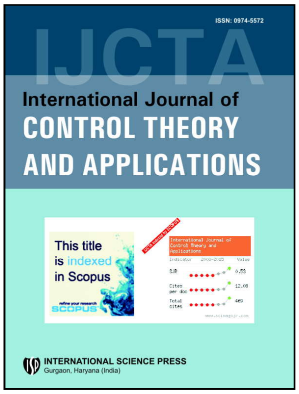

International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 18 • 2017

ECG signal processing using Soft Computing Techniques: A Literature Review

Mangesh Kose[1], Rekh Ram Janghel[2] and Kuhu Gupta[3]

[1,2,3] Department of Information Technology NIT Raipur, Raipur, Chhattisgarh, India, Emails: Mangeshkose@gail.com, rrjanghel.it@nitrr.ac.in, kuhu.gupta08@gmail.com

Abstract: Heart diseases cause death of more peoples than any other. Heart diseases are scientifically referred as cardiac arrhythmia, which differs from normal heart condition. It is very important to analyze the Electrocardiogram (ECG) signal and identify whether the ECG signal is normal or contains some arrhythmia in it. ECG signals are non-stationary in nature again ECG signal for same person can be different at different time. It is tedious job to analyze the ECG signal manually and classify it. An automated method to analyze ECG signal and classify the signal according to type of arrhythmia it contains. There are many computerized methods available for automated feature extraction and classification of the given ECG signal. Feature extraction methods include Wavelet Transform (WT), Principal Component Analysis (PCA) etc. Classification methods available in literature for ECG signal classification includes Support Vector Machine (SVM), Multilayer Perceptron (MLP), Random Forest (RF).

Keywords: Cardiac arrhythmia; ECG; Wavelet Transform; PCA; MLP; RF; ANFIS.

1. INTRODUCTION

Electrocardiography (ECG) is one of the branches of biological science which deals with analyzing and recording behavior of the heart. Electrocardiogram is the device through which the signals of the heart are analyzed. The Electrocardiographic signal also abbreviated as ECG signal is the reflector of status of the heart [1]. Heart is one of the most important parts of the human body, where slight abnormal behavior can cause very high danger to the human life. The abnormal behavior of heart is scientifically referred to as Cardiac Arrhythmia [2].The ECG signal consist of 5 waves i.e. P,Q,R,S,T repeating itself for each heart beat as shown in figure1.

The most important part in ECG signal is QRS complex whose structure identifies the type of abnormality the signal having. Again various other characteristics such as R-R interval, QT interval, ST interval etc. are useful in analyzing the ECG signal.In normal ECG signal theThe computerised analysis of the ECG signal is an important devlopement In the field of Electrocardiography,which also avoided the tedious job of analyzing the large volume of ECG signal manually.Again the manual analysis has the possibility of missing some important information of the ECG signal.Using computerized method of ECG signal analysis it becomes very easy to

