Evolving Systems (2020) 11:15–28
https://doi.org/10.1007/s12530-019-09312-6

ORIGINAL PAPER

Heart disease detection using hybrid of bacterial foraging and particle swarm optimization

Padmavathi Kora[1] · Ajith Abraham[2] · K Meenakshi[1]

Received: 22 November 2018 / Accepted: 6 November 2019 / Published online: 25 November 2019

Abstract
Bacterial-foraging-optimization (BFO) has newly raised and one of the most useful nature inspired optimization algorithm for real parametric optimization. During the process of random walk, the BFO algorithm makes search in the random direction, which increases delay. To overcome the delay in reaching the global optimum and also to boost up the performance of BFO, we proposed an algorithm by mixing the features of BFO and particle swarm optimization (PSO) for detecting the abnormal cardiac beat. Computer simulations illustrate the usefulness of the developed approach compared to the basic versions of BFO and PSO. The main aim of the research is to develop new modifications of BFO and its combination with transform technique such as Wavelet Transform and machine learning method, support vector machines (SVMs) to test their performances in the detection of cardiac arrhythmia. Modification of BFO focuses for improving its convergence in terms of speed and accuracy. Provided results in this paper show that, for the detection of MI and BBB classes, the BFPSO algorithm with SVM gives 98.9% and 99.3% accuracy on MIT-BIH database by including NSR database also. Moreover, the results demonstrate the effectiveness of the proposed method to improve the detection of cardiac arrhythmia.

Keywords Electrocardiography (ECG) · Bacterial foraging optimization (BFO) · Particle swarm optimization (PSO) · Support vector machine (SVM)

1 Introduction

Coronary heart diseases (CHD) are the leading cause of death in India. WHO reports predicts that nearly 2.6 million Indians cause morbidity and death by 2020 due to CHD. In comparison with European countries Indians are highly effected by CHD in early age (In western countries before 70 years the CHD deaths are about 23% whereas in India it is 52%) (Kumar 2014). ECG is an approach for measuring the heart's electrical activity, typically a non-invasive approach to detecting abnormalities in the heart. ECG signal has various characteristics. A typical heartbeat or cycle begins with a P wave followed by a QRS complex. The beat then ends with a T wave. Occasionally, the U wave appears after the T wave. More importantly, rhythm and the morphology of the ECG waveform is altered by cardiovascular diseases and abnormalities such as the cardiac arrhythmias. Hence this research focus on their automatic detection and classification. The aim of this paper is to develop such a computer aided diagnostic system which assists expert cardiologists by providing intelligent, cost effective and time saving arrhythmia diagnostics. To achieve this goal, conventional ECG signal processing techniques along hybrid BF-PSO optimization technique are implemented to identify the arrhythmia recognition. Different diseases are identified, namely AF, MI and BBB are given below.

1.1 Atrial fibrillation

Atrial fibrillation (AFib or AF) is a trembling or unpredictable pulse (arrhythmia) that can prompt blood clusters, stroke, heart attack and other heart-related entanglements. Not less than 2.7 million Americans are living with AF.

Normally heart contracts and relaxes to complete one beat. In atrial fibrillation, the upper assemblies of the heart (atria) beat asynchronously with the ventricles, causing insufficient blood flow into the ventricles. This condition

✉ Padmavathi Kora
padma386@gmail.com

[1] Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India

[2] Machine Intelligence Research Labs, Washington, USA

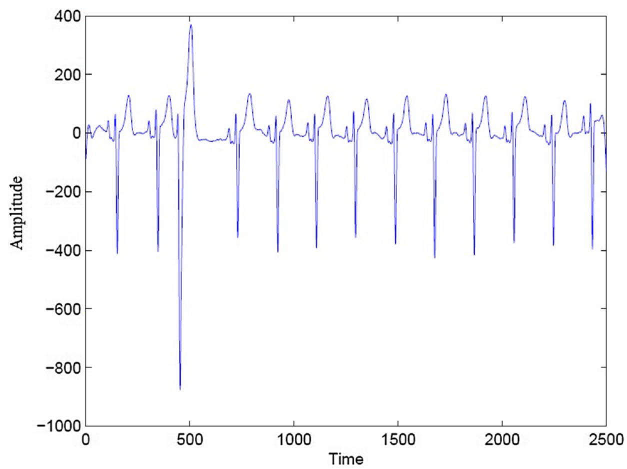

Fig. 1 AF signal

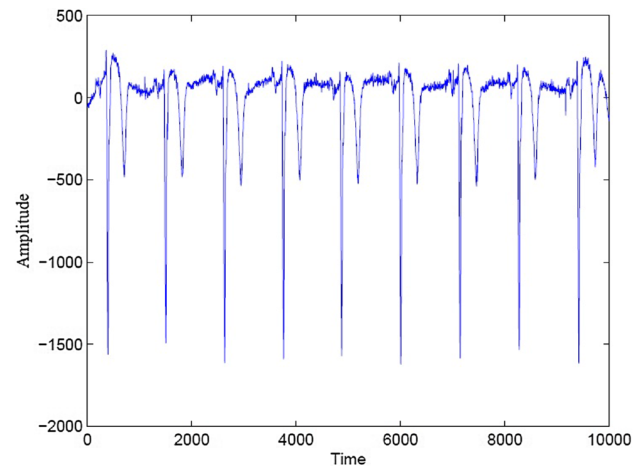

Fig. 2 Myocardial infarction signal

increases the formation of blood clots in the atria. The following morphological changes present in ECG, fluctuating (random) waveforms instead of P waves (f-waves); irregular RR interval are shown in Fig. 1.

