ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645
DOI:10.17485/ijst/2015/v8i30/74576,

Indian Journal of Science and Technology, *Vol 8(30), DOI: 10.17485/ijst/2015/v8i30/74576, November 2015*

An Effective Classification of Heart Rate Data using PSO-FCM Clustering and Enhanced Support Vector Machine

R. Kavitha[1*] and T. Christopher[2]

[1]Department of Computer Science, PSGR Krishnammal College for Women, Coimbatore - 641004, Tamil Nadu, India; rkavitha2509@gmail.com
[2]Department of Research and PG Computer Science, Government Arts College, Coimbatore - 641018, Tamil Nadu, India; chris.hodcs@gmail.com

Abstract

Background/Objectives: Heart Rate Variability is an essential feature which decides the condition of human heart. ECG is used as diagnostic tool to access the electrical function of the heart. **Methods/Statistical Analysis:** The nine linear and nonlinear features are derived from the HRV signals. The feature extraction is carried out with the help of Particle Swarm Optimization (PSO) for data reduction. In proposed scheme Fuzzy C-Means (FCM) clustering and classifier integrated to enhance the accuracy result for ECG beat classification. **Findings:** The Enhanced SVM classifier classifies the heart rate data. Enhanced SVM classifier groups the linear and non-linear parameters as inputs, which are derived from the HRV signal. The denoise signals are classified and identifies the pattern for better classification of ECG signal. **Application/Improvements:** The proposed scheme is experimented with the assistance of the most commonly used MIT-BIH arrhythmia database and adequate results were obtained with an accuracy level of 98.38% than the other well-known approaches.

Keywords: Classification, EGC, ESVM, FCM, HRV, PSO

1. Introduction

Heart Rate Variability (HRV) is mainly utilized for the purpose of measuring the variations in heart signals. To be more exact, it is the variations per unit time of the amount of heartbeats. The most common technique used to detect the heart beats is Electrocardiography (ECG). The Electrocardiogram is a kind of electrical action of heart and generates electrical signals which are called as PQRST waves. In this wave, the most vital wave is QRS complex. A single heart beat is given as the time gap in the middle of two QRS complex waves which is also known as R-R interval[1,2]. Typically, normal healthy persons have considerably larger values of HRV. Since the heart disease is more common in last few decades, prediction of HRV examination is one of the most common areas of research. Through serious research, several techniques have been formulated for effective HRV analysis. Following are the top most available techniques at present, spectral methods based on FFT, nonlinear scheme, together with Markov chain model[3,4]. In general, neural networks and machine learning schemes are one of the dominant schemes to categorize and forecast the HRV patterns. Several investigations have been done by means of different classification approaches like Support Vector Machine and Neural Network etc. The author's in[5], exploited independent components analysis for the purpose of ECG detection. In case of[6], the authors utilized the temporal characteristics. The authors in[7,8], employed wavelet transform and power spectral density

---

* *Author for correspondence*

of ECG signals, in the same way. In case of[9], a scheme based on image was formulated to obtain discriminative details from the trajectories of ECG in the state space. The scheme discussed in[10] used SVM in order to classify the heart beats. In case of[11], the Lyapunov exponents are used as their characteristics and trained a recurrent neural network for the purpose of separating the four categories of beats. In[12], numerous classification systems are investigated in accordance with the linear discriminant classifiers. While considering[13], they employed Hermit functions and Self-Organizing Maps (SOM) for the purpose of clustering of ECG signals. In case of[14], a power spectral based scheme was formulated for the purpose of classifying the cardiac arrhythmias. In[15], an automatic progression of online beat segmentation and classification system by means of a Markovian scheme was developed. In case of[16], a Gaussian mixture model based classifier was developed in order to classify ECG after the registration of ECG signal by means of Pan-Tompkins approach. The classification technique depending on the principle components derived from the ECG features by means of Principal Component Analysis (PCA). In case of[17], a rule-based rough-set decision approach for the assembly of an inference engine for the purpose of effective disease recognition. A patient-adapting heart beat classifier scheme is formulated in[18] with the help of linear discriminants. The authorsin[19] used several classifiers like Support Vector Machine (SVM) and MLP neural networks for the purpose of classification. In[26] the author used PSO hybrid with Fuzzy C-Means for testing and evaluation.

