Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 132 (2018) 11–21

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Intelligence and Data Science (ICCIDS 2018)

Efficient wavelet families for ECG classification using neural classifiers

Ritu Singh[a,*], Rajesh Mehta[b], Navin Rajpal[a]

[a] *USICT,Guru Gobind Singh Indraprastha University,Dwarka, New-Delhi, India*
[b] *Thapar Institute of Engineering and Technology, Patiala , Punjab, India*

Abstract

The emerging technologies in science and engineering helps in developing predictive models for accurate diagnosis (wellness) of diseases such as electrocardiogram (ECG) and electroencephalogram (EEG) related to heart and brain respectively. In this paper, a comparative study of neural classifiers such as back propagation neural network (BPN), feed forward network (FFN) and radial basis function neural network (RBFNN) on ECG signals using classical wavelet transform is presented. The ECG files of normal and abnormal classes depending upon the beats present in the signal are taken from the standard arrhythmia database from MIT-BIH. After removal of noise from the original signal, the features selected using discrete wavelet transform with different wavelet families like daubechies, symlet, coiflet, biorthogonal and reverse biorthogonal are supplied as input to different neural classifiers. The performance of different neural classifiers are figured out by the sensitivity (Se), the positive predictivity (PP), specificity (Sp) and accuracy (Acc).From the experimental results, it is inferred that RBFNN gives Se 100, PP 100, Sp 100, and Acc 100% with the feature extraction using daubechies wavelet as compared to other classifiers. This gives an intellectual diagnosis approach to health care system using signal processing and neural network.

Keywords: ECG; MIT-BIH, DWT; Wavelet families; BPN; FFN; RBFNN.

1. Introduction

Human health care is a major issue for today's ongoing researches in medicine. Cardio vascular disease (CVD) remains number one cause of death. Globally, an average of one person dies from CVD followed by stroke and heart

* Corresponding author.
E-mail address: ritumtech@gmail.com .

10.1016/j.procs.2018.05.054

failure every 40 seconds [1]. This burden of death risk factors has motivated researchers to contribute patterns, models and algorithms. ECG signal processing and feature classification based on machine learning plays vital role in supporting above mentioned system. The biological signals are non stationary, hence abnormalities may not appear throughout the time. So, efficient examination of ECG pattern is required. Operative Anatomy of the heart gives morphological guidelines for better interpretation. The language of heart can be represented in the form of chambers and ECG waveform, as shown in Fig. 1. The electric simulation of heart originates from SA node and moves down to AV node. This is done by pacemaker cells. The potential difference i.e. the change in electric charge in these cells occurs due to depolarization and repolarization.

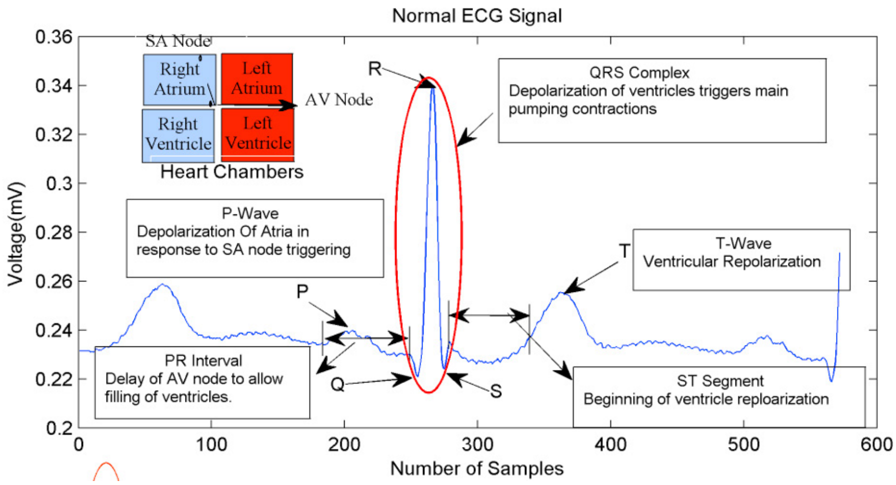


Fig.1. Heart Chambers and ECG Signal

For cardiologist, ECG comes out to be wealth of useful information for assessment even being simplest and oldest. ECG assets are QRS complex, peaks like P, T, PR interval and ST segment. Their misinterpretation can lead to serious chronic illness .So several techniques and algorithms have been developed for accurate analysis and ECG arrhythmia classification .Some proposed techniques are fuzzy c-means clustering[2],extracting cumulant features through principal component analysis (PCA) using neural classifiers[3], Wavelet Coefficient[4] ,RBF networks [5] , using DWT,PCA for feature extraction, support vector machine (SVM)and RBF as classifier [6] and also (SVM) and particle swarm optimization (PSO) for high accuracy ECG classification[7]. H.M Rai, *et al.* used DWT for morphological extract and BPN, FFN and MLP for classification [8].