1.2 Myocardial infarction (MI)

MI (Acharya et al. 2017; Padmavathi and Krishna 2014) is a hazardous condition that happens when blood stream to the heart muscle is suddenly cut off, causing tissue harm. This is normally the consequence of a blockage in at least one of the coronary supply routes. A blockage can create because of a development of plaque, a substance which mainly consists of cholesterol, and cell squander (waste) items. Heart is the primary organ in our cardiovascular system, which likewise incorporates diverse sorts of arteries. Arteries supply oxygen-rich blood to our body and the majority of our organs. The coronary artries supply oxygen rich blood to heart muscle. At the point when these supply routes wind up noticeably blocked or limited because of the development of plaque, the blood stream to heart can diminish fundamentally or stop totally. This can cause a heart assault. A few components may prompt a blockage in the coronary arteries. Two different types MI can be observed from ECG they are ST elevation (type 1) and ST dipression (type 2). In this paper we considered type 2 MI signal. The morphological changes in type 2 MI are shown in Fig. 2.

1.3 Bundle branch block (BBB)

The heart's electrical movement starts in the sinoatrial junction (the heart's common pacemaker), which is arranged on the upper right chamber. The drive goes next through the left and right atria and summates at the atrioventricular point (AV node). From the AV point the electrical drive goes down the bundle-his and partitions into the right and left bundle branches. At last, the branches spread into a large number of purkinje fibres, which thus interdigitise with individual cardiovascular myocytes, taking into consideration fast, organized, and synchronous physiologic depolarization of the ventricles.

BBB condition can be analyzed when the length of the QRS complex on the ECG more than 120 ms. BBB commonly cause prolongation of the QRS complex, and may move the heart's electrical axis marginally to one side. The ECG will demonstrate a terminal R wave in lead V1 and a slurred S wave in lead I. LBBB extends the whole QRS, and much of the time moves the heart's electrical axis to one side. Another ordinary finding with BBB is suitable T wave dipression. The T wave will be diverted inverse the terminal avoidance of the QRS complex. Left block can prompt cardiovascular dyssynchrony. The occurance of left and right blocks at the same time prompts add up to AV block. The changes in the BBB signals shown in Figs. 3 and 4.

There are mainly four stages (steps) in the traditional ECG signal classification [(i) preprocessing; (ii) feature extraction; (iii) feature optimization; (iv) classification] as shown in Fig. 5.

Several transforms have been proposed for signal analysis (Hramov et al. 2015; Vetterli et al. 2014). The choice of signal transforms for such analysis is usually due to some useful properties that these transforms have, including their compact representation of a signal, inversibility, availability of fast versions for computation, capacity of analyzing signals at each frequency independently, among others. One of the most commonly used feature extraction techniques used today is the wavelet transform, has many applications such as signal processing (Xizhi 2008), denoising (Engin 2004), feature extraction (Faust et al. 2015; Ceylan and Özbay 2011) a time-frequency transformation. Wavelet

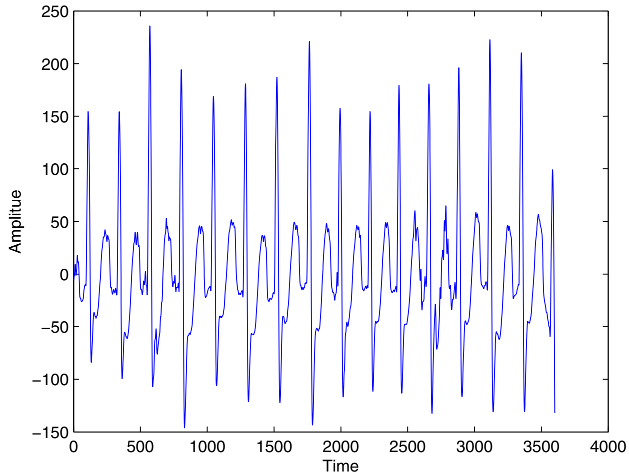

Fig. 3 Left bundle branch block

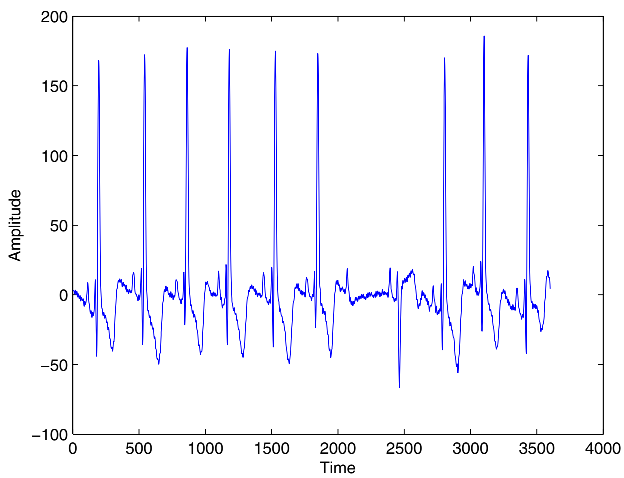

Fig. 4 Right bundle branch block

transforms are preferred for feature extraction as they can express features in both the time domain and frequency domain. Using wavelet transforms has a drawback, though, as additional processing is needed to discern which features are most applicable. Examples of additional processing used to extract the most important features include principal component analysis (PCA), independent component analysis (ICA), generalized discriminate analysis, or linear discriminate analysis. For each of these machine learning feature extraction algorithms, the ECG data are be fed into the classifier model, which uses a predetermined algorithm to separate the features into their various classes. Physiological signals like ECG are considered to be quasi-periodic in nature. They are of finite duration and non stationary. Hence, a techniques like Wavelet coherence, wavelet transform are inefficient because they produce large number of redundant and noisy features. Therefore the common practice is to extract the key features useful for the classification. On the other hand, nature inspired optimization algorithms are the recent addition in this field of research, provides a powerful tool for extracting important features from each ECG beat.