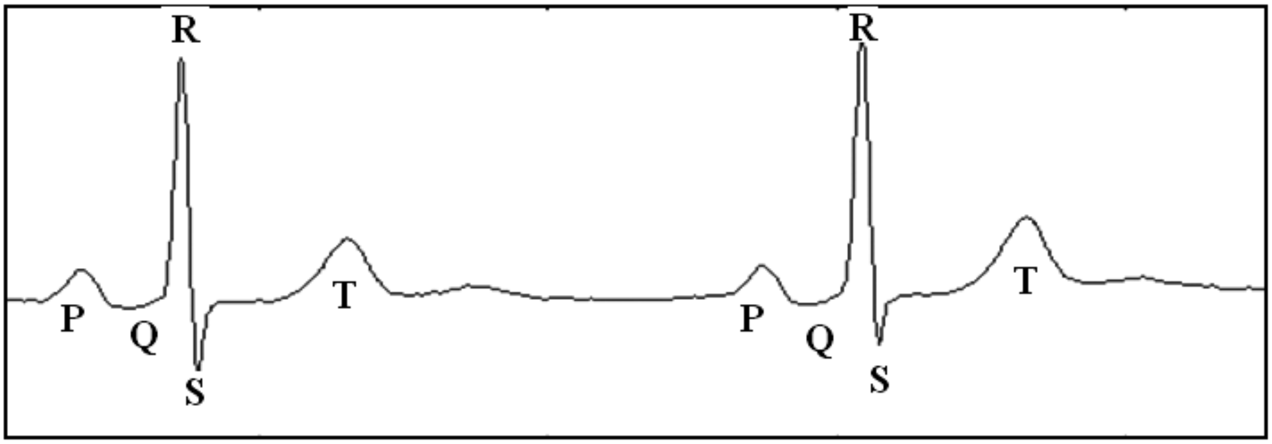

Figure 1: Normal ECG signal beats

classify the signal according to their characteristics. There are large number of techniques availabel in literature to classify the ECG signal into particular class[3].

ECG signal classfication process consist of three stages : a signal preprocessing stage, feature extraction stage and Classification stage. ECG signal contains some artifacts or noise in it. A preprocessing stage is usefull to eliminate the noise from given signal so that we get a pure ECG signal for further processing.There are large number of signal denoising techniques are available in literature.The next stage after signal denoising is Feature Extraction. In this stage the Morphological and Statistical features of the signal are extracted.These extracted features then form the feature vector for the signal classification process.The morphological feature includes: R-R interval,ST interval,width and height of the QRS comlex etc.and statistical features includes:Minimum, Maximum,Mean, Median etc. value of the absolute ECG signal.There are various techniques available in literature for feature extraction[4]. Classification stage uses the feature vector obtained from feature extraction stage to correctly classify the given ECG signal. A large number of classifiers are available in the literature for ECG signal classification purpose.These Feature Extraction and Classification methods are explained in the following sections[5].

2. FEATURE EXTRACTION METHODS

Feature Extraction is an important step to be performed before performing signal classification. There are multiple methods available in literature for feature extraction of the ECG signal as follows:

2.1. Wavelet Transform

It is one of the widely used methods for the feature extraction purpose. As we know ECG signals are Non-stationary in nature, wavelet transform is an effective process to analyze the non-stationary signals. Wavelet transform allows us to choose various time-frequency characteristics of the signal as it is multi-resolution analysis method. There are basically two broad categories of Wavelet Transform, Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT) [5].

The mathematical representation of CWT is as follows

$$\left(W_{\psi} f\right)(a,b) = |a|^{1/2} \int_{-\infty}^{+\infty} f(t)\psi^{*}\left(\frac{t-b}{a}\right)dt = f * \psi_{a}^{*}(b) \tag{1}$$

Where, $\psi_{a,b}(t)$ is a single prototype function called the wavelet mother function, $\psi_{a,b}^{*}(t)$ is the complex conjugate of the analyzing wavelet function $\psi_{a,b}(t)$, a, is the dilation parameter of the wavelet and b is the translation parameter of the wavelet.

DWT decomposes the discrete time signal into lower resolution approximation signal and their associated detail signals .