In this paper, ECG data as normal and abnormal class based data using discrete wavelet transform with three different classifiers named as BPN, FNN and RBFNN have been crafted. Further, same extraction and classification is done using different wavelet families.

2. Materials

The input signal is ECG of (.dat) form. Fig.2 shows proposed block diagram .This data file belongs to arrhythmia database from MIT-BIH from physionet [9]. Each record includes three files. These are a binary .dat file having sampled signals, header .hea file having digitized sample description and an annotation .atr file containing labels regarding features of signal. There are 48 records with 360 Hz sampling frequencies. The window length is 1800 samples for 5 sec and signal length is 650,000 samples for 30 minutes duration having ±5 mV range and 11- bit resolution. The data extracted can be record based or beat based [10]. Records used in this research are mentioned in Table 1.

3. Preliminaries

3.1. Method

The statistics and frequency of ECG keeps on changing with time. It is non-stationary in nature. Frequency content of signal helps to determine pathological condition easily. For biomedical signal, time-frequency representation of the signal is needed and this is efficiently performed by using wavelet transform (WT) which is suitable for all frequency ranges.

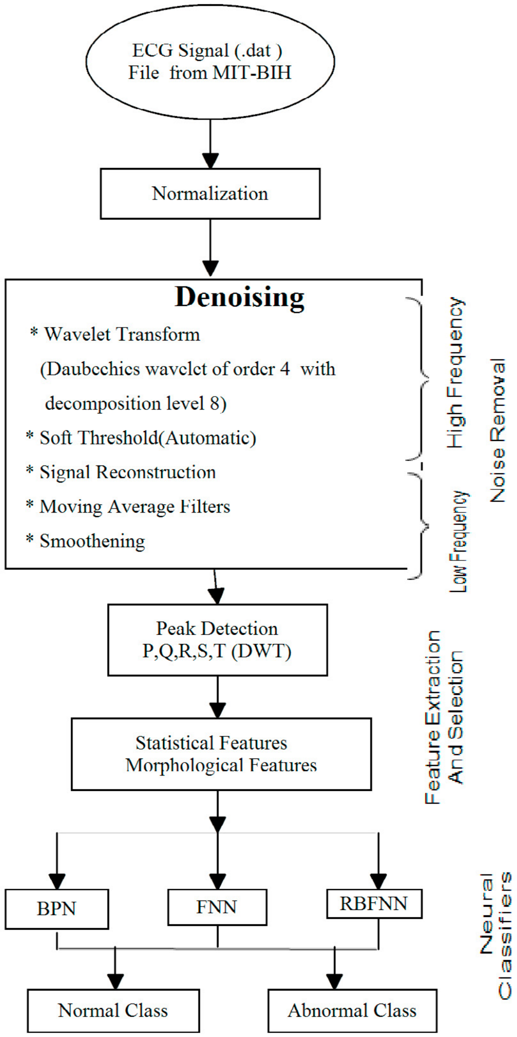


Fig.2. Proposed Block Diagram

Wavelet Analysis

Fourier transform gave birth to wavelet transform which is based on multi scale concept. It provides time and frequency information at same particular instant. WT can be broad, more spacious at low frequency and narrow or tapering at high frequency depicting varying window size. It can be computed discretely on the time-frequency plane to reduce redundancy. Feature selection and reduction are major application of WT [11].

Multi-resolution Wavelet Transform

The DWT analyse ECG signal at different resolution by disintegrating into several successive frequency bands. And multi-scale attribute empowers the signal. Wavelets are generated from mother wavelet having single prototype. This is done by the process of dilation and shifting.

$$\phi_{m,n}(t) = \frac{1}{\sqrt{a^m}}\phi\left(\frac{t - nba^m}{a^m}\right) \tag{1}$$

Where φ is the wavelet function having m as the wavelet dilation control integer and n as the wavelet translation control integer. ‘a’ and ‘b’ are the scale and location parameter respectively. The logarithmic scaling of both the factors by the power of two is called as the dyadic grid arrangement. Seeing the special case by putting a=2 and b=1, in Eq (1) comes out to be

$$\phi_{m,n}(t) = 2^{-m/2}\phi\left(2^{-m}t - n\right) \tag{2}$$

Decomposition is followed by separating time domain signal into high and low pass coefficients. This is done by passing sampled discrete time signal x[n] through a high pass and low pass filters giving A and D, approximation and detail coefficients. First level decomposition results in Eq (3) and Eq (4) also known as multiresolution decomposition.

$$h1[n] = y_{high}(k) = \sum_{n} x[n]g[2k-n] \tag{3}$$

$$a1[n] = y_{low}(k) = \sum_{n} x[n]h[2k-n] \tag{4}$$

h1 [n] and a1 [n] are the output of the high pass and low pass filters respectively. The input ECG signal is having sampling frequency 360 Hz and is down sampled by a factor 2 a shown in Fig.3. The choice of mother wavelet and its decomposition level depends on the shape and biology of the authentic ECG signal [12]