In the classification of cardiac arrhythmia, the high dimensional features often brings a barrier for researchers. Actually, in classification tasks, the unlabeled instances are required to be classified into specific groups according to informative features. The reality is that not all features are useful for classification, and some redundant and irrelevant features may even serve as obstacles. Consequently, the accuracy of classification is also decreased. The challenge is that, without previous knowledge, it is difficult to distinguish the most representative and compact features from the useless ones when the dimensional space is rather high. So the feature selection, as the effective tool, have been wildly applied to select a relative small but useful feature subsets from available features. The purpose of feature selection is to eliminate the redundant and irrelevant features in great deal to achieve the highest classification accuracy.

Evolutionary based algorithms like differential evolution (DE), GA, genetic programming (GP) which have been frequently shown to be effective in solving the problems in complex conditions are presented for feature selection (Angelov and Kasabov 2006). Most of them are binary description algorithms, and every bit represents a feature, e.g. the value "1" represents a selected feature while value "0" means a feature that is not selected. However, parameters used in genetic or evolutionary based algorithms need to be properly defined to obtain the considerable performance.Since the swarm algorithms like PSO, ACO are good at local research, while evolutionary-based algorithms like

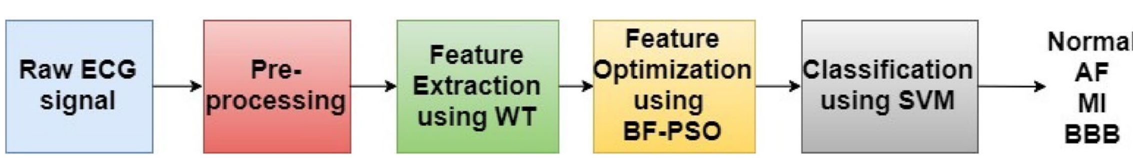

Fig. 5 Arrhythmia classification steps

GA have a better global searching ability.The combinations of the swarm and evolutionary algorithms are presented in the literature, such as hybrid GA-ACO (Roeva et al. 2016; Lee et al. 2008) Hybrid GA-ACO-PSO (Tam et al. 2008) are good at both local and global search. Even so, the effectiveness of feature selection would be decreased when the size of search space increases. Most of those feature selection methods are suffering from the computational complexity or local optima. To address high dimensionality feature selection problems, an effective feature selection method with global search ability is demanded. During the process of random search, the BFO try different search directions which may lead to delay in reaching global solution. In order to speed the convergence of Bacterial Foraging Optimization, (Rajasekhar et al. 2011) had proposed an improved BFO namely BF-PSO. The BF-PSO algorithm borrowed the ideas of velocity updating from PSO, the search directions specified of the bacteria are oriented by the individual best location and global best locations concurrently.

2 Related work

Different research groups around the world developed and used novel meta-heuristic optimization methods like nature inspired PSO (Ordonez De León et al. 2019; Moeini and Babaei 2017, Cuckoo Search Gandomi et al. 2013) Firefly Algorithms (Yang 2010), BFO (Das et al. 2009) in machine learning applications such as biomedical signal processing (Kora 2017), data mining (Fong et al. 2016), image retrieval (Younus et al. 2015), manufacturing design (Raju et al. 2018) and robotic platform (Garcia et al. 2009), etc. The former researchers (Biswas et al. 2007; El-Wakeel et al. 2015; Anguluri et al. 2011) also applied the hybrid BF-PSO algorithms for solving optimization problems of numerical benchmark, permanent magnet brush less dc motor, and FOPI Speed Controller etc. They all reported that the performance of hybrid BF-PSO algorithm is far better than classical PSO and BFO algorithm. Furthermore, the hybrid BF-PSO algorithm involves the elite-selection mechanism to reproduce the near optimal solutions, thereby balance the exploration and exploitation abilities of the search space (Padmavathi and Kalva 2015). BFO algorithms moves in the in the random directions in the search space (tumble behavior), each bacterium moves in a random direction, this random search behavior may lead to delay in reaching the local and global optimum point. So in hybrid BF-PSO (Biswas et al. 2007) algorithm social information ability of PSO and elimination and reproduction behavior of BFO are combined to overcome the delay in reaching the global optimum and also to boost up the performance operation of BFO. Lee et al. (2013) extracted linear and nonlinear features and classified using different types of classifiers to detect coronary artery diseases. Kim et al. (2007) used multiple discriminant analysis to differtiate three heart diseases. Sridhar et al. (2016) compared performance of DWT and nonlinear techniques techniques to detect normal and coronary heart diseases using SVM, Decision Tree (DT), K-nearest neighbor (KNN) and probability neural network (PNN) classifiers. Schreck et al. (1988) used Emperical mode decomposition technique to classify normal and resting ECGs. Lehtinen et al. (1988) utilized multilayer perceptron neural network to classify coronary artery diseases and showed that computer aided diagnosis increased the detection accuracy. Lewenstein (2001) used neural network to detect coronary artery disease and classified as healthy and unhelathy patients. Babaoglu et al. (2010) used binary particle swarm optimization and genetic algorithm as feature optimization techniques and SVM as classification technique to detect coronary heart diseases. Kaveh and Chung (2013) used electrocardiogram exercise stress test data collected from Physionet database to detect coronary artery atherosclerosis. Features are extracted and optimized using DWT and principal component analysis (PCA) then classified using SVM. Acharya et al. (2017) utilized higher-order statistics and spectra and classified different coronary heart diseases using KNN and decision tree classifiers. Kumar et al. (2017) used flexible analytic Wavelet Transform to decompose ECG signals and then classified using least squares-support vector machine. Classification accuracy is 99.6% (Morlet wavelet kernel) in comparison to 99.56% (radial basis function kernel) (Table 1).

3 Methods

In the preprocessing, the data has been collected from MIT-BIH AF data base consisting normal sinus rhythm database (18 patients, 128 Hz) and AF database (26 patients, 250 Hz) of 30 min duration. The MI data has been collected from Physionet PTB database consisting 52 normal individuals and 148 MI patients data at a sampling rate of 1000 Hz. From the PTB data base, 6 normal and 7 MI files of duration 30 min have been used. The BBB data has been collected from the MIT-BIH arrhythmia data base consisting 48 half channel ambulatory ECG recordings obtained from 47 individuals. From the arrhythmia data base 3 LBBB and 3 RBBB files of duration of 30 min at 360 Hz sampling rate have been used.