In this research work, Enhanced Support Vector Machine (ESVM) is employed for the purpose of classification of HRV. SVM is one of the most common techniques in the field ofpattern recognition and data mining. Moreover, the next version of SVM, ESVM is very activein several fields, due to the feature of hyper plane with extensive margin that divides the sample into two classes and to kernel techniques for non-linear classifications. In this scheme, subsequent to feature extraction, the selection of best features is carried out with the help of Particle Swarm Optimization (PSO) which is utilized to diminish the dimension of features in heartbeat classification with clustering by means of FCM. The proposed work is shown in Figure 1. This paper is organized as follows. Section 2 describes the preprocessing and feature extraction along with feature selection process. Section 3 describes the ESVM system for HRV classification. Section 5 describes the experimental design, including the data description and experiment settings. Some of the evaluation results of this system are given. Finally, Section 5 concludes the paper.

2. Proposed Methodology

With the effective studyof the HRV and its bondingwith other physiological signals can be determined without any difficulty. The following schemes are used for the purpose of quantifying the HRV: Time domain, spectral or frequency domain, geometric and nonlinear.

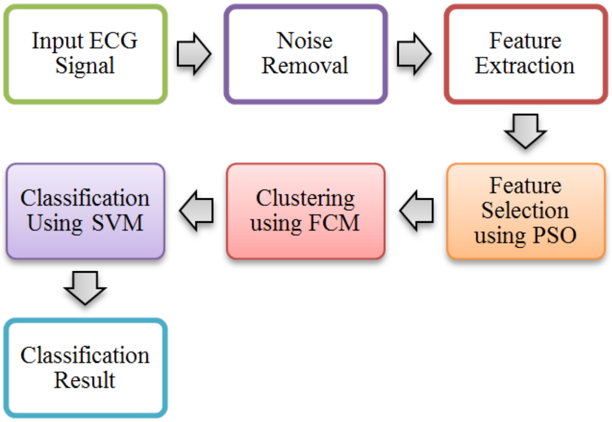

Figure 1. Architecture diagram of the proposed work.

2.1 Preprocessing the Signal

In this research work, HRV datais employed which is produced from the ECG signals given by the MIT-BIH database. In the beginning, it is essential to derive the HRV signals from the ECG. In most scenarios, a small number of intrusive signals close to the mains 50 Hz, the electromyogram (EMG) signals and in addition the baseline wandering can have an effect on the extraction procedure. Therefore, these interfering signals are eliminated from the input ECG signal by means of a 5-15 Hz band pass filter. Subsequently, the signal is processed by means ofthre sholding and the resulting signal is recognized as HRV signal.

2.2 Feature Extraction

The subsequent step in the architecture is the most vital, feature extraction step. Normally, the cardiovascular system demonstrates linear and non linear characteristics mutually. Accordingly, a blend of both linear and nonlinear features is considered for the further process.

2.2.1 Linear Analysis

2.2.1.1 Time Domain Features

Following are the four commonly applied time domain constraints of the HRV signal which are unambiguously derived from the RR interval time series:

- STD HR: This feature denotes the standard deviation of instantaneous heart rate in all segments.
- Mean HR: This feature indicates the mean value of the heart rate with in the range of one minute in all segments.
- HRV triangular index: This feature indicates the integral of the histogram(that is to say, Overall amount of RR intervals) to the altitude of the histogram. At this instant, a bin width of 1/128 is taken for account.
- pNN50: This feature indicates the quantity of consecutive dissimilarity of 64 R-R intervals that fluctuates higher than 50 ms, correspondingly, divided by 64.

2.2.1.2 Frequency Domain Features

It must be noted that the time domain constraints are almost exceptionally proficient; moreover they do not have the adequacy to successfully make a distinction in the midst of the sympathetic and parasympathetic contents of the HRV signal. Then, the High-Frequency (HF) band, in the range of 0.2-0.5 Hz, exemplifies the cardiac vagal behaviors similar to Respiratory Sinus Arrhythmia (RSA). In reality, HF elements are taken as the source of parasympathetic activities of the cardiovascular system. On the contrary, the Low-Frequency (LF) band, in the range of 0-0.2Hz, is fully united with the baroreceptor control and is determined through the use of sympathetic systems. In this research work, the Power Spectral Density (PSD) in case of HF and LF bands are computed and the relative quantity of the LF and HF bands power (LF/HF) is taken as the frequency domain feature of the HRV.