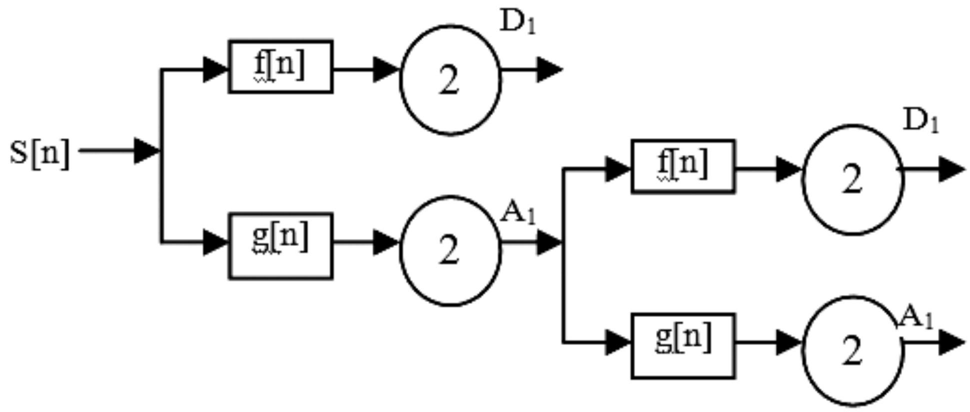

Figure 2: Decomposition of signal S[n] into approximation (A_i) and detail (D_i) signal. F[n] is high pass filter and g[n] low pass filter

2.2. Principal Component Analysis (PCA)

PCA is well known technique for feature extraction which based on belief that more information about particular class is contained in the direction of maximum variability. It is the statistical procedure uses orthogonal transformation to convert set of correlated variables to the principal component. For P number of variables N numbers of principal components are obtained where N<=P. The transformation is occurred in such a way that the maximum variable component is considered as first principal component and each successive component in turn has highest variance with the restriction that it must be orthogonal to preceding components[6].

In addition to feature extraction, PCA performs feature reduction in which the number of features required to perform the classification process is reduced without affecting the accuracy of the classification. As the large size of data is not only problem for system hardware, but also it affects the performance of machine learning algorithms. In such situations it is an important and popular tool for the feature reduction process.

Steps of PCA algorithm

1. Standardization of the Dataset.
2. Calculate the Eigenvectors and corresponding Eigen values from the covariance or correlation matrix or obtain the singular vector decomposition.

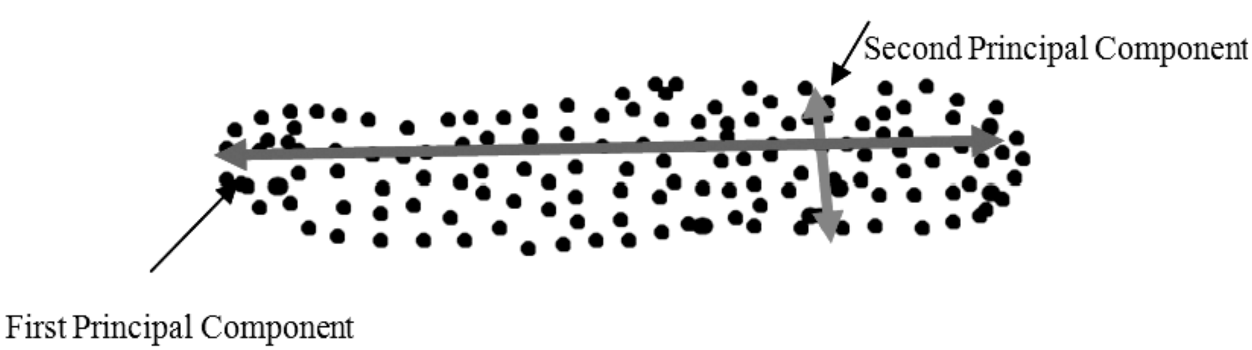

Figure 3: Principal Component representation

3. Choose the 'N' Eigenvectors corresponding to 'N' largest Eigen values where 'N' will be the number of dimension of new feature space.
4. Obtain the projection matrix 'P' from these 'N' Eigenvectors.
5. Perform the transform of original feature vector 'X' via 'P' to obtained new feature vector 'Y' with 'N' dimensional datasets.

3. CLASSIFICATION METHODS

After completion of Feature extraction methods the user will get the feature vector. This feature vector is use to classify the given ECG signal into proper class. These classification methods have high importance in pattern recognition applications to identify whether the given pattern belongs to given class or not depending on its feature vector.