This section describes the proposed methods for classification of electrocardiogram (ECG) signals based on a set of features extracted from various time-frequency transforms. Different classifiers are employed and evaluated over a collection of signals as in Fig. 6.

Table 1 Accuracy comparison for the identification of cardiac arrhythmia

Author	Approach	Performance
Kim et al. (2007)	Linear features nonlinear features multiple discriminant analysis	Angina sen = 72.5% Spe = 81.8% coronary syndrome sen = 84.6% Spe = 91.5%
Sridhar et al. (2016)	DWT, nonlinear features and classifiers: SVM, DT, KNN and PNN	Sensitivity = 95.02 % Specificity = 99.2% Accuracy = 98.67%
Schreck et al. (1988)	Emperical mode decomposition	Men sen = 84.3% Spe = 81.8% Women sen = 76.2% Spe = 80%
Lehtinen et al. (1988)	ANN	ROC = 91.5%
Lewenstein (2001)	RBF Neural Network	Sensitivity = 97% Specificity = 98%
Babaoglu et al. (2010)	PCA	Accuracy = 80%
Kaveh and Chung (2013)	DWT & Principal component analysis and SVM	Accuracy = 80%
Acharya et al. (2017)	Bispectrum and cumulant	Sen = 94.8%, SPe = 99.3%, Acc = 98.2%
Kumar et al. (2017)	Features flexible analytic wavelet transform cross information potential(CIP) classifiers KNN, DT	Sen = 94.8%, SPe = 99.3%, Acc = 98.2%
Proposed	BFPSO & SVM	AF Acc = 99.1% MI Acc = 98.9% BBB Acc = 99.3%

3.1 Wavelet transform

Wavelet Transforms are efficiently useful for the identification of abrupt changes in the signal. Wavelet is a rapidly decaying wave like oscallation that has zero mean. Unlike the the sinusoids which extends upto infinity a wavelet exists for finite duration. Wavelets are in different sizes and shapes. The availability of wide range of wavelets is the key strength of wavelet analyis. Scaling and shifting are the two important concepts of wavelets. For each ECG segment coefficients of wavelets d(j, n) and scaling functions c(n) are calculated via

$$f(t) = \sum_{n=-\infty}^{\infty} c(n)\phi(t-n) + \sum_{j=0}^{\infty}\sum_{n=-\infty}^{\infty} d(j,n)2^{j/2}\psi(2^j t - n) \tag{1}$$

$$c(n) = \int_{-\infty}^{\infty} f(t)\phi(t-n)dt \tag{2}$$

$$d(j,n) = 2^{j/2}\int_{-\infty}^{\infty} f(t)\psi(2^j t - n)dt. \tag{3}$$

We have extracted 6 six approximated coefficients A1–A6 and six detailed coefficients D1–D6 using the above expression as shown in Fig. 7.

3.2 Hybrid bacterial foraging and particle swarm optimization

Biologically inspired systems-based algorithms were developed by modeling the behavior of animals in doing something collectively, for example foraging in groups, rather than individually. Bacterial foraging Optimization (BFO),

