more prominent at slow heart rates	after depolarization which interrupt or follow repolarisation

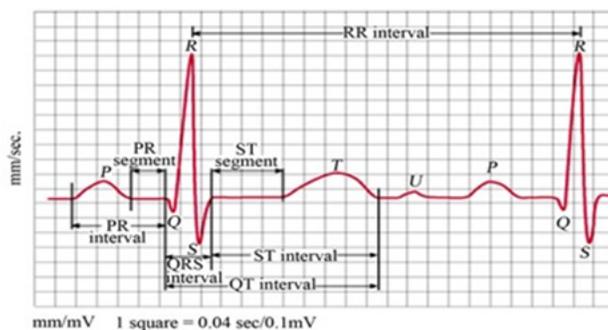


Fig 1. Features of normal sinus rhythm

2 Materials and Methods

To detect the cardiac abnormalities based on machine learning, the primary need is to analyse the abnormal beat segmented from the continuous ECG signal. The overall block diagram of the proposed system is given in Figure 2. The entire process is divided in five stages: pre-processing of ECG signal, segmentation, ECG feature extraction in wavelet domain, beat classification and performance analysis of the classifier. The pre-processing aims at improving the ECG signal quality by base line wander removal and elimination of noise. The signal is then segmented into single beats with R-point at the centre. The feature extraction step aims at wavelet feature extraction from the segmented ECG beats to construct the feature matrix. The feature matrix has N rows (F_{11} - F_{1N}) indicating the various features of the single ECG beat and n columns (F_{11} - F_{n1}) indicating the similar features of every ECG beat under consideration. This matrix serves as input to train and test the classifier model to perform required classification. The identification of the appropriate input features for classification is included in the performance analysis stage.

This paper considers wavelet coefficients and relative wavelet energy as the input features for classification.

2.1 ECG data acquisition

The ECG data is procured from the open source MIT/BIH arrhythmia database is used for training and testing the ECG classifier model. The database consists of 48 ECG records- each record has ECG signals from two channels for 30 minute duration segmented from 24-hour recordings of 47 individuals. The ECG signals are filtered by a 0.1–100 Hz band pass filter and digitized at 360 Hz. The database has annotation for both timing and beat class information verified by independent domain experts. A

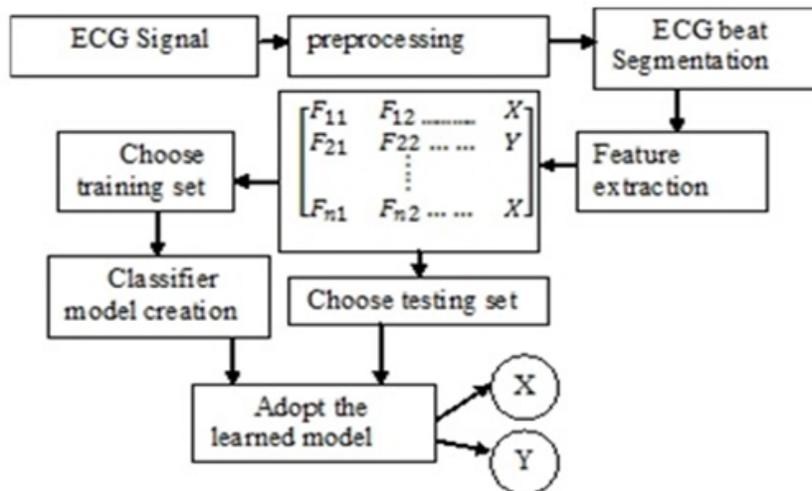


Fig 2. Overall flow diagram of arrhythmia classification

total of 13,648 beats are chosen from all 43 records are considered as test patterns for performance evaluation of classifier. In all records, only the modified limb lead II (MLII) signals are considered and labels are utilized to locate beats in ECG digitized signal⁽¹¹⁾. In this work, seven different arrhythmias are considered in accordance with standard of Association for the Advancement of Medical Instrumentation (AAMI). Table 2 indicate the different classes of arrhythmias along with timing information considered in this work. Several noises overlapping with the bandwidth of ECG signal is shown in the Figure 3⁽¹²⁾.

Table 2. ECG Arrhythmia classes

ECG Class	Class No	Figure Number	Record	Time
Normal beat (NL)	1	4(a)	115	0.939 to 3.783sec
Atrial premature beat(AP)	2	4(b)	223	435.514 to 438.358sec
Fusion of ventricular ad Normal beat(fVN)	3	4(c)	223	309.886 to 312.731sec
Junctional Escape beat(NE)	4	4(d)	222	47.883 to 550.728sec
Nodal premature beat(NP)	5	4(e)	234	844.475 to 847.319sec
Right Bundle Branch Block beat(RB)	6	4(f)	232	3.186 to 6.031sec
Fusion of paced and normal beat(fPN)	7	4(g)	217	161.731 to 164.575sec

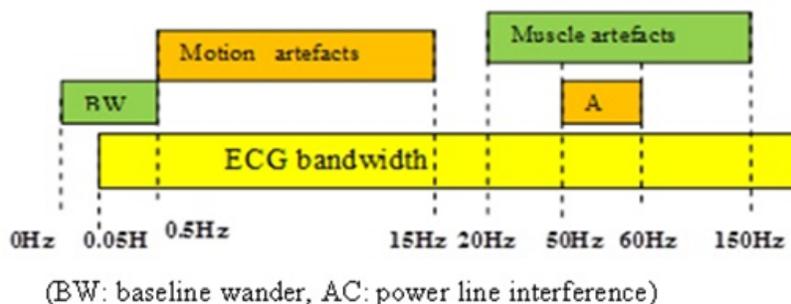


Fig 3. Various noise sources present in ECG signal

2.2 Pre-processing of ECG signal

The ECG pre-processing stage, prior to ECG feature extraction involves in segmentation, baseline correction and alignment of ECG (4,13,14). ECG segmentation is done to separate the ECG beat under consideration. Figure 4 shows the segmented ECG with 1025 samples for a duration of 3.075sec. Each segment may contain 3 to 5 beats depending on the heart rate and abnormality. Pre-processing of the beats include the following steps:

1. Normalization to compensate the probable error introduced during the acquisition of the ECG signal.
2. AC interference elimination by notch filter of second order.
3. Noise filtering- the noise present in the ECG is filtered by band pass IIR filter.
4. Correction of ECG baseline-the alignment of PQ segment of ECG beat with the reference (zero) line, achieved by subtracting the 10th level approximation signal from the filtered ECG beat with the discrete wavelets. In this work, Daubechies wavelet (db6) is taken because of its morphological similarities to normal ECG beat (15). The pre-processed ECG is further smoothed by an averaging filter to eliminate the flaws and to aid the peak detection process. The data acquisition, segmentation is done using cygwin-an open source tool and pre-processing, feature extraction and classification are implemented with MATLAB 2018A computing environment. Finally, the ECG features in wavelet domain are extracted from the ECG beats.