3.1. Support Vector Machine (SVM)

Support Vector Machine is most popular classification technique. It uses the concept of supervised learning. Initially SVM was developed only for the binary classification purpose but later it is used for multiclass classification as well. The working principal of SVM includes creation of hyper planes in multidimensional space. SVM creates two parallel hyper planes on both sides of data searching hyper planes. Hyper plane which is used as separator is chosen in such a way that distance between parallel hyper planes will be maximizing [3].

As shown in figure 4, middle hyper plane is Separator hyper plane which separates different classes from each other. The elements represented by circle belong to class1 and elements represented by square are belongs to class2.Misclassified elements from class1 and class2 are represented by M1 and M2 respectively. Accuracy of classification in this method is depending only on the elements which are near to the hyper planes. These elements near to the hyper planes are called support vectors. As this classification method depends on support vectors this classification technique is called as Support Vector Machine.In this method the dataset is divided as

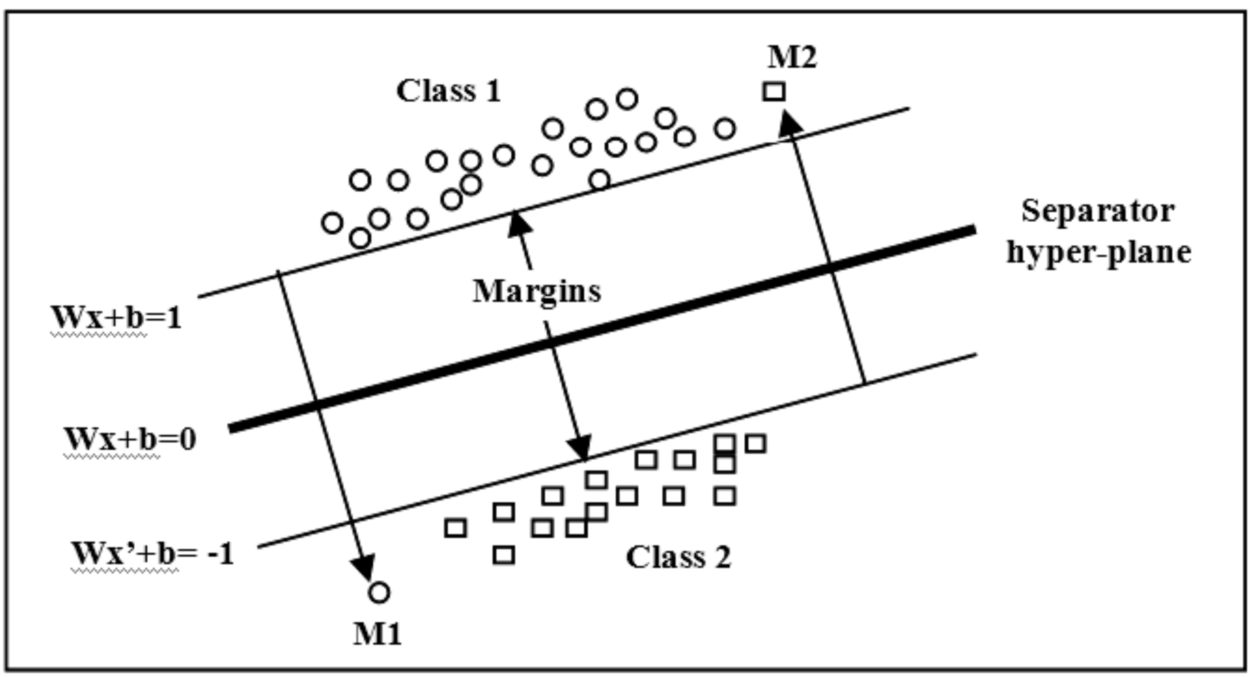

Figure 4: Representation of Hyper planes

Training set and Testing set. Training set consist of $\{x_{i;} y_i\}$, i=1 …i, where x_i is feature vector and y_i is class label corresponding to that feature vector. The SVM has to solve the following optimization problem[13].