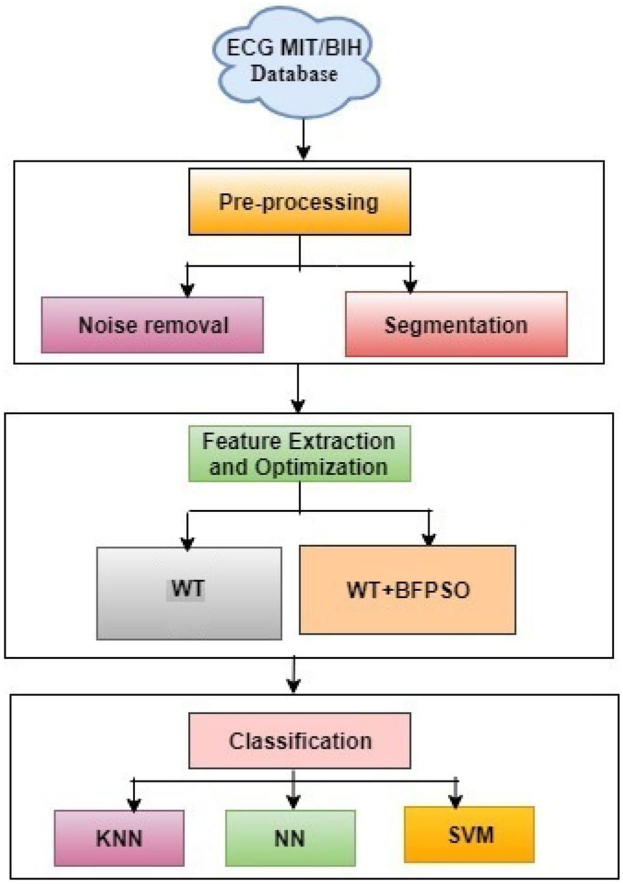

Fig. 6 Classification flow diagram

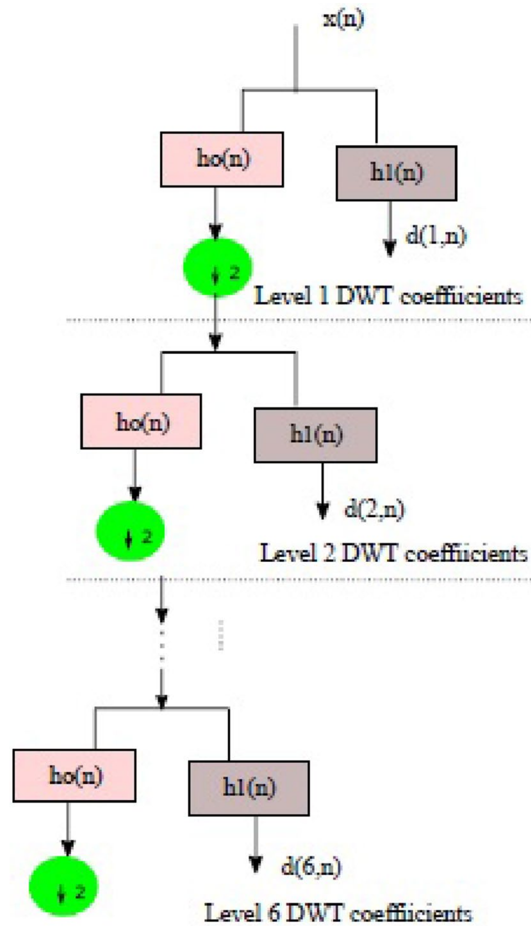

Fig. 7 Discrete wavelet transform

which has been developed based on foraging strategies of *Escherichia Coli* bacteria, emerged as one of relatively the most newly developed biologically-inspired optimisation methods. *E. Coli* bacteria always try to find a place which has high level of nutrition and avoid a place which has noxious substance. From the optimisation point of view, the optimum value is the place which has the highest nutrient level. The initial position of bacteria could be set in certain predetermined position or dispersed randomly across the nutrient media (search space). Bacteria will move towards global optimum position by applying "biased" random walk called chemotaxis with constant "step size". After performing a predetermined number of chemotactic steps, the health of bacteria is sorted based on the nutrient level they get. Half of the population with high nutrient level, the healthy bacteria, will reproduce and takes the same place of their mother while the remaining half with lower nutrient level, unhealthy bacteria, will die. By using this mechanism, the number of bacteria in the population will remain constant. Because of external factors, the number of bacteria in the population could be decreased sharply or dispersed to other regions of nutrient media. This event makes bacteria able to explore most parts or whole regions of nutrient media.

From the literature, it can be noted that, since it is a relatively new optimization technique, there are still spaces that can be explored on the modification of BFO especially to increase its convergence speed and accuracy and also its possible combination with other soft computing techniques such as fuzzy logic and neural networks. Moreover, despite its potential, BFO has not been reported yet for modeling and control of flexible manipulator systems. The main aim of the research is to develop new modifications of BFO and its combination with PSO technique and test their performances in detecting abnormal cardiac beat. Modification of BFO is focusing on maximizing the detection accuracy of cardiac beat and the performance of modified BFO will be compared to that of the standard BFO.

Natural selection tends to eliminate animals with poor foraging strategies and favor the propagation of genes of those animals that have successful foraging strategies since they are more likely to enjoy reproductive success. After many generations poor foraging strategies are either eliminated or shaped into good ones. Until date BFO has found its successful implementation real world problems such as PI/PID controller design (Debbarma et al. 2014), stock market prediction (Majhi et al. 2009), and power systems (Kumar and Jayabarathi 2012). However, during the process of chemotaxis, the BFO depends on random search directions which may lead to delay in reaching global solution. In order to speed the convergence of Bacterial Foraging Optimization, Raja sekhar et al. (Rajasekhar et al. 2011) had proposed an improved BFO namely BF-PSO. The BF-PSO algorithm borrowed the ideas of velocity updating from PSO, the search directions specified of the bacteria are oriented by the individual best location and global best locations concurrently. In the tumble behavior each bacterium moves in a random direction, this random search behavior may lead to delay in reaching the local and global optimum point. So in hybrid BF-PSO algorithm social information ability of PSO and elimination and dispersal behavior of BFO are combined. The chemotaxis step of BFO is updated by velocity equation of PSO as shown in Eq. 4:

$$v(j+1) = W \times v(j) + C1 \times rand \times \left(p_{best} - P_{current}\right) + C2 \times rand \times \left(g_{best} - P_{current}\right) \tag{4}$$

where pbest is the best position of each bacterium and gbest is the global best position. The BF-PSO procedure given in algorithm 1 and flow chart shown in Fig. 8.

The initial the parameters: D,N_p, N_c, N_r,C and $x\ i$, where D is the dimension of search space, N_p is the total number of population (bacteria), N_c is the total chemotatic steps, N_{re} is the total reproduction loops, $x\ i$ is the ith bacterium location. C1, C2, w is the PSO parameters as shown in Table 2.