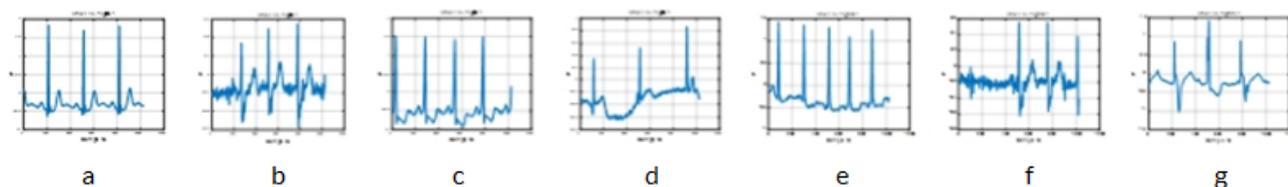


Fig 4. Typical ECG segments of various classes

2.3 ECG Feature Extraction

An ECG feature can be any intelligence extracted from the ECG beat to discriminate its type from others. The ECG features can be extracted in various forms directly from the signal's morphology and time information PQRST inter wave intervals. Features in frequency domain can be extracted with Fourier transform or using wavelet descriptors. Feature extraction involves the description of ECG beat and feature selection involves in choosing a subset of most significant features to improve the performance of the classifier. Discrete wavelet decomposition involves in selection of appropriate wavelet descriptor and decision on number of decomposition level to extract time and frequency information of the ECG beat. The levels are chosen based on the regions of the ECG beat that are correlated with the frequencies required for arrhythmia classification. The decomposition results in the approximate and detailed coefficients, which is the representation of original ECG beat. In this work, time-frequency representation involves single level one-dimensional Daubechies wavelet filters (5,8).

2.4 Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) is an effective method for representing the ECG beats in wavelet domain to find the abnormalities in ECG beat. The DWT implementation disintegrates the ECG beat in the form of mutually perpendicular set of wavelets. The accurate ECG features of time and frequency domain at high and low frequencies are extracted. The ECG beat is decomposed using a high scale, low pass filter (LPF) to yield "Approximation coefficients" (A). Simultaneously ECG beat is decomposed using a low scale, high pass filter (HPF) to result in "Detailed coefficients" (D). The entire process repeated to increase the frequency resolution. Hence, a set of filters have been used for decomposition of the ECG beats using DWT (13,15,16). Figure 5 (a) indicates the wavelet decomposition process. The sampling frequency(F_s) of ECG is 360Hz informs that frequency component of ECG upto 180Hz($F_s/2$). The eight levels ECG decomposition is considered in order to include all the useful frequency components. The Figure 5(b) shows the ECG decomposition process using DWT of the ECG signal to include all frequencies from 180Hz to 0.5Hz.

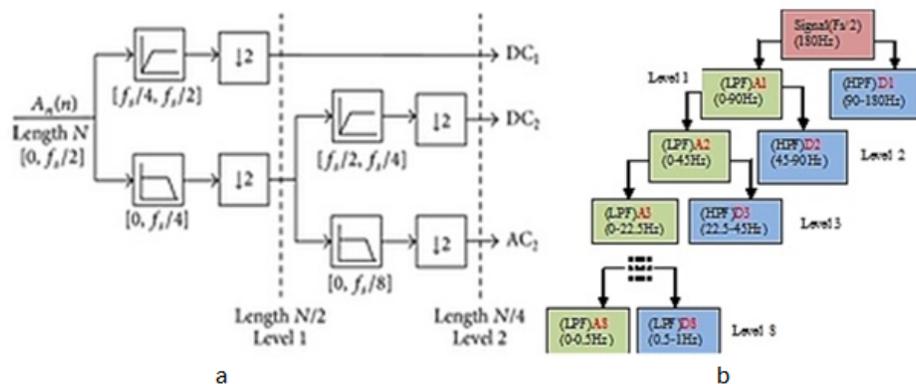


Fig 5. (a) Wavelet decomposition, (b) DWT decomposition of ECG

The energy distributions of the ECG signal for different arrhythmia classes are not similar at various frequency bands under consideration. The energy associated with wavelet coefficients indicates the energy of the signal at different composition level and can be singled out at distinct resolution levels. The energy associated with detailed coefficients at j^{th} level is computed as (16):

$$E_j = \sum_{k=1}^{\frac{N}{2^j}} |d_{j,k}|^2 \quad \text{---(1)}$$

Where $j=1, 2, \dots, J$ decomposition levels; $k = 1, 2, \dots, \frac{N}{2^j}$ Index of decomposition coefficients and N –number of samples in the ECG beat.

Energy of the approximation coefficients at J^{th} level is calculated as:

$$E_{J+1} = \sum_{k=1}^{\frac{N}{2^J}} |c_{J,k}|^2 \quad \text{---(2)}$$

The total energy of the ECG beat is:

$$E_{Total} = \sum_{j=1}^{J+1} E_j \quad \text{---(3)}$$

The time scale density or relative wavelet energy (RWE) of the ECG beat is computed in accordance with the equation $\sum D_j + A = 100, j = 1, 2, \dots, J$ for all wavelet coefficients. This is the representation of the energy corresponding to different frequency bands of the ECG beat.

$$D_j = \frac{E_j}{E_{Total}} \times 100, \quad j = 1, 2, \dots, J \quad \text{---(4)}$$

$$A = \frac{E_{J+1}}{E_{Total}} \times 100$$

The normalized relative wavelet energy at each level is computed for all wavelet coefficients as given in the equation (4). The mean \pm standard deviation (std) for the coefficients is calculated and this serves as the meaningful feature of ECG beat. Finally, the variance of the obtained coefficients is computed for each decomposition levels in order to extract the features from the ECG beats for abnormality classification (17).