$$\min_{w,b,\S} \frac{1}{2} w^T w + c \sum_{i=1}^{l} \S i \tag{2}$$

$$\text{Subject to } y_i(w^T ø(x_i)+b) \geq 1\text{- } \S_i \tag{3}$$

$$\S_i \geq 0$$

3.2. Multi Layer Perceptron (MLP)

A Multilayer Perceptron is neural network based classification method .MLP is a feed forward structure that produce output using given input values.MLP consist of multiple layers of neurons. Every layer of neurons is fully connected to next layer neurons. It is supervised learning based technique which needs actual target value for learning purpose. Structure of the MLP network is presented in figure 4. Many users do not consider the input layer as the layer of the neural network as no computation is performed at that point. But in this study we considered input layer as one of the layer of neural network [16].

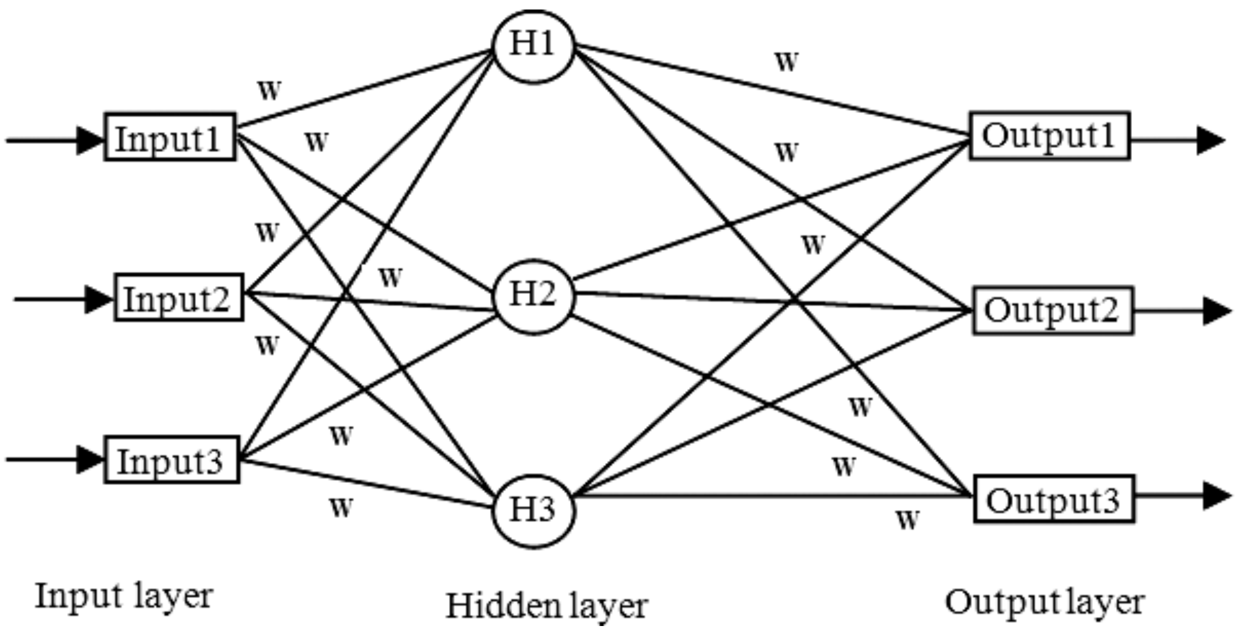

Figure 5: The general structure of MLP

MLP consist of more than two layers where each neuron from one layer is connected to the neurons in the next layer with certain weight 'w'. MLP network contains single input and output layer but the number of hidden layers can be more than one. The above example contains only single hidden layer represented by neurons (H1, H2, H3) . MLP network performs learning by changing weight after processing of piece of data. Learning is performed in MLP by using Back Propagation algorithm. In this algorithm error is calculated by using the actual target value and value produced by the network at output layer. According to this error value the correction is made in the weights so as to minimize the value of the error calculated. So repeating this process of updating weights over certain time the final value of error will be minimized.