2.3 Nonlinear Analysis

In case of non-linear analysis, HRV signal assessment using nonlinear dynamics provides very valuable details for the effective physiological understanding of the heart. Following are the four dissimilar nonlinear constraints of the HRV signal.

- LLE: Largest Lyapunov Exponent gives most beneficial details on the subject of the reliance of system on opening states and a positive Lyapunov exponent confirms the incidence of chaos in the system. With the aim of computing the LLE, a particular point is chosen in the reconstructed stage space of the system and the complete neighbor points located within a predetermined radius are decided. In view of the fact that the system progresses, the mean distances in the midst of the trajectory of the primary point and the trajectories of the close proximity points are computed. After that, the logarithm of these mean values plots in opposition to the time and the slope of the resulting line are taken as LLE. It must be noted that, the embedding dimension and the lag are considered as $m = 10$ and $t = 1$, in the same way. Moreover, the threshold distance e is considered as the mSD, in which SD points out the standard deviation of the RR time series.
- SD1/SD2: Poincare plot is a kind of graphical depiction of the association in the midst of successive RR intervals. It is accomplished through the process of plotting each RR interval (RR (n+1)) as a function of the prior interval (RR (n)) in RR interval time series. Subsequently, this is quantitatively examined through the method of determining the SD of the distances of the time series points from the lines $y = x$ and $y = x + 2RR_m$, in which RR_m indicates the mean of the overall values of RR interval time series. The corresponding values are named as SD1 and SD2 respectively. Here, SD1 indicates the quick beat-to-beat variability, at the same time SD2 indicates the reasonably long-standing inconsistency in the HRV signal. Here, SD1/SD2 is utilized as the primary nonlinear feature which is obtained from the HRV segments.
- D2: Correlation Dimension is a determination of complication of the time series and makes a decision on the smallest quantity of dynamic variables which can model the system.
- SpEn: Spectral Entropy reveals the obstacle of the input time series in the given frequency domain. Massive values of SpEn shows sign of elevated abnormality and lesser values shows sign of standard time series. Here, the Shannon's channel entropy is utilized for the purpose of approximating the spectral entropy as given below:

$$SpEn = -\sum_{f} p_f \log p_f$$

- where p_f represents the PDF value ona specific frequency f. However, the entropy is interpreted as a determinationof uncertainty concerning the event at f. Thus, the entropy can be used as a determination of system complication. The spectral entropy describes the complication of the HRV.

2.4 Feature Selection using PSO

Feature Selection (FS) is a kind of process for the purpose of choosing a subset of the features taking place in the training set and using only this subset as features in the process of classification. Feature selection provides two major functions. Initially, it puts together training and executing a classifier more effective by lessening the size of the efficient features. Then, feature selection typically boosts classification accuracy by removing noisy elements. A noisy element is one that when included to the document representation, augments the classification error on new data. PSO is a type of stochastic optimization procedure initially formulated by Kennedy and Eberhart[20] that initiated from the simulation of the actions of a collectionof a group of birds or the social activities of a group of people or school of fish. Every bird/individual flies in the region of search space with a speed which is vigorously fine-tuned consistent with its individual flying experience and its companions' flying experience, rather than employing evolutionary operators to operate the individuals like in other evolutionary computational approaches. Every individual is taken as a volume-less particle in the N-dimensional search space. During a m particular time stept, the ith particle is indicated as. $X_i(t) = (x_{i1}(t), x_{i2}(t), \dots, x_{in}(t))$. The collection of positions of particles in a' multidimensional space is found as $X = \{X_1, \dots, X_j, \dots, X_l, \dots, X_m\}$. The finest preceding position (the position provided the best fitness value) of the ith particle is obtained and indicated as $P_i(t) = (p_{i1}, p_{i2}, \dots, p_{iN})$. The index of the best particle in the midst of the entire particles in the population (global model) is denoted by the symbol g. The index of the best particle in the midst ofthe entire particles in a particular topological area (local model) is indicated by the index subscript l. The speed of progress (velocity) for particle i at a particular time step t is given as $V_i(t) = v_{i1}(t), v_{i2}(t), \dots, v_{iN}(t))$. The particle variables are operated in accordance with the equation below (global model):

$$v_{in}(\mathrm{t}) = w_i \times v_{i1}(t\text{-}1) + c_1 \times rand\,1\,() \times (p_{in} - x_{in})(t\text{-}1)) + c_1 \times rand\,1\,() \times p_{in} - x_{in}(t\text{-}1))$$