2.5 ECG feature Selection and Classification

In this work, the support vector machine (SVM) is considered to classify the different arrhythmias. SVMs are statistical, supervised learning models—primarily described by partitioning hyper planes in multidimensional space and used for both classification and regression challenges. SVM is a binary classifier that takes a pair of input features and class labels, to separate cases of different labels (5,13,18,19). It analyses the input features, trains itself and then classifies the data by generating the support vectors. The support vectors are nothing but the coordinates of each data point. In this work, the SVM model adopts RBF kernel, One Versus One encoding scheme, 10 fold cross validation and simplex optimization routine.

3 Results and Discussion

3.1 Feature Extraction

The typical single ECG beat for 7 classes are as shown in the Figure 6. The morphology of ECG beats appears similar for the visual observation. These beats serve as input to the feature extractor. The ECG beat input is decomposed with dB6 wavelet. Nine wavelet features viz., D1, D2,...,D8 and A8 are extracted from every ECG beat of the database. The frequency components of ECG beat at each level of wavelet decomposition to generate the wavelet coefficients are given in the Table 3. The mean ± stdev of the mean, variance of wavelet coefficients at each level for all 7 classes are tabulated in Table 4 (a) and Table 4(b) respectively. They are included row wise in the feature matrix. The relative wavelet energy at each level computed as per equation (1) and are indicated in Table 4(c). The column of the feature matrix represents the similar features for different ECG beats. A set of 27 features are extracted from each ECG beat resulting in 702 combinations of input pairs taking two features at a time. Out of these only 351 feature pair combinations are valid without repeating any input feature pair.

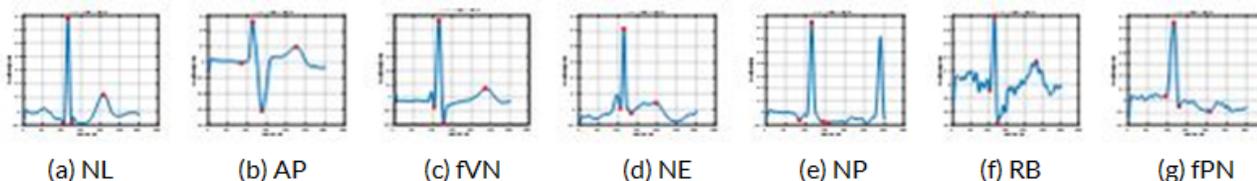


Fig 6. (a-g) Typical single ECG beats

Table 3. Frequency coefficients of ECG at each level of DWT

Level	1	2	3	4	5	6	7	8
Low freq(Hz)	A1(0-90)	A2(0-45)	A3(0-23)	A4(0-12)	A5(0-6)	A6(0-3)	A7(0-2)	A8(0-1)
High freq(Hz)	D1(90-180)	D2(45-90)	D3(23-45)	D4(12-23)	D5(6-12)	D6(3-6)	D7(2-3)	D8(1-2)

Table 4. (a) The mean ± standard deviation of Wavelet coefficients Mean

Coeff.	Normal	APB	fVN	NE	NP	RBBB	fPN
D1	5.6E-05± 0.000129	1.94E-05± 0.0001	-1.1E-05± 6.03E-05	2.09E-05± 0.000113	-8.3E-06± 4.93E-05	7.18E-06± 6.21E-05	-1.3E-05± 5.88E-05
D2	6.3E-05± 0.00013	6.78E-05± 0.0001	4.06E-05± 7.37E-05	9.68E-05± 0.0001156	-2.1359E-05± ±0.0001	1.78E-05± 0.000128	4.48E-05± 0.000111
D3	-0.0010± 0.001729	-0.001± 0.002174	-0.00016± 0.00148	7.8E-05± 0.001482	0.000428± 0.001264	-0.00145± 0.00185	0.001328± 0.002335
D4	-0.0095± 0.016663	-0.00312± 0.011545	-0.00156± 0.013694	-0.00875± 0.007459	-0.01432± 0.007069	-0.00748± 0.016431	-0.0053± 0.012197
D5	-0.0352± 0.046099	-0.04821± 0.024309	-0.03402± 0.067718	-0.0392± 0.008631	-0.09183± 0.028748	-0.08752± 0.048355	-0.04386± 0.020543
D6	-0.0161± 0.01662	-0.03917± 0.0291	-0.02622± 0.034624	-0.02187± 0.011867	0.017865± 0.011446	-0.04196± 0.033968	-0.03163± 0.0136

Continued on next page

Table 4 continued

D7	-0.0004± 0.002329	-0.00059± 0.0028	-0.00063± 0.00334	-0.0011± 0.002123	-0.00312± 0.001456	-0.00111± 0.001931	-0.00209± 0.001252
D8	9.25E-05± 0.000492	-0.00016± 0.0006	-0.00044± 0.000912	0.000106± 0.000588	-0.00013± 0.00058	-0.0005± 0.000876	-4.1E-05± 0.000409
A	-0.0912± 0.21282	0.042± 0.58	0.04246± 0.53708	-0.0176± 0.402959	-0.17516± 0.401611	-0.02743± 0.23439	-0.01491± 0.314685