Back Propagation algorithm for MLP consists of following steps:-

1. Initialization: - In this step we have to set the values of weights and biases to small real values.

2. Providing input and desired output values:- Input vector x(1), x(2),…….,x(N) and corresponding desired output d(1), d(2), ……d(N) is provided to the MLPNN.

   Calculation of actual Output:- Equation used to calculate actual output is as follows:-

$$y_i = \varphi\left(\sum_{j=1}^{N_{M-1}} w_{ij}^{(M-1)} x_j^{(M-1)} + b_i^{(M-1)}\right),\ i = 1,\ldots,NM-1 \tag{4}$$

3. Adaptation of weights (wij) and biases(bi):-

$$\Delta w_{ij}^{(l-1)}(n) = \mu . x_j(n) . \delta_i^{l-1}(n) \tag{5}$$

$$\Delta b_i^{l-1}(n) = \mu . \delta_i^{(l-1)}(n) \tag{6}$$

   Where

$$\delta_i^{(l-1)}(n) = \begin{Bmatrix} \varphi'\left(net_i^{(l-1)}\right)\left[d_i - y_i(n)\right], \\ \varphi'\left(net_i^{(l-1)}\right)\sum_k w_{ki} . \delta_k^{(l)}(n), \end{Bmatrix} \begin{matrix} 1 = M \\ 1 \le 1 \le M \end{matrix} \tag{7}$$

In which $x_j(n)$ is output at iteration n at j^{th} node, k is the total number of output nodes of neural network, M is output layer, and the activation function is ∅. The rate of learning is represented by μ[1].

3.3. Random Forest (RF)

These are an ensemble learning method used as a classifier and other tasks. The RF method creates multiple Classification Trees which themselves used as classifiers. In classification process using RF method, each classification Tree which is constituent of the whole Random Forest produces output i.e. in the classification process class label for the feature values. It is possible that different classification tree in a RF give different class label for same feature values. The final output for the feature i.e. class label is obtained by using vote system. In this process each classification tree will vote for the class that it given as output. Final output for given feature will be the class having highest vote.

In short RF is the method of averaging the decision trees which trained on different part of training set. In this way variance of model is decreased. Random Forest is not sensitive to the noise. RF recursively built the classification trees by splitting the node. A split of node is made in such a way that reduces the uncertainty in the data and hence the chances of misclassification using the Gini's criteria of node impurity. The process of splitting of node is repeated until the "forest" is created as a collection of multiple trees.

RF algorithm can be stated as follows:-

1. Take a number n such that $n \sim \sqrt{N}$ where N is total number of variables.
2. Each classification tree of maximum level is grown without pruning.
3. At each node, n variables are selected randomly.
4. Then by using Gini node impurity criteria best split on these n variables is determined.

Reducing value of n reduces the strength of each classification tree and vice versa. The complexity of computation for each tree in RF is $\sqrt{N}$ Slog(S); having S number of training cases [15]. The other advantages of RF method is that it works very fast, the performance of the method is excellent, cross validation is not needed and useful for long term ECG classification purpose.

4. RESULT AND DISCUSSION

4.1. Result

The result after application of classifiers on the Dataset describe in section 4.2 is shown in table in terms of percentage of Accuracy in classifying the ECG signal.

Table 1
Accuracy of classifiers

Classifiers	*Accuracy (%)*
SVM	92.16
MLP	88.23
RF	98.8

Among the classification methods shown in Table I, Random Forest has highest accuracy percentage than all the other classifiers. Random Forest is the ensemble learning method in which it builds multiple classification trees. Random Forest is the more time complex classifier than SVM. Support Vector Machine is less complex method and performs fast execution. The Random Forest classifier must be used in the applications where accuracy is main concern. Support Vector Machine is less complex classifier which takes less time to get executed than RF. MLP is feed forward neural network structure with multiple layers. MLP classifier is more efficient for multiclass classification problems.