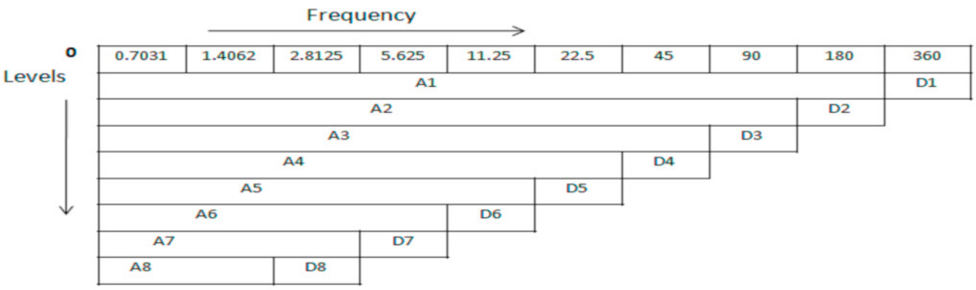


Fig.3. DWT Decomposition Tree

Different wavelet families are evaluated with different orders and finally the best efficient one is selected. Table 3 shows different wavelet families. The Daubechies (db4) wavelet with decomposition level 8 provides much appropriate results.

3.2. *Pre-processing*

The raw ECG is impaired with disturbances. This is due to skin electrode contact, motion artifacts and power line interference [13]. In this stage, first normalization is done to reduce the DC offset and to eliminate the amplitude variance. Then denoising is done in two stages as shown in Fig.2. Daubechies wavelet with vanishing moments four is being used with eight decomposition level. In the first stage, WT decomposes the signal into successive coefficients and after that automatic soft thresholding technique is applied to remove high frequency noise. Using original and changed coefficients, signal is again obtained. In the second stage, Baseline wander noise is removed. Its range is 0.15 - 0.5 Hz and is due to respiration of human body and electrode impedance. It is removed by using moving average filter and smoothing the data. The outcomes are shown in Fig.4 for record no. 103 and Fig 5 for record no. 217.

3.3. Feature Extraction and Selection

WT has ability to give compressed values known as features. Banerjee *et al.* [14] used multiresolution concept with adaptive thresholding to extract features. The combination of statistical and morphological features (79) is used to make the dataset and bi-frication is shown in Table 1. To have efficient and the best result, input to the classifier should be selected in such a way that the hidden layer neurons should not get exceeded in number. Statistical features are calculated after wavelet decomposition. And for calculating morphological features R-peak detection is required. To find the QRS complex and R peak, load denoised ECG signal and decompose up to eight level using

Daubechies (db4) wavelet. Select the best similar (A2) down sampled or decomposed low frequency component and select a threshold value which is equal to 0.60 times of maximum value of selected approximation coefficient. Find values which are greater than threshold and R-peaks which are 10 samples apart according to A2. Map the detected positions in decomposed signal to original signal. Lastly search within a window of ±20 samples in original signal. Detected maximum peaks collection is stored in R_Loc array and Ramp represents amplitude at the original scale. Similarly, other peaks are also detected by creating and selecting a window keeping R_Loc as reference. Fig.6 shows all detected peaks of record no. 234.