Table 4b. The Mean ± Standard Deviation of Wavelet Coefficients Variance

Coeff.	Normal	APB	fVN	NE	NP	RBBB	fPN
D1	1.03E-06± 1.38E-06	1.52E-06± 3.12E-06	1.19E-06± 2.1E-06	1.34E-06± 9.5E-07	4.08E-07± 3.29E-07	1.11E-06± 1.12E-06	8.94E-07± 4.25E-07
D2	9.87E-05± 6.77E-05	0.000105± 8.97E-05	6.83E-05± 5.56E-05	0.000121± 9.4105E-05	3.7515E-05 ± 2.84352E-05	0.00021± 0.00016	0.000113± 7.3006E-05
D3	0.007566± 0.006555	0.007221± 0.0042	0.005337± 0.005666	0.00339± 0.002064	0.004804± 0.003861	0.01195± 0.0113	0.00277± 0.00161
D4	0.03149± 0.01099	0.03614± 0.021333	0.051037± 0.030506	0.01414± 0.007774	0.064294± 0.03714	0.0564± 0.04016	0.01826± 0.006245
D5	0.11594± 0.063926	0.11635± 0.075917	0.2266± 0.190437	0.04665± 0.032662	0.279557± 0.177452	0.28507± 0.3131	0.12± 0.073633
D6	0.1042± 0.0677	0.114706± 0.072531	0.26466± 0.271385	0.0403± 0.030327	0.24571± 0.170415	0.2118± 0.2575	0.1334± 0.056666
D7	0.3415± 0.556043	0.1786± 0.135366	0.4608± 0.541579	0.0996± 0.087643	0.54617± 0.393645	0.292± 0.2089	0.476± 0.375839
D8	0.3496± 1.028468	0.2296± 0.496333	0.40412± 1.11757	0.1232± 0.162529	0.23275± 0.155766	0.168± 0.1141	0.637± 1.1786
A	0.1404± 0.5005	0.5011± 1.228	0.52496± 1.096088	0.279± 0.827522	0.2981± 0.634533	0.1692± 0.2591	0.194± 0.60438

Table 4c. The Mean ± Standard Deviation of Normalized Wavelets Energy(RWE)

Coeff.	Normal	APB	fVN	NE	NP	RBBB	fPN
D1	0.00403± 0.004091	0.006± 0.0097	0.00174± 0.002811	0.01544± 0.011958	0.0023± 0.00115	0.00322± 0.003574	0.0063± 0.005467
D2	0.26335± 0.19168	0.295± 0.308	0.1052± 0.143757	0.76153± 0.627745	0.14034± 0.094226	0.3573± 0.266616	0.4122± 0.276355
D3	11.675± 9.0691	11.16± 8.84	4.7831± 3.787	11.5321± 6.591866	10.423± 8.51377	10.337± 7.000688	7.2091± 5.98137
D4	29.7144± 11.50642	26.48± 9.0	27.0184± 13.61289	26.4585± 7.602671	74.571± 16.91835	27.622± 8.844162	25.412± 11.85122
D5	55.2264± 19.04488	47.8± 15.13	59.389± 17.44655	47.4702± 17.5763	4.1094± 9.2548	59.3717± 14.77807	60.1597± 19.22822
D6	0.0073± 0.0161	0.016± 0.022	0.00426± 0.006008	0.00269± 0.004237	0.0032± 0.005987	0.00744± 0.01559	0.0034± 0.016189
D7	0.0035± 0.0104	0.013± 0.018	0.00299± 0.005611	0.0008± 0.00099	0.00358± 0.006904	0.00077± 0.001005	0.00274± 0.014021
D8	0.00415± 0.01024	0.007± 0.0099	0.00344± 0.00514	0.00096± 0.003639	0.00122± 0.002573	0.00037± 0.000494	0.00159± 0.00868
A	3.1021± 7.772	14.222± 21.68	8.69173± 13.09087	13.758± 19.47969	10.746± 14.08593	2.30022± 2.3378	6.79304± 12.0298

3.2 Classification

The extracted features are plotted for various classes of arrhythmias under consideration. Not all the features support perfect classification of the ECG abnormalities. Typical plots of extracted features across various class of arrhythmia are indicated in Figures 7, 8 and 9. The plots 7(a)-9(d) indicate that features extracted can be categorized as perfect, partially perfect and poor classification. This gives an approximate estimate of classification. The Table 5 indicates the summary of features useful for further classification. The perfect category is indicated as (i)NL and NP (ii)APB and NP(iii)NP and NE(iv)NP and RB (vi)NP and fPN classes which are perfectly classifiable with respect to mean of D8 coefficients as indicated in Figure 7(d). The NL and NE with respect to variance of D4, NE and NP with respect to variance of D6 as shown in Figure 8 (b) and Figure 8(c)

respectively. The NP and other classes are perfectly separable with respect to RWE of D4 and D5 coefficients as indicated in 9(b) and 9(c). The partially perfect category shows the partial overlapping of mean \pm std of different classes- Mean of D3 for NP and RB class as shown in Figure 7(a), NL and NP class as indicated in Figure 7(c). Total overlapping indicates the poor classification as seen in Figure 8 (a) with NL and APB for mean of D3. NL and NE for mean of D6 as in Figure 7(c), NP and RB for RWE of D3 coefficients as seen in Figure 9 (a). These are only indicative and do not give the accuracy of classification.

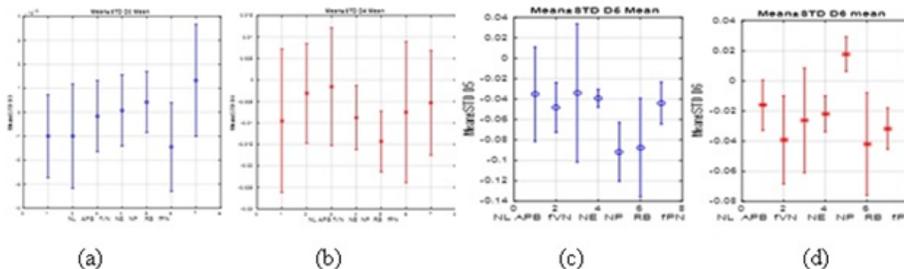


Fig 7. (a — d) Comparison plot of mean of wavelet coefficient at D3, D4, D6, D8 levels