where n indicates the dimension ($1 \le n \le N$), c_1 and c_2 represents positives constants, rand1() and rand2() represents two random functions in the limit of [0, 1], and w indicates the inertial weight. The inertial weight is linearly declining[21] in proportion to the equation given below:

$$w_t = w_{max} - \frac{w_{max} - w_{min}}{t_{max}} \times t$$

In case of the neighborhood (lbest) model, the only adjustment is to replace p_{in} for p_{gn} in the equation for velocity. Same equation in the global model is employed to compute a particle's new velocity in accordance with its earlier velocity and the distance of its existinglocation from its own best experience (pbest) and the group's best experience (gbest). The local model computation is the same,apart from that the pbestis utilized rather than the gbest. PSO has been employed for approaches that can be exploited across anextensivecollection of applications, in addition tothe particular applications concentrated on anexplicit requirement. Its beauty over several other optimization approaches relies in its comparative simplicity,for the reason that only a small number of parameters require to be fine-tuned.

- **Particle representation:** The particle includes three components: The input features, *C,* and *y,* if the RBF kernel is selected. Table1 exhibits the demonstration of a particle with dimension n+2, where n indicates the total quantity of input features of a specified data set. The range of the n variables is stuck between 0 and 1. In case if the value of a variable is under or equal to 0.5, at that point in time its interrelated feature is not taken into account. In contrast, if the value of a variable isbeyond 0.5, at that moment its related feature is selected.
- **Fitness function:** The fitness function is utilized to assess the superiority of each particle which has to be planned before looking for the best values of both the SVM parameters and the feature selection. Fitness function completely depends on the accuracy of the results of SVM classifier.

Table 1. Representation of a particle with dimension

	Input features	C	y
Particle	$x_1, x_2, \dots, x_n$	x_{n+1}	x_{n+2}

2.5 Fuzzy C-Means Clustering

Structure recognition of fuzzy systems is feasible by building sufficient rules with suitable input and output membership functions. The recognized model can then be employed to illustrate the behavior of the objective system, in addition to prediction use. In this research, the fuzzy layer is trained with the help of the Fuzzy C-Means clustering algorithm. The concept of fuzzy clustering is to segregate the data into fuzzy partitions that extend beyond with one another. As a result, theinsertion of data in a cluster is characterized by a membership grade in [0, 1]. In this scenario, clustering an unlabeled data $X = \{X_1, X_2, \ldots, X_N\} \subset R^h$ where N indicates the number of data vectors and h indicates the dimension of each data vector, is the allocation of c-partition labels to the vectors in X and c-partition of X constitutes sets of $(c.N)$ $\{u_{ij}\}$ membership values that can be suitably organized as a $(c \times N)$ matrix $U = [u_{ij}]$. The fuzzy clustering discovers the optimum membership matrix U. The most extensively employed objective function for fuzzy clustering is the weighted within-groups sum of squarederrors J_m, which is utilized to describe the following constrained optimization complication.

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m}\left\|x_i - c_j\right\|$$

where $1 \le m \le \infty$, i.e., m indicates any real number beyond 1, u_{ij}, represents the degree of membership of x_i is the ith component of d-dimensional measured data, c_j indicates the d-dimension center of the cluster, as well as ||*|| represents any norm conveying the resemblance among any computed data and its center. Fuzzy partitioning is performed by means of an iterative optimization of the objective function exposed above, with the update of membership u_{ij}.

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{\left\|x_i - c_j\right\|}{\left\|x_i - c_j\right\|}\right)^{\frac{1}{m-1}}}$$

In addition tothe cluster centers c_j by:

$$c_j = \sum_{i=1}^{N}\frac{u_{ij}^{m}x_i}{u_{ij}^{m}}$$

This iteration will terminate when error,

$$\left\{\left\|u_{ij}^{k+1} - u_{ij}^{k}\right\|\right\} \le \varepsilon$$

where ε indicates a termination criterion in the range of 0 and 1, while k indicate the iteration steps. This process meets to a local minimum otherwise a saddle point of J_m.