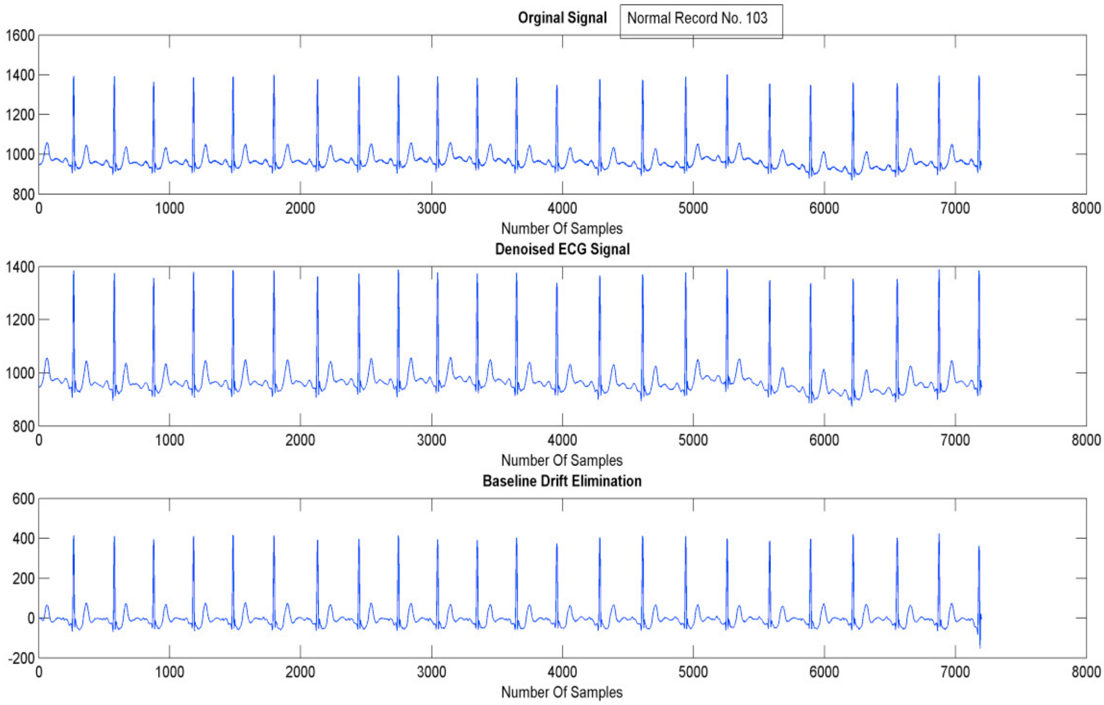


Fig.4. Pre-processed ECG Signal No.103 (Original, Denoised, Baseline Eliminated).

3.4. Classification Techniques

The neural network (NN) provides optimized results in classifying patterns which have similar characteristics [18]. Its classifiers BPN, FFN and RBFNN are powerful tools which are used in present method. The general architecture consists of 79 input features, one hidden layer having 20 neurons and 2 output class as shown in Fig.7 (a). RBFNN is a hybrid network which has non-linear organization having radial basis function in the hidden layer. Its neurons get trained and store prototype from training set. For new input classification, each neuron calculates the Euclidean distance between new input and it trained neuron [19]. RBFNN is known for its high speed and accuracy.

3.5. Performance Parameters

From the confusion matrix results, conclusion related to the classifier performance can be judged. The measures are sensitivity (Se), the positive predictivity (PP), specificity (Sp) and accuracy (Acc).

$$Se(\%) = \frac{TP}{TP+FN} \times 100\,\% \tag{5}$$

$$PP(\%) = \frac{TP}{TP+FP} \times 100\% \tag{6}$$

$$Sp(\%) = \frac{TN}{TN+FP} \times 100\% \tag{7}$$

$$Acc(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100\,\% \qquad (8)$$

TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative.
The true word signifies that the input exactly matches targeted output i.e. TP means normal class detected as normal and TN means abnormal class classified as abnormal. Similarly, false means opposite i.e. FP means normal class detected as abnormal and FN means abnormal class classified as normal. All above terms are shown in Fig.8. (d) extracted from confusion matrix of neural network.

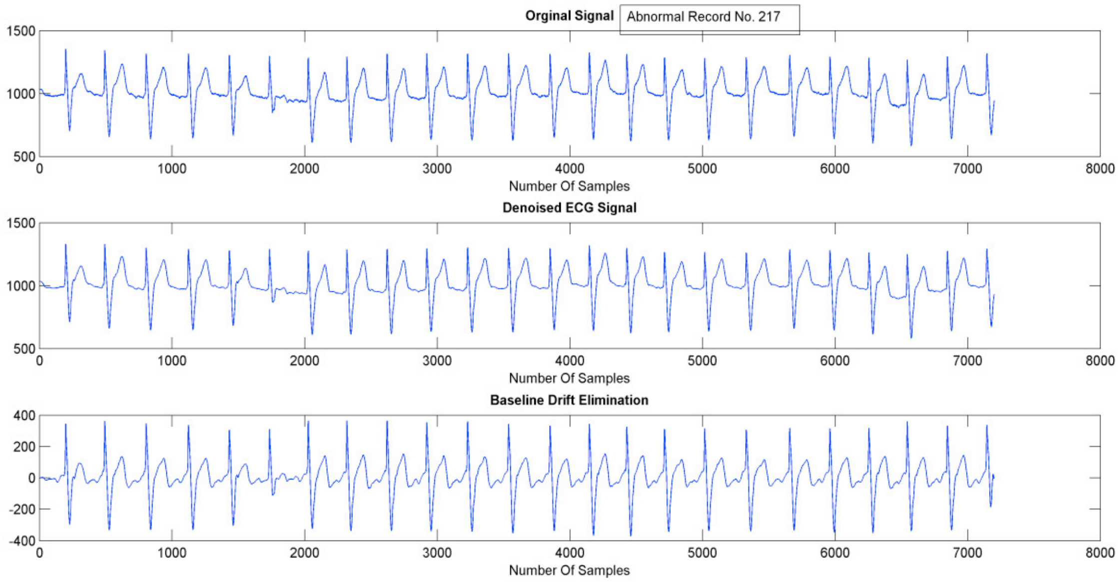


Fig.5. Pre-processed ECG Signal No.217 (Original, Denoised, Baseline Eliminated)