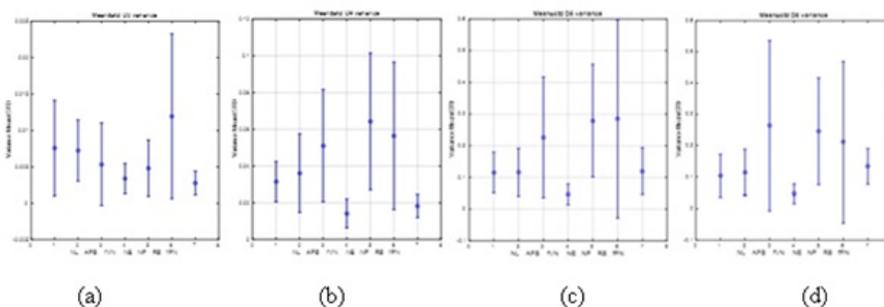


Fig 8. (a — d) Comparison plot of variance of wavelet coefficient at D3, D4, D6, D8 levels

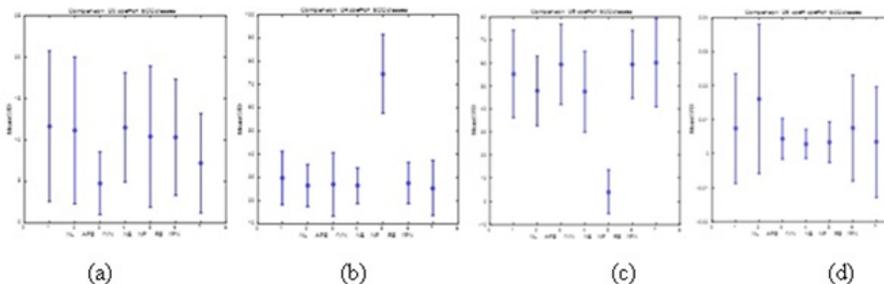


Fig 9. (a — d) Comparison plot of RWE of wavelet coefficient at D3, D4, D5 and D6 levels

3.3 Classification with SVM

The classification can be further quantised by the implementation of the multiclass support vector machine (MSVM). It gives a measure of classification in terms of accuracy. The SVM takes two features at a time(column wise) from the feature matrix as input and trains itself with training data set and classifies the testing dataset. The training and test datasets are selected at random from the feature database but their ratio can be set by the programmer. The accuracy of the classifier is the benchmark to select the particular feature pair for the ECG classification. The different input to MSVM leads to various degrees of classification

Table 5. Summary of the suitability of the ECG feature pairs for the SVM

Figure Number	Perfect	Partially perfect	Poor
7(a)-D3Mean	-	NP-RB, NE-RB, RB-fPN	NL-APB, NL-fVN
7(b)-D4Mean	-	fVN-NP, APB-NP	NL with all others
7(c)-D5Mean	NE-NP, NP-fPN	APB-NP, APB-RB	NL-APB, NE, fPN, NE-RB
7(d)-D6Mean	NP-NL, APB, NE, RB, fPN	fVN-NP, NL-RB, NL-APB	NL-fVN, NL-NE,
8(a)- D3-Var	-	-	NL-All other
8(b)-D4-Var	NE-NP, NL, fPN-NP,	NE-RB, APB,	NL-APB, fVN, RB
8(c)-D6-Var	NE-NP	fVN-NE, APB, NL	NP-RB, NP-fVN
8(d)-D8-Var	NE-fPN, NE-NP	APB-NE, APB-NP	NL-APB, fVN, NP, RB, fPN
9(a)- D3-Var	-	fVN-fPN, fVN-NE	NL-All other, NE- NP, NE-RB
9(b)-D4-Var	NP-All other	-	NL-APB, NL-fVN, NL-NE, NL-RB, NL-fPN
9(c)-D6-Var	NP-All other	APB-fVN, fVN-NE	NL-APB, NL-fVN, NL-NE, NL-RB, fPN
9(d)-D8-Var	-	-	NL-All other

accuracies. The accuracy above 90% are considered as good and 80-90% as moderate and below 80% as poor for the selected input feature pairs from the dataset. Table 6 and Figure 10 indicate the segregation of the input pairs leading to different ranges of accuracy of MSVM. Table 8 (a & b) lists input feature pair and the corresponding accuracies for the good and moderate category with accuracy greater than 80%. Only those pairs of input with resulting accuracy more than 90% are considered for further processing, to obtain better accuracy of classification.

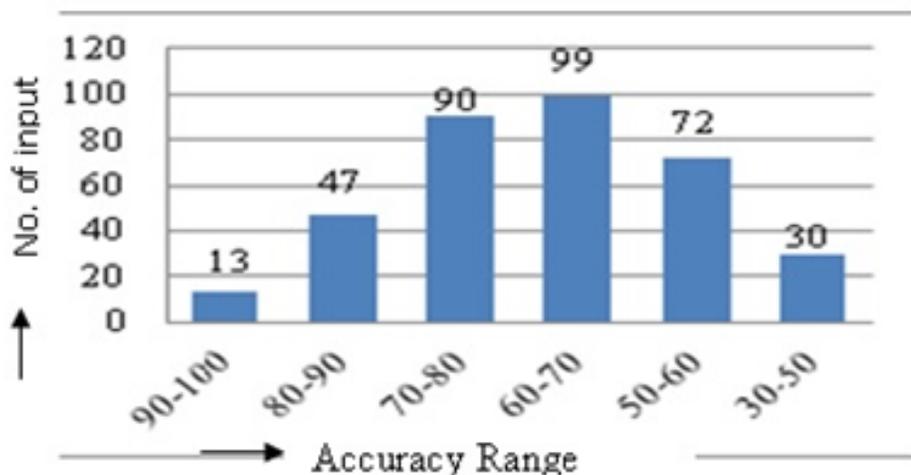


Fig 10. Distributions of input feature pair vs accuracy across SVM inputs

The input features are fed to the SVM with different ratios of training and testing data.

Observations indicate that classification accuracy is improved for the 7:3 and 7.5:2.5 ratios with 10 successive iterations. The results are tabulated in Table 8 (a) and Table 8 (b) indicate the accuracies for typical input pairs D4Var-D4RWE and D4Var-D7Var respectively Table 8 (a).