The algorithm includes the following steps:

i. Initialize the quantity of clusters (c), weighting exponent (m), iteration limit, termination criterion ($\varepsilon > 0$) and $U = [u_{ij}]$ matrix $U^{(0)}$,

ii. Approximate the initial location of cluster centers.

iii. During the k the step,figure out the center vectors $c^k = [c_j]$

$$C_j = \sum_{i=2}^{N}\frac{u_{ij}^{m}x}{u_{ij}^{m}}$$

iv. Revise $U^{(k)}$ to $U^{(k+1)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{\left\|x_i - c_j\right\|}{\left\|x_i - c_j\right\|}\right)^{\frac{2}{m-1}}}$$

If $\| U^{(k+1)} - U^{(k)} \| \le \varepsilon$ then STOP;
If not, return to step(i).

3. Classification of Heart Rate Data using Enhanced Support Vector Machine

Several SVM kernel functions are formulated for users to decide from for several applications[22,23]. The most familiar kernel functions are the radial basis function, polynomial function, linear function and sigmoid function. All these kernel functions do not take the variations between characteristics of data. Based on the standard SVM kernel function determine $K(x_i, x)$, it can be noted that the overall features of the training or test datasets are utilized equally. This process possibly will not be successful and it will have a considerable effect on the accuracy of SVM. In order to solve this complication, it is essential to take the importance of different features into consideration is to append weights to a kernel function. The weights are employed to compute the significance of all features. A generic structure of the new kernel function is given as $K(wx_i, wx)$, where w indicates a vector incorporates weights of features of data set. In the same way, a nonlinear discriminant function with feature weights is given as follows,

$$f(x) = sgn\left(\sum_{i=1}^{l}\alpha_i y_i K(wx_i, wx) + b\right)$$

This enhanced kernel is self-regulating to specific kernel functions. For dissimilar applications, one might prefer the most appropriate kernel function to implement the feature weights on. Here, utilized rough set theory to compute and produce these weights based on training data. The elemental standards of weight computation are given as follows: 1. In case if a feature is not included in any reducts, at this stage the weight of this feature is taken as 0; 2. In case if the maximum times a feature comes into view in the reducts, at this stage it is taken as the more significant feature; 3. In case if the lesser number of features comes into view in a reduct, at this point the more significant these features appearing in this reduct are. In such scenario, when a reduct includes only one feature, the feature belonging to this reduct is the most significant. In consistent with the above principles, an algorithm is formulated and is depicted in Algorithm 1 that implements rough set theory for the purpose of ranking features and compute feature weights. Subsequent to the feature ranking process, considered those features with 0 weights as the least significant features and remove them. In Algorithm 1, feature ranking and feature selection are carried out in the same process.

Algorithm 1. Feature weights calculation using ESVM

```
Input:Database DB.
Output: A weight vector W.
Detect the entirereducts of DB by means of PSO;
$N_{feature}$ ← amount of features in DB;
$N_{reduct}$ ← amount of reducts of DB;
//Initialize the weight of each feature.
for (i ← 0 to $N_{feature}$ ) do
    $w_i$ ← 0;
end
// Compute the weight of all features.
for (i ← 0 to $N_{feature}$) do
    for (j ← 0 to $N_{reduct}$)
    do
        if (feature i in the jth reduct $R_j$)
        then m ← number of features in $R_j$;
        ;
    end
end
end
Scale the values of feature weights into the interval [0, 100];
```

4. Experimental Results and Discussion

Here, MIT-BIH arrhythmia database[24] is taken as the data source. It comprises of 48 recordings. Each one includes a phase of 30 minutes and includes two leads; the modified limb lead II and one of the modified leads V1, V2, V4 or V5. Sampling frequency is taken as 360 Hz, the data are bandpass filtered during 0.1–100 Hz and the resolution is taken as 200 samples per mV. 23 of the recordings are intended to make available as a representative sample of customary clinical recordings and 25 recordings incorporate complex ventricular, junctional and supraventricular arrhythmias. It includes more than 109,000 labelled ventricular beats from 15 diverse heartbeat classes. There is anenormous variation in the quantity of examples in all heart beat category. The most common and huge class is Normal beat with approximately 75,000 examples and the least class is Supraventricular Premature beat (SP) with just two instances. The database is indexed mutually in timing details and beat classification. In order to get more details please refer[25]. Here for the experimentation, 12 records are taken from the database. A sum of 30,873 beats was obtained; 22,476 normal beats, 5,394 abnormal PVC arrhythmia beats and 3,003 additional arrhythmic beats. The database index files from database are utilized to trace beats in ECG signals. The performance parameters are assessed from the confusion matrix as follows. Sensitivity (SE) is defined as, a measure of the capacity of test the positive samples.