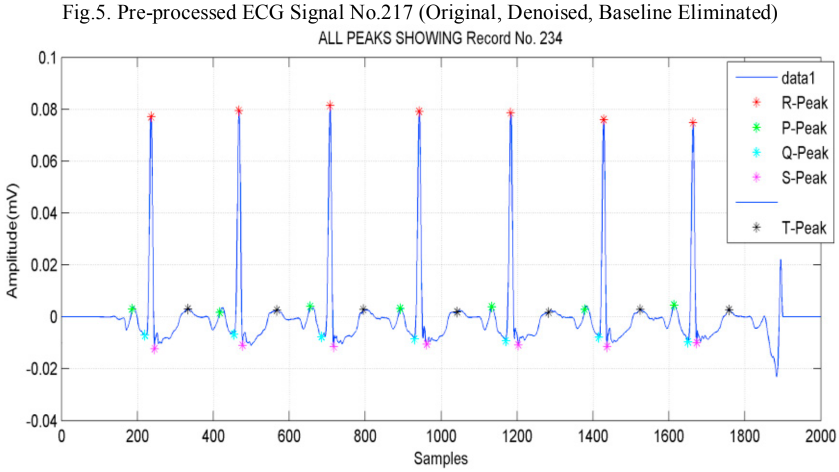


Fig.6. All peaks detected normal ECG signal (record no. 234)

3.6. Wavelet Families

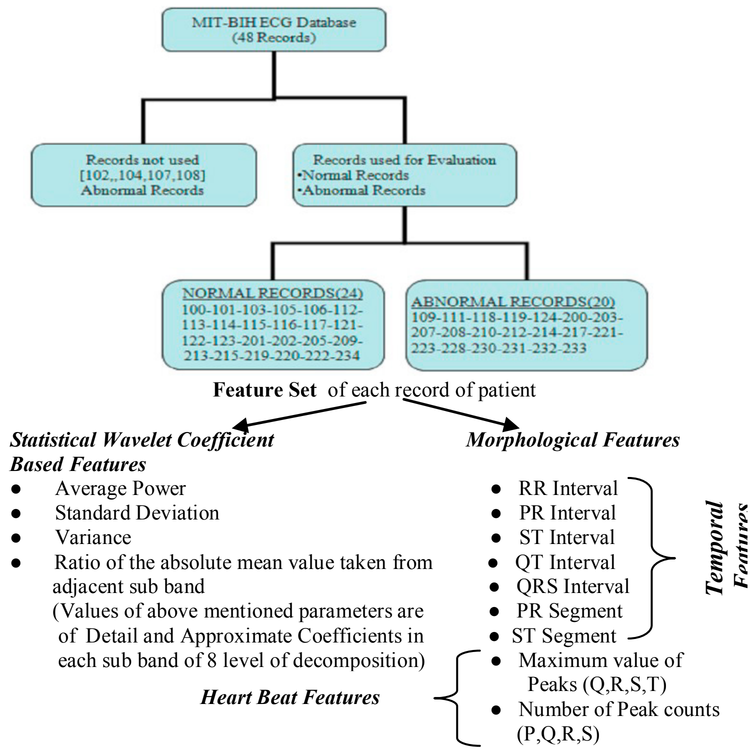
ECG classification into normal and abnormal class is performed with different wavelet families. Daubechie, symlet, biorthogonal, reverse biorthogonal and coiflets are used in proposed method. Orthogonality, smoothness, symmetry, order of filtering, vanishing moments etc. are the distinguishing characteristics of wavelet families [16, 17]. These factors help in determining accuracy, frequency resolution and required decomposing filters.

4. Experimental Results and Analysis

The algorithm used in the evaluation of signals and classification are implemented in Matlab working environment (R2013 B). Dataset used is categorized into normal and abnormal class. The feature set (79×44) is prepared by selecting files listed in Table1. There are total 44 records in which 24 are normal and rest 20 are abnormal record numbers. Table1 shows the extracted feature using WT.

4.1 Signal Detection and Evaluation using Daubechies Wavelet

Table 1.
Representation of extracted features of files of dataset used as input to classifier



A normal ECG signal 103 and an abnormal ECG signal 217 are de-noised using WT. The high and low frequency disturbances are removed i.e. Baseline wander correction is shown in Figs.4, 5. Figure 6 shows all peaks of ECG signal record (234). All required peaks of ECG signal are detected successfully. The processing time involved in feature extraction of 44 files for 1 minute duration is around 140.1 s.