The performance of the SVM is measured in terms of classification accuracy. Classification accuracy of 97.67% is obtained for input feature pairs D4Var-D4RWE with 8.5:1.5 training and testing data as indicated in Table 8 (a). An output accuracy of 98% is achieved for selected input feature set for D4Var - D7Var with 8: 2 of training and testing data as seen in Table 8 (b). The MSVM output for selected input feature set is indicated in Figure 11 (a) and Figure 12 (a). As the training and test datasets are chosen at random, the MSVM is tested for multiple iterations with the same input feature pairs and testing datasets. The

Table 6. Performance of MSVM

Accuracy Range (%)	No. of SVMInput Pairs
>95	4
90<A<95	9
80<A<90	47
70<A<80	90
60<A<70	99
50<A<60	72
<50	33

summary of performance analysis of the SVM for typical input features are indicated in the Table 9. An accuracy upto 98.63% is achieved for the database under consideration. Figure 13 represents the overall accuracy of SVM for different training and testing data ratios [Figure 11(a)-(b)], [Figure 12(a)-(b)], [Table 9], [Figure 13].

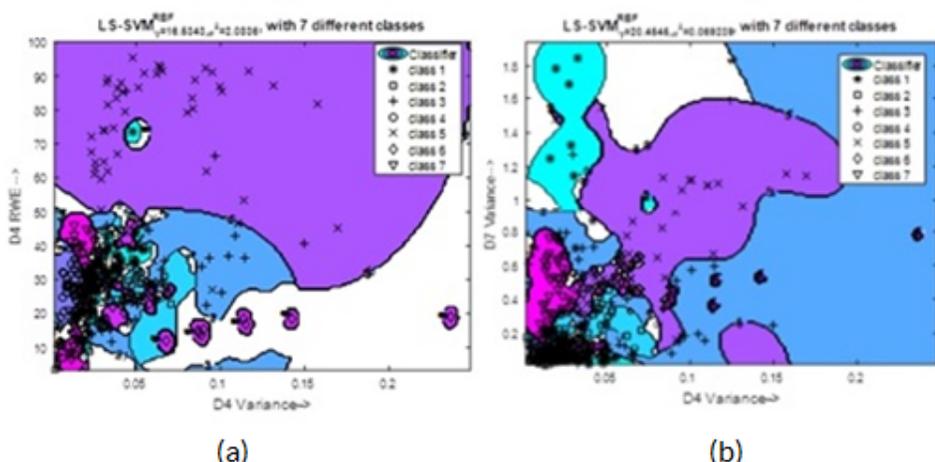


Fig 11. (a) D4Var:D4 RWE 8 5:1.5 (train:test) Accuracy 98.28% (b) D4Var:D7Var 8:2 (tain:test) Accuracy 97.67%

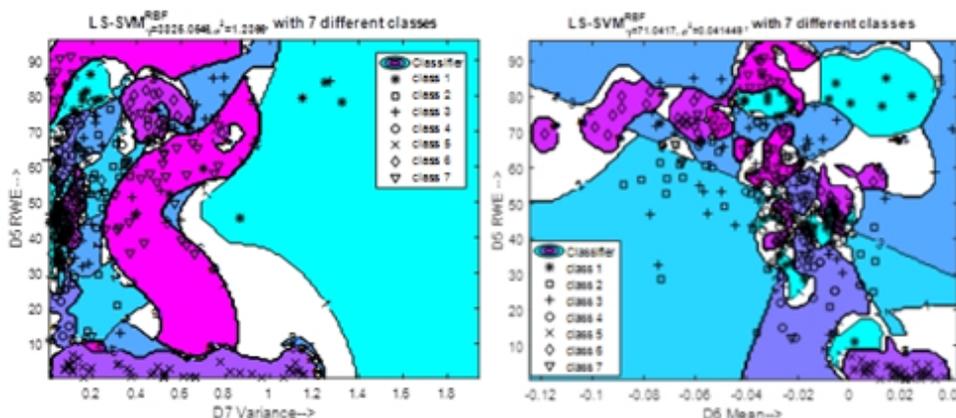


Fig 12. (a) D7Var RWE7:3 (train:test) Accuracy 95.4%, (b) D6Mean–D5 RWE 7:3 (train:test) Accuracy 95.4%

The performance of the proposed method is compared with other related works. Seven arrhythmia classification methods were considered to compare with this method. These methods differ in method of ECG feature extraction, number of arrhythmia

Table 7. (a) List of feature pairs whose SVM accuracy is above 90%

Sl. No	Feature pair	Accuracy	Sl. No	Feature pair	Accuracy	Sl No	Feature pair	Accuracy
1	D2M-D6Mean	89.66	6	D5Mean-D6Mean	97.26	10	D6Mean-D7Var	94.25
2	D4Var-D7Var	91.78	7	D6Mean-D6Var	91.95	11	D6Mean-D8Var	91.95
3	D4Var-D4RWE	91.78	8	D6Mean-D7Mean	90.8	12	D6Mean-D3RWE	93.1
4	D6Mean-D5RWE	93.1	9	D6Mean-D6RWE	95.4	13	D6Var-D6RWE	91.95
5	D7Var-D5RWE	95.4						

Table 7(b). Feature pairs whose SVM accuracy is between 80 and 90%

1	D1M-D4Var	83.56	17	D2Var-D5Mean	85.06	33	D4Var-D6mean	82.19
2	D1M-D6M	80.46	18	D2Var-D2RWE	86.21	34	D4Var-A8RWE	79.45
3	D1Var-D4Var	79.31	19	D3Mean-D6Mean	80.82	35	D5Mean-D8Var	80.82
4	D1Var-D6Mean	85.06	20	D3Mean-D6Var	82.19	36	D5Mean-D2RWE	81.61
5	D1Var-D7Mean	79.31	21	D3Mean-D5RWE	89.04	37	D5Mean-D5RWE	86.21
6	D1Var-D2RWE	80.46	22	D3Var-D5var	84.93	38	D5Var-D6Mean	86.21
7	D1Var-D3RWE	80.46	23	D3Var-D5RWE	80.82	39	D6Mean-A8Mean	85.06
8	D2Var-D2RWE	86.21	24	D6Mean-D1RWE	80.46	40	D6Mean-D4RWE	87.3
9	D2Var-D5RWE	87.36	25	D6Mean-D2RWE	85.06	41	D6Mean-D7RWE	81.61
10	D6Mean-D8RWE	83.91	26	D6Var-D7Var	87.36	42	D6Var-D8Var	81.61
11	D6Var-D5RWE	80.46	27	D7Mean-D7Var	83.91	43	D7Mean-D4RWE	85.06
12	D7Mean-D5RWE	83.9	28	D7Var-D1RWE	87.36	44	D7Var-D2RWE	79.31
13	D7Var-D4RWE	83.91	29	D8Mean-D2RWE	79.31	45	D8Mean-D4RWE	81.31
14	D7Var-D4RWE	81.26	30	D8Mean-D5RWE	87.36	46	D8Var-D4RWE	79.31
15	A8Mean-D5RWE	88.51	31	A8Var-D4RWE	85.06	47	A8Var-D5RWE	88.51
16	D4RWE-D7RWE	85.06	32	D4RWE-A8RWE	82.76			