$$SE = \frac{TP}{TP + FN}$$

Specificity (SP) is defined as, a measure of capacity of test the negative samples.

$$SE = \frac{TN}{TN + FP}$$

Accuracy is defined as

$$Accuracy = \frac{TN + TP}{TP + TN + FN + FP}$$

Where TP, TN, FP, and FN stand for true positive, true negative, false positive and false negative, respectively. For example the proposed work classifies with theses five classes such as Premature Ventricular Contraction (PVC), Normal Sinus Rhythm (NSR), Atrial Fibrillation (AF), Ventricular Fibrillation (VF) and 2° heart block (BII). Consider a division of HRV along with the VF arrhythmia iscategorized as the VF, subsequently it is considered that the segment is classified TP. In contrast, if a non-VF segment is categorized as non-VF, subsequently

Table 2. Sensitivity, specificity, and accuracy for each class (%) for ESVM method

Classes	Proposed ESVM with PSO based FS			Existing SVM classifier without FS		
	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
NSR	100	99.94	99.98	97.67	93.52	95.97
PVC	90	99.15	98.76	40.71	99.91	98.03
AF	96.48	99.19	98.53	95.38	97.66	97.13
VF	95.38	100	99.73	75.84	96.23	93.35
BII	100	99.88	99.89	60.85	94.65	95.45
Average	96.37	99.63	99.38	74.09	96.394	95.986

it is considered that the segment is classified TN. If any non-VF segment which is categorized as a VF segment by error will generate a FP, at the same time any VF segment which is categorized as a non-VF segment by fault will generate a FN result. In order to assess the performance of the proposed classifier, here taken a total of 1317 segments, which are taken from the MIT-BIH arrhythmia database, were utilized and it has 835 NSR segments, 57 PVC segments, 322 AF segments, 78 VF segments and 25 BII segments. The results are presented in Table 2.

4.1 Accuracy Comparison

The proposed PSO-SVM algorithm produces better accuracy rate shown in Figure 2 which is much greater accuracy results than existing SVM algorithm. When the number of features increases the accuracy of the result is increases. This approach produces high accuracy rate when compared to existing system.

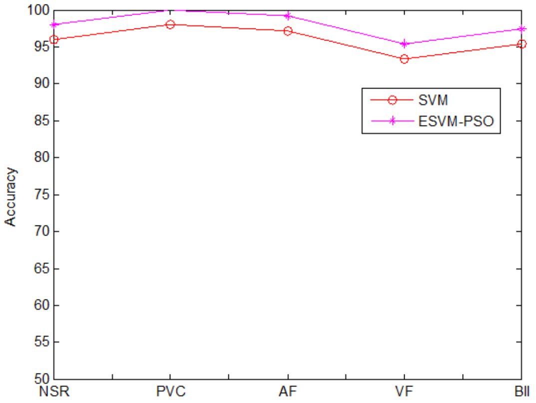

Figure 2. Accuracy comparison.

4.2 Sensitivity Comparison

The proposed ESVM with PSO based FS algorithm produces high sensitivity shown in Figure 3 which is much greater accuracy results than existing SVM algorithm without FS. When the number of features increases the sensitivity of the result is increases. This approach produces high sensitivity rate when compared to existing system.

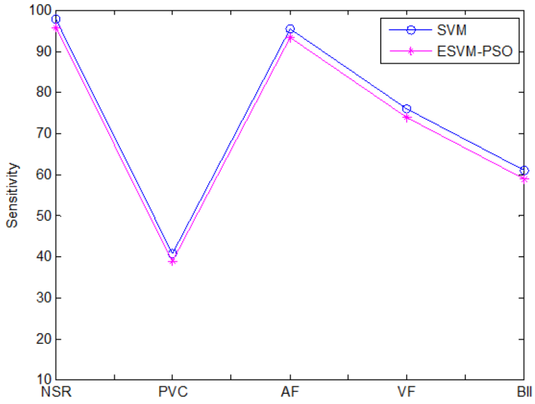

Figure 3. Sensitivity comparison.