4.2. Dataset

The MIT-BIH Arrhythmia dataset downloaded from physionet.org is used in this study. This database contains 48 recordings. Each recording has duration of 30 min. The recording is made using two electrodes MLII and one of the modified leads V1, V2, V4, or V5. The dataset has sampling frequency of 360 Hz. Signal recording is divided into two parts in which first part containing 23 recordings and second part consist of 25 recordings.

This database provides three files for each recording. These files include text header file, a binary annotation file and a binary file. The text header file contains all the information related to the signal, such as sampling frequency, number of samples, number of leads, type of ECG leads, clinical information of the patient and format of the signal to be downloaded. Binary final contains the original ECG signal that is to be used for processing. Binary annotation file consists of beat annotation. The signal contains some noise in it. Preprocessing is necessary in order to get the noise free signal. Performing operation on noise free signal provides efficient results.

The types of arrhythmias occur in the ECG signals are as follows

Table 2
ECG Signal arrhythmias

Symbol	*Meaning*	*Characteristics*
N	Normal Beat	Regular occurrence of each wave without any fluctuation in width or phase.
P	Paced beats	Wave deformation based on paving of ventrium and atrium.
LBBB	Left Bundle Branch Block	R wave takes shape of alphabet “M
RBBB	Right Bundle Branch Block	Wide S wave, delayed activation of right ventrium produces a slurred.
AFIB	Atrial fibrillation	P wave vanishes
VFL	Ventricular flutter	Whole PQRST cycle gets deformed and mixed with each other.
PREX	Pre-excitation beat	The initial portion of QRS complex takes shape of delta wave.
SBR	Sinus Bradicardia	Heart rate becomes lower than normal i.e. 60 beats per min.
PVC	Premature ventricular Contraction	QRS complex gets altered.
F	Fusion beats	Normal and PVC beats combined to get fusion beat.

5. CONCLUSION

In this paper we have presented thorough study on the signal processing methodologies. These methodologies used for the purpose of feature extraction and classification of the given ECG signal. In the feature extraction methods, Wavelet transform is most popular and widely used method for feature extraction. It represents the given signal in multiple domain i.e. time and frequency domain. It means, wavelet not only specifies that whether an element is present or not but also the location of element. Wavelet can be efficiently used for the extraction of Statistical as well as morphological features. Other feature extraction methods are also used depending on the need of the algorithm. For example, PCA algorithm will be more useful in the situation where, in addition to feature extraction, feature reduction is also needed. Among the classification methods shown in Table I, Random Forest has highest accuracy percentage than all the other classifiers. Random Forest is the ensemble learning method in which it builds multiple classification trees. Random Forest is the more time complex classifier than SVM. Support Vector Machine is less complex method and performs fast execution.

After deep study of the soft computing techniques for feature extraction and classification of ECG signal following research gaps found. Classification accuracy of the signal depends on the feature vector obtained from that signal. ECG signals are non-stationary in nature; the adaptive method for the feature extraction from the signal is an important challenge in front of researchers. We are working on the method of feature extraction using adaptive approach. Accuracy is the key factor in the ECG signal classification process. Certain methods provide great accuracy but require complex calculation and long time to get executed. So, the research in order to develop a classifier in combination with optimization method to reduce Miss Classification can be developed. Most of the classifiers work efficiently for binary classification meanwhile for multiclass classification they are not as much efficient. Efforts in order to develop an efficient method for multiclass classification are needed.

REFERENCES

[1] H. M. Rai, "ECG Signal Classification using Wavelet Transform and Back Propagation Neural Network," vol. 3, pp. 212–215.

[2] T. M. Nazmy, H. El-Messiry, and B. Al-Bokhity, "Adaptive Neuro-Fuzzy Inference System for classification of ECG signals," *Informatics Syst. (INFOS), 2010 7th Int. Conf.*, pp. 1–6, 2013.

[3] E. D. Übeyli, "ECG beats classification using multiclass support vector machines with error correcting output codes," vol. 17, pp. 675–684, 2007.