Table 8. (a) Performance of MSVM (Accuracy) for various ratios of training: test data over multiple iterations for input D4 Var-D4 RWE

Ratio of training : test data	7:3	93.1	93.1	91.95	91.95	86.21	86.21	93.1	94.25	94.25	94.25
	7.5:2.5	91.78	93.95	91.78	80.82	90.41	83.56	90.41	91.78	90.41	90.41
	8:2	94.83	94.83	93.1	91.38	91.38	79.31	91.38	81.03	84.48	94.83
	8.5:1.5	93.02	97.67	93.02	93.02	68.38	84.1	84.1	93.02	97.67	72.09

Table 8(b). Performance of MSVM (Accuracy) for various ratios of training : test data over multiple iterations for input D4Var- D7Var

Ratio of training : test data	7:3	72.4	91.95	72.4	87.4	67.8	90.8	73.56	87.36	66.67	72.4
	7.5:2.5	65.8	76.7	87.7	79.5	89.0	89.0	95.89	67.12	94.52	67.1
	8:2	89.7	98.3	89.7	93.1	94.8	77.6	89.66	82.76	87.93	68.97
	8.5:1.5	83.7	74.4	79.1	88.4	93.0	90.7	79.07	86.05	93.02	86.05

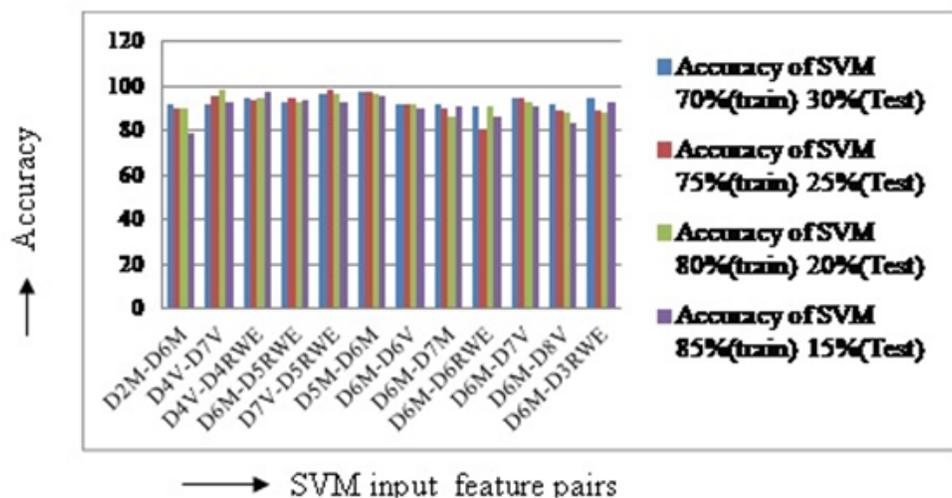


Fig 13. Performance of the SVM for various input features pairs

Table 9. Performance analysis of MSVM for chosen input feature pairs

Sl No	Feature pair	Accuracy of MSVM							
		70%(train) 30%(Test)	I. No	75%(train) 25%(Test)	I. No	80%(train) 20%(Test)	I. No	85%(train) 15%(Test)	I. No
1	D2M-D6M	91.95	3	90.41	8	89.66	11	79.07	8
2	D4V-D7V	91.95	2	95.89	7	98.28	2	93.02	5
3	D4V-D4RWE	94.25	8	93.78	2	94.83	1	97.67	2,9
4	D6M-D5RWE	93.1	4	94.52	4	93.1	9	94.03	1,2
5	D7V-D5RWE	96.55	3,5,8	98.63	8	96.55	5,6	93.02	1,8,9
6	D5M-D6M	97.7	5	97.26	6	96.55	1,3	95.35	6
7	D6M-D6V	91.95	1	91.78	10	91.38	9	89.55	7
8	D6M-D7M	91.95	6	90.41	4	86.21	5	90.8	9
9	D6M-D6RWE	90.8	9	80.46	3,8	90.8	9	86.05	1
10	D6M-D7V	94.25	3,7, 11	94.52	8	93.1	2,3,4,5,8	91.04	3,5,6,8
11	D6M-D8V	91.95	10	89.04	8	87.93	3,7,8,9	83.58	1,5,6,10
12	D6M-D3RWE	94.25	4,10	89.04	8	87.93	6	92.54	7

M-Mean, Var-Variance, RWE-Relative Wavelet Energy, I. No-Iteration Number

beat type and type of classifier considered. The performance of the methods are indicated in Table 10. The results shows that the proposed method considers increased number of wavelet features of ECG and SVM classifier results in considerable accuracy of classification of higher number of arrhythmia beat types in comparison with other methods using the same database. [Table 10]

Table 10. Comparison of classification accuracy of the proposed method and some of the work performed based on the MIT-BIH database