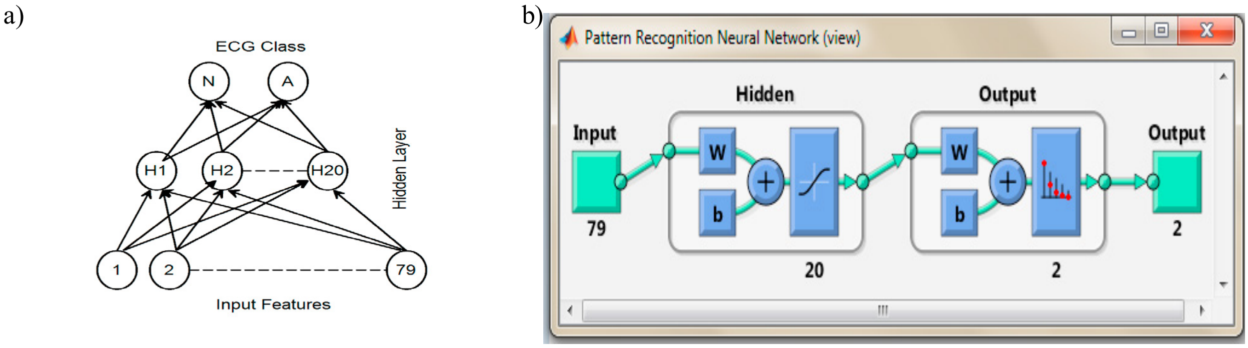


Fig.7. (a) Proposed Neural Architecture; (b) Back Propagation Network

4.2 Classifier Results

For classifier, the input is the feature dataset matrix (79×44) and its output result is compared with pre defined binary targeted values. The selected record number with feature set are randomly assigned into training and testing subset using training data as 70% and testing data as 30% . The simulation has been performed by using BPN, FFN and RBFNN [15]. 20 neurons are considered for hidden layer in each case. Two neurons represent the output layer. Fig. 7. (b) shows a pictorial view of all three layers i.e. network architecture. Table 2 summarizes the detail performance of all three neural classifiers and shows comparison in accuracy, total processing time, sensitivity, specificity and positive predictivity. Fig.8 (a-c) shows Confusion Matrix.