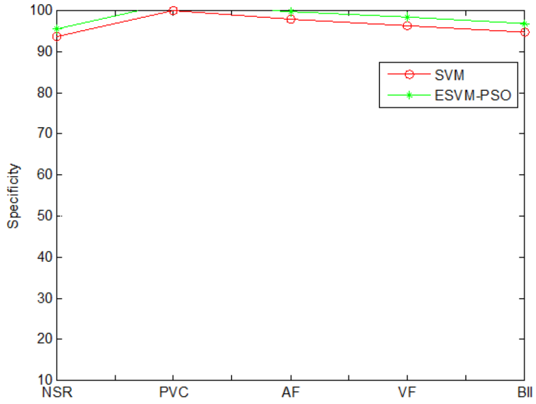

Figure 4. Specificity comparison.

4.3 Specificity Comparison

The proposed ESVM with PSO based FS algorithm produces high specificity shown in Figure 4 is higher than the existing SVM algorithm. When the number of features increases the specificity of the result is increases. This approach produces effective specificity rate when compared to existing system.

4.4 F1-Measure Comparison

F1-measure is defined as the harmonic mean of precision and recall. A good classifier is assumed to have a high F1-measure, which indicates that the classifier performs well with respect to both Precision (P) and Recall (R). The proposed ESVM with PSO based FS algorithm produces high F-measure shown in Figure 5 is higher than the existing SVM algorithm. When the number of features increases the F-measure of the result is increases. This approach produces effective F-measure rate when compared to existing system.

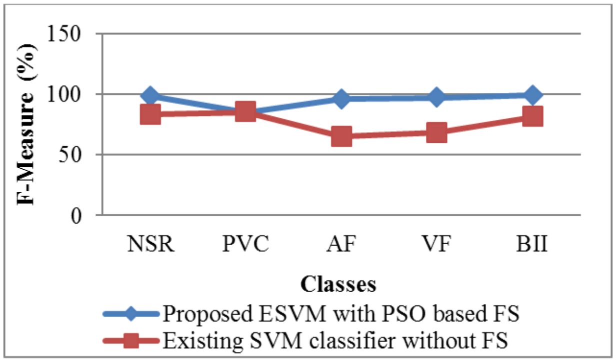

Figure 5. F-measure comparison.

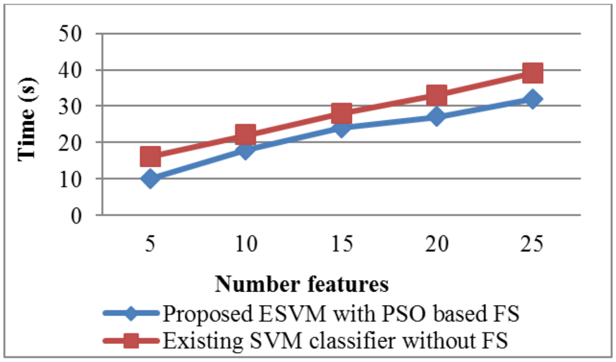

Figure 6. Execution time comparison.

4.5 Execution Time Comparison

Figure 6 shows the Number of selected features for different feature vector dimensions using the ESVM with PSO based FS and SVM without feature selection algorithms. The best accuracy rate is achieved using the linear and non-linear feature vector and the PSO-based feature selection algorithm using only less number of selected features and with approximately 35% less selected features which is significantly improves the accuracy and reduces the computational complexity.

5. Conclusion

In this paper, the Enhanced SVM classifier is presented as classifier to classify the heart rate data. Enhanced SVM classifier gets the grouping of linear and non-linear parametersas inputs, which are derived from the HRV signal. Enhanced SVM classifier demonstrated adequate results in discriminating five categories of arrhythmia with PSO feature selection algorithm with FCM clustering. The accuracy of NSR, PVC, AF, VF and BII were 99.98%, 98.76%, 98.53%, 99.73%, and 99.89%, respectively. The results prove that this method is extremely effective for classification of heart rate data, with a satisfactory high accuracy. It is apparent that the mixture of the linear and nonlinear features accompanied by the classifier is exceptionally efficient. The major benefit of the scheme, when compared to other schemes is that the feature can be extracted with better accuracy even in case of noisy or complicated ECG recordings; at the same time the selection of optimal features is done using PSO that effectively increase the accuracy in classification result. Finally, because of its low complexity and high accuracy, the proposed method can be used as a real-time arrhythmia classifier for reliable clinical monitoring of patients. As a future work, the hybrid techniques with swarm based methods can effectively reduce the execution time.