Feature extraction Method	Beat types	No. of ECG features	classifiers	k-fold cross validation	Accuracy
(15) Wavelet features	3	-	NN	-	96.5
(3) Morphological features	2	-	RNN	-	79% for VEB* 96% SVEB**
(4) Higher order spectra	5	22	NN,SVM(RBF)	10	99.03
(5) wavelet multi-resolution analysis(bior6.8 wavelet)	6	-	SVM(RBF)	10	99.70
(13)Morphological features	7	17	SVM(RBF)	10	88.5
(14)Time domain and HoS features	7	14	SVM(RBF)	10	95.3
(this work) Wavelet features	7	27	SVM(RBF)	10	98.3

*ventricular ectopic beats ** supraventricular ectopic beats

4 Conclusion

Automatic detection of cardiac abnormalities is the need of the day for medical professionals for early diagnosis of heart disease. The study of cardiac activities through ECG provides a mean for researchers to develop various algorithms using machine learning techniques. This work presents the classification of cardiac arrhythmia with the ECG single beat analysis. It focuses on ECG feature extraction with multiresolution analysis in wavelet domain and identification of suitable ECG features for arrhythmia classification with multiclass support vector machine. A set of input feature pairs for classifier model are identified to find the correctness of arrhythmia classifier with different training and test dataset ratios and successive iteration of MSVM. Classification accuracy upto 98% is achieved for selected input feature pairs. Higher accuracy of classification can be achieved by incorporating the feature reduction techniques and considering larger size of the database in each class of abnormalities.

References

- Liu F, Liu C, Jiang X, Zhang Z, Zhang Y, Li J et al. Performance Analysis of Ten Common QRS Detectors on Different ECG Application Cases. *Journal of Healthcare Engineering*. 2018;2018:1–8. Available from: <https://dx.doi.org/10.1155/2018/9050812>.
- Osman J, Tan SC, Lee PY, Low TY and Jamal R. Sudden Cardiac Death (SCD) – risk stratification and prediction with molecular biomarkers. *Journal of Biomedical Science*. 2019;26(1):39–40. Available from: <https://dx.doi.org/10.1186/s12929-019-0535-8>.
- Scirè A, Tropeano F, Anagnostopoulos A and Chatzigiannakis I. Fog-Computing-Based Heartbeat Detection and Arrhythmia Classification Using Machine Learning. *Algorithms*. 2019;12(32):1–21. Available from: <https://dx.doi.org/10.3390/a12020032>.
- Chashmi AJ and Amirani MC. An efficient and automatic ECG arrhythmia diagnosis system using DWT and HOS features and entropy- based feature selection procedure. *Journal of Electrical Bioimpedance*. 2019;10(1):47–54. Available from: <https://dx.doi.org/10.2478/joeb-2019-0007>.
- Qin Q, Li J, Zhang L, Yue Y and Liu C. Combining Low-dimensional Wavelet Features and Support Vector Machine for Arrhythmia Beat Classification. *Scientific Reports*. 2017;7(1):6067–6067. Available from: <https://dx.doi.org/10.1038/s41598-017-06596-z>.
- da S Luz EJ, Schwartz WR, Cámara-Chávez G and Menotti D. ECG-based heartbeat classification for arrhythmia detection: A survey. *Computer Methods and Programs in Biomedicine*. 2016;127:144–164. Available from: <https://dx.doi.org/10.1016/j.cmpb.2015.12.008>.
- Alfarhan KA, Mashor MY, Saad ARM, Azeez HA and Sabry MM. Effects of the Window Size and Feature Extraction Approach for Arrhythmia Classification. *Journal of Biomimetics, Biomaterials and Biomedical Engineering*. 2017;30:1–11. Available from: <https://dx.doi.org/10.4028/www.scientific.net/jbbbe.30.1>.
- Rahouma HK, Muhammad HR and Eldahab FAHMAA. Analysis of Electrocardiogram for Heart Performance Diagnosis Based on Wavelet Transform and Prediction of Future complications. *Egyptian Computer Science Journal*. 2017;41(2):11–30. Available from: <https://doi.org/10.13141/RG.2.2.35948.41608>.
- Kora P and Kalva SR. Hybrid Bacterial Foraging and Particle Swarm Optimization for detecting Bundle Branch Block. *SpringerPlus*. 2015;4(1):1–19. Available from: <https://dx.doi.org/10.1186/s40064-015-1240-z>.
- Yanowitz GF . 2016,. Available from: <http://library.med.utah.edu/kw/ecg/>.
- "the MB . . Available from: <http://www.physionet.org/physiobank/database/mitdb>.
- Hadji MSE, Salleh M, Rohani MF and Kamat M. Wavelet-based Performance in Denoising ECG Signal. In: and others, editor. International Scientific conference of Pure Science (ISCPs) . 2016,. Available from: <http://dx.doi.org/10.1145/3015166.3015212>.
- Tribhuvanam S, Nagaraj HC and Naidu VPS. ECG Abnormality Classification with Single Beat Analysis. In: and others, editor. International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN) . 2019,. Available from: <https://doi.org/10.1109/ViTECoN.2019.8899410>.

- 14) Tribhuvanam S, Nagaraj HC and Naidu VPS. Arrhythmia Classification with Single Beat ECG Evaluation and Support Vector Machine”. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*. 2019;12:2278–3075.
- 15) Sarkaleh MK. Classification Of Ecg Arrhythmias Using Discrete Wavelet Transform and Neural Networks. *International Journal of Computer Science, Engineering and Applications*. 2012;2(1):1–13. Available from: <https://dx.doi.org/10.5121/ijcsea.2012.2101>.
- 16) Rajani J and Naidu VPS. Bearing Fault Diagnosis using DWT and SVM. In: International Journal of Engineering Research & Technology, NCESC-2018 Conference Proceedings. vol. 6 . 2018,. p. 1–6.
- 17) García M, Ródenas J, Alcaraz R and Rieta JJ. Application of the relative wavelet energy to heart rate independent detection of atrial fibrillation. *Computer Methods and Programs in Biomedicine*. 2016;131:157–168. Available from: <https://dx.doi.org/10.1016/j.cmpb.2016.04.009>.
- 18) Zhe, Wang X and Xue. Springer International Publishing Switzerland . 2014,. Available from: <https://www.scribd.com/document/322745147/Support-Vector-Machines-Applications>.
- 19) Hsu CW and Chang CC . 2016,. Available from: <https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>.