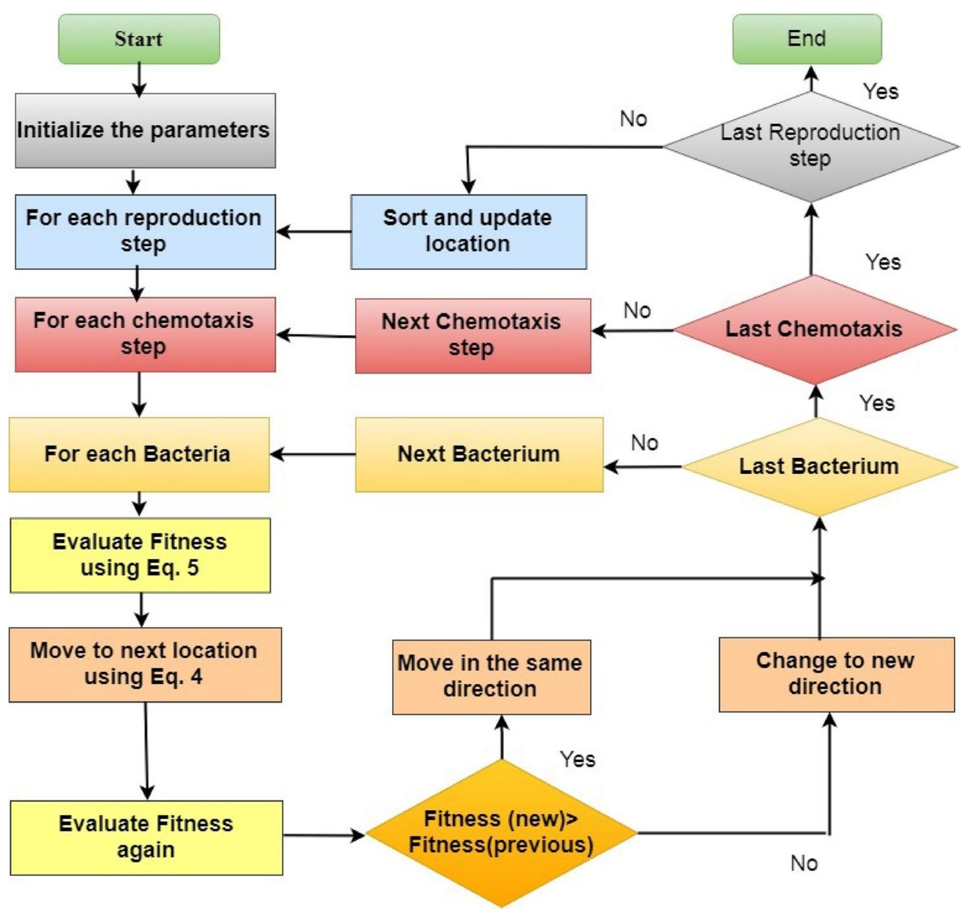

Fig. 8 BF-PSO flowchart

Algorithm 1: Hybrid BF-PSO algorithm

```
step-1 : Initialize the parameters.
step-2 : Reproduction loop k;
step-3 : Chemotaxis loop j;
step-4 : For each bacterium i;

for j=1:Np do
  for i=1:D-1 do
    step-5 : Compute the fitness at initial
    position using objective function and
    save the fitness J.
  end for
end for

for k=1:Nr do
  Jnew=J;
  for j=1:Nc do
    for i=1:Np do
      step-6 : Tumble: move the
      bacterium using Eq. 4 (from PSO).

      Again compute the fitness J.
      step-7: Swim
      if present_fitness >
      previous_fitness then
        move the bacterium the direction
        of gbest and pbest.
        Again compute the fitness J.
      else
        Move the bacterium in the using
        x(j + 1)=x(j) + v(j + 1)
        Compute new fitness Jnew.
      end if
    end for
  end for
end for

for i=1:Np do
  Jhealth(i)=sum(Jnew(i,:));
  Based on health of bacterium sort and
  update location.
end for
```

In this hybrid combination, PSO performs tumble operation and produces a near optimal solution very rapidly which is then followed by a global search elimination of dispersal, which fine-tunes the solution and gives an optimum solution. PSO has an inherent disadvantage of being trapped in the local optimum but has high convergence speed whereas BFO has the drawback of having a very poor convergence speed but has the ability of not being trapped in the local optima. The fitness function (ECG equation based on Gaussian model) shown in Eq. 5.

$$x(i) = \sum_{i\in P,R,S} \left(\pm\frac{d}{dt}\right)^{i} M_i e\left(-\frac{t-ti}{\sqrt{2}w_i}\right) + \sum_{i\in Q,j\in i} M_{ij} e\left(-\frac{t-t_{ij}}{w_{ij}}\right)^2 \tag{5}$$

If $i \in P, Q, R, S$ then Gaussian wave for each component of ECG wave have following parameter: M_i is height of curves peak, t_i is the center position of the peak and w_i controls the width of ECG signal.

$(\pm\frac{d}{dt}) = \left(\frac{d}{dt}\right)$ if $i \in R$ and $(\pm\frac{d}{dt}) = -1$ if $i \in S$.x

3.3 Clasification of ECG beats: support vector machine (SVM)

supervised learning in which at the each instant of input is applied, the desired class label must be provided to the system. SVM is basically a binary classifier. It is a supervised learning algorithm and can be used for both classification and regression. It adjusts weights of the network so error between the actual and predicted one decreases. It views the given database as two sets of vectors x_1 and x_2 in an 'n' dimensional space and constructs a separating hyper plane that maximizes the margin between the signal relevant to training and the signals for testing. SVM is a kernel method and the kernel function used in SVM is very crucial in determining its performance. Kernel function "svmtrain" uses to map the training data into kernel space. The kernel function can be one of the following character vectors or a function handle: The kernels may be linear, quadratic, polynomial, Gaussian Radial Basis Function and Multi-layer Perceptron. In this Gaussian RBF is used. But in this work, there are P classes where P > 2. There are three different multiple-class SVM techniques. They are

- One against one.
- One against all.
- Error correcting.

1. *One against one* This approach involves constructing a binary classifier for each pair of classes resulting in $\frac{N\times N\text{-}1}{2}$ binary classifiers. When applied to a test feature vector, each binary classifier gives one vote to the winning class

and the point is labeled with the class having majority votes.

2. *One against all* For N different classes. One vs all will train one classifier per class in total N classifiers. For class i it will assume i-labels as positive and the rest as negative. It is computationally complex.
3. *Error correcting* To speed up the process, Error correcting SVM is designed in the context of error correcting code, inspired by the coding schemes used in communications. For 'P' classes, the number of binary classifiers required is ${}^{P}C_2$ denoted by L.

$$\begin{bmatrix} -1\& -1\& -1 \\ +1\& -1\& +1 \\ +1\& +1\& -1 \end{bmatrix}$$

Each class is represented by a binary string of length L. During training, the *k*th classifier, k = 0, 1, 2, ⋯, L, the desired class labels may be either ±1. For understanding purpose, we choose M = 3. The number of binary classifiers required is $L = {}^{3}C_2$. Therefore L = 3. For each class, the desired labels may be different for various classes. This is equivalent to constructing a matrix of M × L of chosen labels.