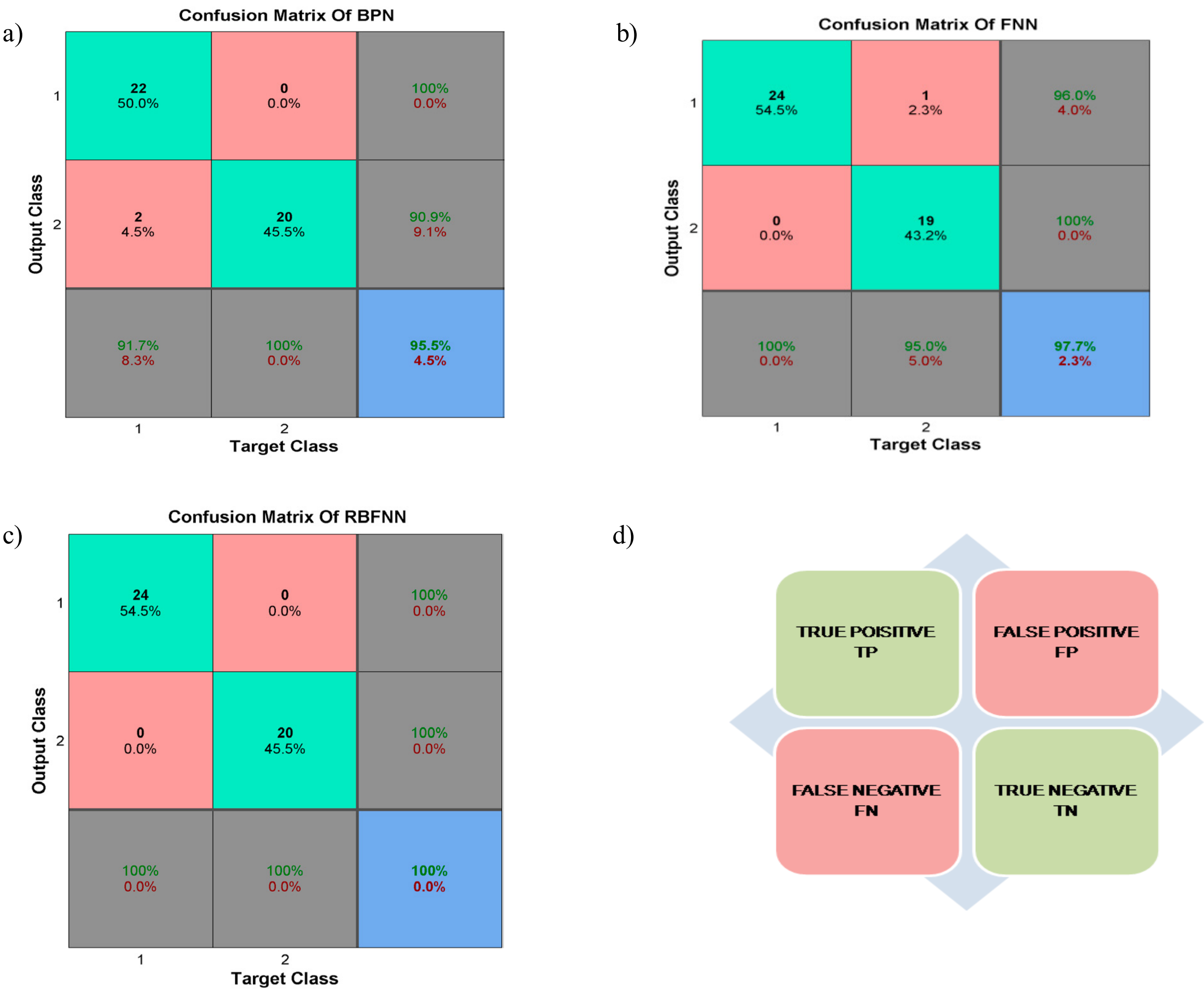


Fig.8. (a-c) Confusion Matrix of ECG signal classifiers; (d) Performance Matrix Evaluation

Table 2. Performance Comparison of ECG signal classifiers

Classifying Methods	Targeted Output (Class)	Normal Class	Abnormal Class	Accuracy (%)	Processing Time (s)	Sensitivity (%)	Positive Predictivity (%)	Specificity (%)
RBFNN	Normal	24	0	**100**	**0.734**	**100**	**100**	**100**
	Abnormal	0	20	**100**				
	Total	24	20	**100**				
BPN	Normal	22	0	91.7	0.991	91.6	100	100
	Abnormal	2	20	100				
	Total	24	20	95.5				
FFN	Normal	24	1	100	3.453	100	96	95
	Abnormal	0	19	95				
	Total	24	20	97.7				

4.3 Comparison Results with other Wavelet families :

The accuracy and processing time taken by different wavelet families are calculated and listed in Table 3. The best performance results are indicated in bold. It is observed that in terms of classification and processing time bior 4.4, rbio 5.5 and sym 4 perform almost similar with db4. These can be promising wavelets for future. Daubechies wavelet family depicts shape like QRS complex and energy concentration at low frequency helps in better analysis.

Table 3. Comparison results of accuracy (%) and total processing time of classifier executed on Different wavelet families

Technique / Wavelet	Order	BPN Acc (%)	BPN T(s)	FNN Acc (%)	FNN T(s)	RBFNN Acc (%)	RBFNN T(s)
Daubechie	**db2**	**95.5**	**1.052**	**100**	**3.485**	**100**	**0.798**
	db4	**95.5**	**0.991**	**97.7**	**3.453**	**100**	**0.734**
	db6	95.5	1.039	88.6	2.739	100	0.788
	db8	95.5	1.058	93.2	2.725	100	0.761
Symlet	sym2	95.5	1.002	100	3.186	100	0.771
	sym4	**95.5**	**1.020**	**100**	**2.695**	**100**	**0.782**
	sym6	95.5	1.003	100	3.211	100	0.801
	sym8	90.9	1.023	95.5	2.691	100	0.805
Bior	bior1.5	95.5	1.075	95.5	2.753	100	0.790
	bior2.6	95.5	0.963	95.5	2.218	100	0.729
	bior4.4	**95.5**	**1.001**	**100**	**3.213**	**100**	**0.702**
	bior5.5	95.5	0.959	100	3.224	100	0.724
Rbio	rbio1.5	95.5	1.010	95.5	3.196	100	0.738
	rbio2.6	95.5	0.982	95.5	3.251	100	0.754
	rbio4.4	95.5	0.960	95.5	3.520	100	0.772
	rbio5.5	**95.5**	**0.908**	**97.7**	**3.221**	**100**	**0.712**
Coif	coif2	93.2	0.931	95.5	2.627	100	0.700

5. Discussion

In this section, comparison of classification performance has been done. Sayadi *et al*. [20] showed 99.1% accuracy in three class classification. Bayesian filtering is used for normal, ventricular premature contraction (VPC) and other beats. Martis *et al*. [21] modeled the ECG signal to classify it into normal and abnormal ECG beats. They reported 98.4% accuracy with SVM-RBF kernel function and Meyer's wavelet function. The authors of reference [8] donot include the record no. 102,107,217 and 233 of MIT-BIH arrhythmia database where as in present method record no. 217 and 233 are included. It is observed that the experimental outcomes of RBFNN gives better results of classification accuracy which is achieved 100 % in proposed method as compared to BPN and FFN which is 97.8 % in both the cases. [8].