[4] E. Alickovic and Z. Babic, "The effect of denoising on classification of ECG signals," *2015 25th Int. Conf. Information, Commun. Autom. Technol. ICAT 2015 - Proc.*, 2015

[5] S. A. Shufni, "Wavelet Transform , Time Domain and Frequency Domain Features," no. March, pp. 30–31, 2015.

[6] R. Ceylan and Y. Özbay, "Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network," *Expert Syst. Appl.*, vol. 33, no. 2, pp. 286–295, 2007.

[7] Barhatte, Alka, and Rajesh Ghongade. "A multiclass cardiac events classifier using clustering and modified adaptive neuro-fuzzy inference system." *2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*. IEEE, 2015.

[8] P. De Chazal, B. G. Celler, and M. Zeee, "selectin $ a neural network structure for ECG diagnosis," vol. 20, no. 3, pp. 1422–1425, 1998.

[9] N. García-Pedrajas and D. Ortiz-Boyer, "A cooperative constructive method for neural networks for pattern recognition," *Pattern Recognit.*, vol. 40, no. 1, pp. 80–98, 2007.

[10] Emanet, Nahit. "ECG beat classification by using discrete wavelet transform and Random Forest algorithm."*Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control, 2009. ICSCCW 2009. Fifth International Conference on*. IEEE, 2009.

[11] M. K. Gautam, "A Neural Network approach and Wavelet analysis for ECG classification," no. March, pp. 0–5, 2016.

[12] A. Khazaee and A. Ebrahimzadeh, "Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features," *Biomed. Signal Process. Control*, vol. 5, no. 4, pp. 252–263, 2010.

[13] S. Kumar, V. K. Sharma, and R. Kumari, "A Novel Hybrid Crossover based Artificial Bee Colony Algorithm for Optimization Problem," *Int. J. Comput. Appl.*, vol. 82, no. 8, pp. 18–25, 2013.

[14] J. A. Nasiri, M. Naghibzadeh, H. S. Yazdi, and B. Naghibzadeh, "ECG arrhythmia classification with support vector machines and genetic algorithm," *EM.*

[15] Emanet, Nahit. "ECG beat classification by using discrete wavelet transform and Random Forest algorithm."*Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control, 2009. ICSCCW 2009. Fifth International Conference on*. IEEE, 2009.

[16] Güler, Ýnan, and Elif Derya Übeyli. "A modified mixture of experts network structure for ECG beats classification with diverse features." *Engineering Applications of Artificial Intelligence* 18.7 (2005): 845-856.

[17] Y. Özbay and G. Tezel, "A new method for classification of ECG arrhythmias using neural network with adaptive activation function," *Digit. Signal Process.*, vol. 20, no. 4, pp. 1040–1049, 2010.

[18] I. S. S. Rao and T. S. Rao, "a Comparitive Study on Classification of Ecg Using," vol. 6, no. 2, pp. 1–5, 2016.

[19] M. Srinivasulu and K. C. Reddy, "Accuracy Assessment for Multi-Channel ECG Waveforms Using Soft Computing Methodologies," vol. 4, no. 7, pp. 41–49, 2014.

[20] E. D. Übeyli, "ECG beats classification using multiclass support vector machines with error correcting output codes," vol. 17, pp. 675–684, 2007.

[21] L. Zhang, "An Approach for ECG Classification Based on Wavelet Feature Extraction and Decision Tree," pp. 1–4, 2010.

[22] S. Osowski and L. T. Hoai, "Analysis of features for efficient ECG signal classification using neuro-fuzzy network,"*IEEE Int. Conf. Neural Networks - Conf. Proc.*, vol. 3, pp. 2443–2448, 2004.

[23] G. Soto and S. Torres, "Feature Extraction Based on Circular Summary Statistics in ECG Signal Classification," *Chil. Comput. Sci. Soc. (SCCC), 2009 Int. Conf.*, no. 2, pp. 2–4, 2009.