During training, the first classifier is designed to respond (−1, +1, +1). The second classifier will be trained to respond (−1, −1, +1) and so on. The procedure is equivalent to grouping the classes in to L different pairs, and, for each pair, a binary classifier is trained accordingly. In this, for the first binary classification, class w_2 has been grouped together with w_3 and has been assigned label '1' and w_1 with label 0. Similarly, in the second binary classifier, class w_1 has been grouped together with w_2 with label −1 and w_3 with 1. This is repeated for all the classifiers. Each row must be distinct and may corresponds to a class label. In the absense of errors, the outputs of the L classifier for a pattern for class 1 is [−1, −1, −1]. When an unknown pattern is applied, the output of each L binary classifiers is calculated. Then, Hamming distance of this code word is compared with M code words and the pattern is classified as class corresponding to the small distance. The RBF kernel is employed with a sigma of 0.1 for training the network. Total 1800 cases are used for training and 1006 beats are used for testing in order to analyze the performance of the classifier.

4 Results

This section presents the experimental results obtained through the proposed ECG signal classification method, describes the data set used in the experiments, and performs a comparative analysis against other methods available in the literature.

The data has been collected from MIT-BIH AF data base consisting normal sinus rhythm database (18 patients, 128 Hz) and AF database (26 patients, 250 Hz) of 30 minutes duration. The MI data has been collected from Physionet (MIT-BIH) PTB database consisting 52 normal individuals and 148 MI patients data at a sampling rate of 1000 Hz. From the PTB data base, 6 normal and 7 MI files of duration 30 minutes have been used. The BBB data has been collected from the MIT-BIH arrhythmia data base consisting 48 half channel ambulatory ECG recordings obtained from 47 individuals. From the arrhythmia data base 3 LBBB and 3 RBBB files of duration of 30 minutes at 360 Hz sampling rate have been used.

4.1 Output of wavelet transform

The ECG signal, consisting of many characteristic points. These points characterize the behaviour of the ECG signal. Representing these points (features) with a lesser number of parameters to represent the ECG signal is particularly important for recognition and diagnostic functions. These multiple-scale features of the DWT allows the decomposition of a signal into a number of scales, every scale representing a particular coarseness of the signal under this study. The procedure of multiresolution decomposition of a signal x[n] is schematically shown in 7. The approximation (A) coefficients are the lowpass representation of the signal and the details (D) are the wavelet coefficients. At each subsequent level, the approximation coefficients are again divided into a coarser approximation (lowpass) and highpass (detail) part. Every stage of this scheme consists of two digital filters and two down samplers. The down sampled outputs of first high pass and low-pass filters give the detail, D1 and the approximation, A1, respectively. The first approximation, A1 is more decomposed and this process is continued as shown in Figs. 9, 10, 11, 12,13, 14, 15 and 16. A ten-second ECG waveform during LBBB and RBBB all along with its detailed coefficients of levels D1 to D6. As the sampling rate o f BBB signal is 360Hz, this signal contains 3600 samples. It can be clearly shown that the BBB activity can be apparent in the wavelet domain, especially in the detail coefficients.

4.2 Performance evaluation parameters

In the present work, hybrid BF-PSO algorithm was tested at single function for 100 population for 25 iterations and the upper/lower bound limits of all parameters are selected based upon the specified range of parameters. The results in terms of mean solution, standard deviation, power spectral density and mean square error are the observed values in Table 3.

From the given results, it has been seen that mean value of 2.8592e−005 , standard deviation (SD) value of 0.0286,