6. References

1. Kumar AS. ECG in 100 steps!2013 Mar 05. Available from: http://www.lifehugger.com/doc/120/ecg100-steps
2. Kampouraki A, Manis G, Nikou C. Heartbeat time series classification with support vector machines. IEEE Transactions on Information Technology in Biomedicine. 2009 Jun; 13(4):512–8.
3. Kamath M, Fallen E. Power spectral analysis of HRV: A noninvasive signature of cardiac autonomic functions. Crit Rev Biomed Eng. 1993; 21(3):245–311.
4. Silipo R, Deco G, Vergassola R, Gremigni C. A characterization of HRV's nonlinear hidden dynamics by means of Markov models. IEEE Trans Biomed Eng. 1999 Aug; 46(8):978–86.

5. Yu SN, Chou KT. Selection of significant for ECG beat classification. Expert Syst. 2009; 36(2):2088–96.
6. Sarvestani RR, Boostani R, Roopaei M. VT and VF classification using trajectory analysis. Nonlinear Anal. 2008; 25–33. DOI: 10.1016/j.na.2008.10.015.
7. Ebrahimzadeh A, Khazaee A. An efficient technique for classification of Electrocardiogram signals. Advances in Electrical and Computer Engineering. 2009; 9(3):89–93.
8. Ince T, Kiranyaz S, Gabbouj M. A generic and robust system for automated patient-specific classification of electrocardiogram signals. IEEE Trans Biomed Eng. 2009; 56(5):1415–26.
9. Lin CH. Frequency-domain features for ECG beat discrimination using grey relational analysis-based classifier. Comput Math Appl. 2008;55(4):680–90.
10. Ubeyli ED. Support vector machines for detection of electrocardiographic changes in partial epileptic patients. Engineering Applications of Artificial Intelligence. 2008; 21(8):1196–203.
11. Derya EU. Recurrent neural networks employing Lyapunov exponents for analysis of ECG signals. Expert Systems with Applications. 2010; 37(2):1192–9.
12. Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE Trans Biomed Eng. 2004; 51(7):1196–206.
13. Lagerholm M, et al. Clustering ECG complexes using Hermite functions and self-organizing maps. IEEE Trans Biomed Eng. 2000;47(7):839–47.
14. Khazaee A, Ebrahimzadeh A. Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features. Biomedical Signal Processing and Control. 2010; 5(4):252–63.
15. Andreao RV, Dorizzi B,Boudy J. ECG signal analysis through hidden Markov models. IEEE Trans Biomed Eng. 2006 Aug; 53(8):1541–9.
16. Joy Martis R, Chakraborty C, Ray AK. A two stage mechanism for registration and classification of ECG using Gaussian mixture model. Pattern Recognition. 2009; 42(11):2979–88.
17. Mitra S, Mitra M, Chaudhuri BB. A rough set based inference engine for ECG classification. IEEE Trans Instrum Meas. 2006; 55(6):2198–206.
18. Philip C, Richard RB. A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features. IEEE Transactions Biomedical Engineering. 2006; 53(12):2535–43.
19. Majid M, Khorrami H. A qualitative comparison of artificial neural networks and support vector machines in ECG arrhythmias classification. Expert Systems with Applications. 2010; 37(4):3088–93.
20. Kennedy J, Eberhart RC. Particle swarm optimization. Proceedings of IEEE International Conference on Neural Networks; Piscataway, NJ. 1995.p. 1942–8.
21. Shi YH, Eberhart RC. Empirical study of particle swarm optimization. Proceedings of the Congress on Evolutionary Computation; Washington, DC. 1999. p. 1945–50.
22. Zadeh AE, Khazaee A. High efficient system for automatic classification of the electrocardiogram beats. Ann Biomed Eng. 2011; 39(3):996–1011.
23. Ebrahimzadeh A, Khazaee A. An efficient technique for classification of Electrocardiogram signals. Advances in Electrical and Computer Engineering. 2009; 9(3):89–93.
24. Mark RG, Moody GB. MIT-BIH arrhythmia database. 1997. Available from:http://ecg.mit.edu/dbinfo.html
25. Moody GB, Mark RG. The impact of the MIT/BIH arrhythmia database. IEEE Eng MedBiol Mag. 2001 May-Jun; 20(3):45–50.
26. Soleimanian F, Gharehchopogh, Ebrahimi L, Isa M, Gourabi SJ. A novel PSO based approach with hybrid of Fuzzy C-means and learning automata in software cost estimation. Indian Journal of Science and Technology. 2014 Jun; 7(6):795–803.