6. Conclusion and Future Scope

Three types of neural classifiers BPNN, FFN and RBFNN are used to categorize the ECG signals into normal and abnormal class based on the beat counts present in the signals are examined in this study. The statistical and morphological features are extracted after removal of noise using wavelet transform with different wavelet families and it is observed that daubechies wavelet gives best results among all wavelet family. From the experimental results, it is concluded that RBFNN outperforms FFN and BPN with sensitivity 100, positive predictivity 100, specificity 100, and accuracy 100%. This system performance shows that daubechies, symlet, biorthogonal and reverse biorthogonal wavelets prove to be efficient with RBFNN as neural classifier which is accurate and high speed neural network for ECG signal classification. In future, this proposed arrhythmia monitoring support system can be designed using Wavelet Packet Transformation for better time frequency resolution. The proposal presented in this paper can be extended for multiclass classification including diverse datasets. Furthermore, an efficient ECG classifier can be proposed using robust feature extraction through dimensionality reduction methods.

References

[1] Benjamin, Emelia J., et al. "Heart disease and stroke statistics—2017 update: a report from the American Heart Associatio n." *Circulation* 13 5.10 (2017): e146-e603.

[2] Yuksel Ozbay, Rahime Ceylan, Bekir Karlik (2011), "Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier*", Expert Systems with Application*s, Vol. 38, pp. 1004–1010.

[3] R.J. Martis, U.R. Acharya, C.M. Lim, K. Mandana, A.K. Ray, C. Chakraborty(2013), "Application of higher order cumulant featur es for cardiac health diagnosis using ECG signals*", Int. J. Neural Syst.* 23(04) .

[4] P. de Chazal, B. G. Celler, R. B. Rei (2000), "Using Wavelet Coefficients for the Classification of the Electrocardiogram", *Proceedings of The 22nd Annual EMBS International Conference*, Chicago IL.

[5] Hafizah Hussain, Lai Len Fatt(2007) , "Efficient ECG Signal Classification Using Sparsely Connected Radial Basis Function Neural Network", *Proceeding of the 6th WSEAS International Conference on Circuits, Systems, Electronics, Control and Signal Processing*, pp. 412- 416.

[6] R.J. Martis, U.R. Acharya, L.C. Min(2013), "ECG beat classification using PCA, LDA, ICA and discrete wavelet transform", *Biomed. Signal Process. Control* 8 (5) 437–448.

[7] F.Melgani, Y.Bazi (2008), "Classification of electrocardiogram signals with support vector machines and particle swarm optimization", *IEEE Transactions on Information Technology in Biomedicine*, 12(5), 667–677.

[8] H.M. Rai, et al. (2013), "ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier" *Measurement* 46(9), 3238–3246.

[9] Physiobank Archive Index, MIT-BIH Arrhythmia Database,http://www.physionet.org/physiobank/database.

[10] Chen, Shanshan, et al. (2017), "Heartbeat classification using projected and dynamic features of ECG signal." *Biomedical Signal Processing and Control* 31: 165-173.

[11] Sahoo, Santanu, et al. (2017),"Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities." *Measurement* 108: 55-66.

[12] J. Pan, W.J. Tompkins (1985), "A real-time QRS detection algorithm", *IEEE Trans.Biomed. Eng.* 32 (3) (1985) 230–236.

[13] Hari Mohan Rai, Anurag Trivedi (2012), "De-noising of ECG waveforms using multiresolution wavelet transform", *intenational journal of computer application,* vol. 45, issue 18, pp. 55-60.

[14] S. Banerjee, R. Gupta, M. Mitra (2012),"Delineation of ECG characteristic features using multiresolution wavelet analysis method", *Measurement* 45 (3) 474–487.

[15] Neural Network Toolbox. Available: http://www.mathworks.com/

[16] Soman, K. I., & Ramachandran (2004). "Insight into Wavelets from theory to practice", Prentice-Hall of India

[17] Daubechies, (1992)" Ingrid. Ten lectures on Wavelets." Vol. 61. Philadelphia: Society for industrial and applied mathematics.

[18] Dokur, Zümray, and Tamer Ölmez. (2001) "ECG beat classification by a novel hybrid neural network." *Computer methods and programs in biomedicine* 66.2: 167-181.

[19] RBFN Tutorial by Chris McCormick.https://chrisjmccormick.wordpress.com/2013/08/15/radial-basis-function-network-rbfn-tutorial/

[20] Sayadi, O., Shamsollahi M. B., Clifford G. D. (2010)," Robust Detection of Premature Ventricular Contractions Using a Wave-Based Bayesian Framewor*k," IEEE Transactions on Biomedical Engineering*, Vol. 57, Issue 2, pp 353-362

[21] R.J. Martis, U.R. Acharya, A.K. Ray, C. Chakraborty(2011), "Application of higher order cumulants to ECG signals for the cardiac health diagnosis", *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pp. 1697–1700.