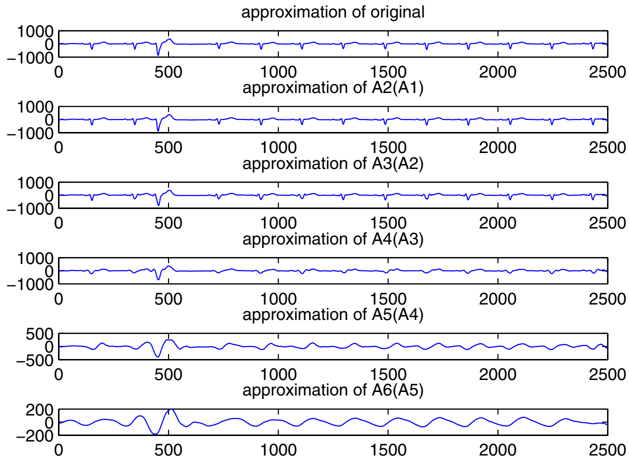

Fig. 9 Approximate coefficients A1–A6 of AF signal

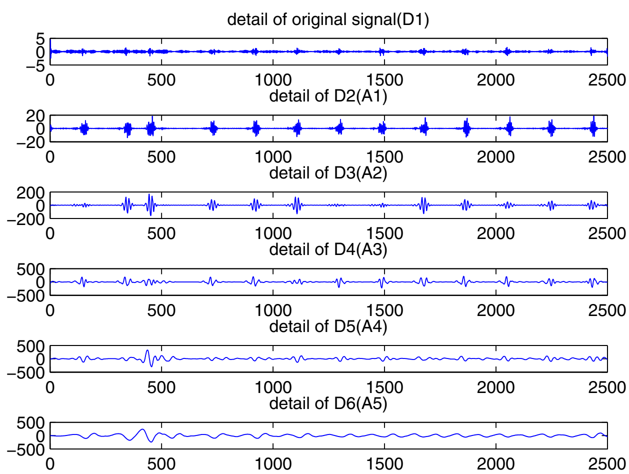

Fig. 10 Detailed coefficients D1–D6 of AF signal

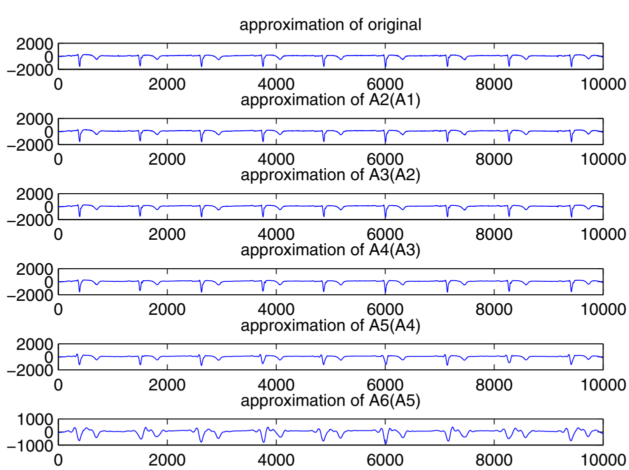

Fig. 11 Approximate coefficients A1–A6 of MI signal

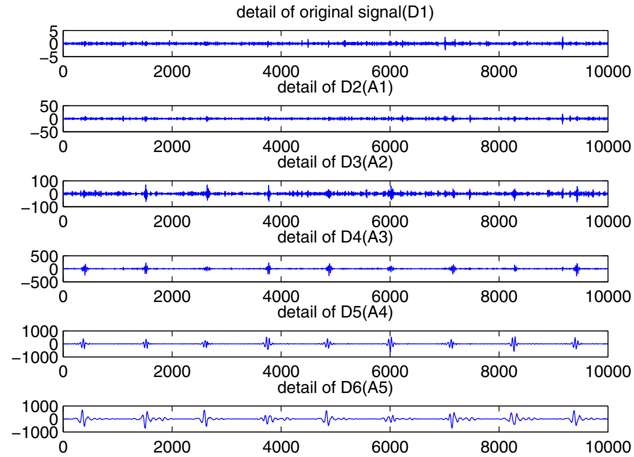

Fig. 12 Detailed coefficients D1–D6 of MI signal

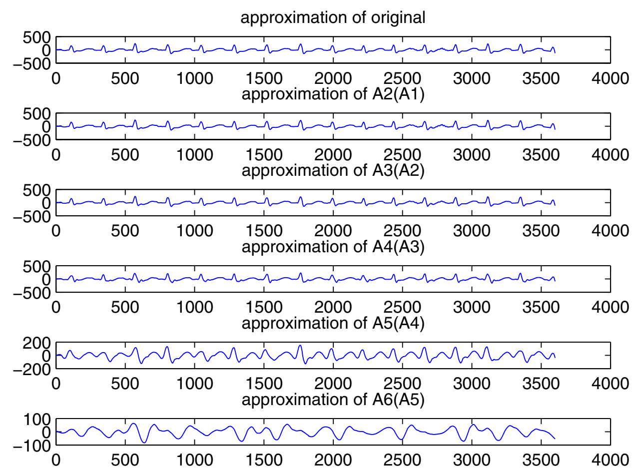

Fig. 13 Approximate coefficients A1–A6 of LBBB signal

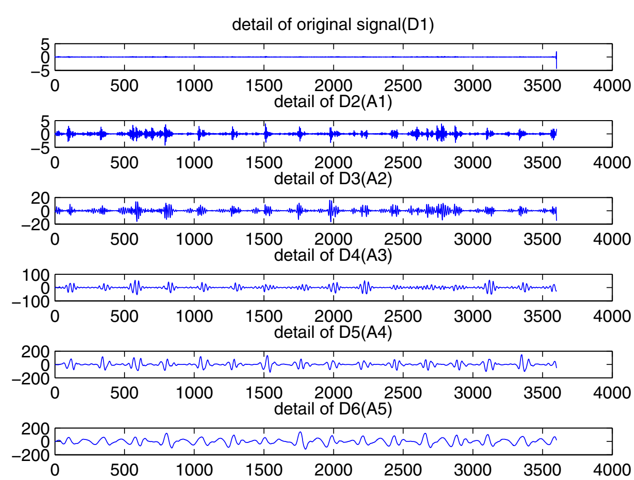

Fig. 14 Detailed coefficients D1–D6 of LBBB signal

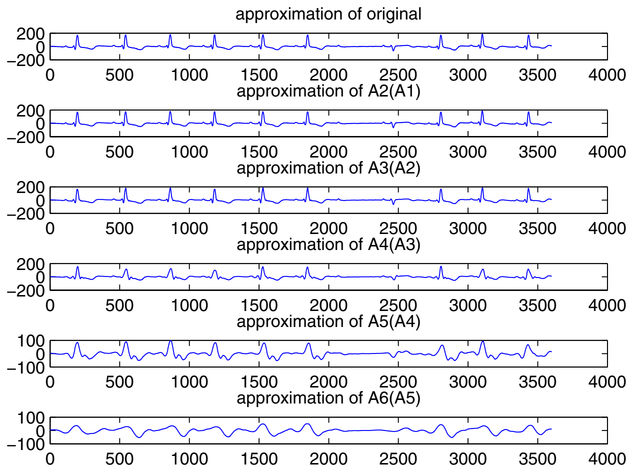

Fig. 15 Approximate coefficients A1–A6 of RBBB signal

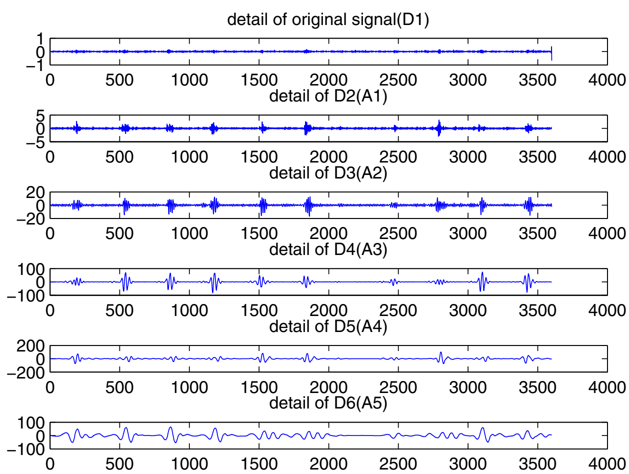

Fig. 16 Detailed coefficients D1–D6 of RBBB signal

Table 2 BF-PSO algorithm parameter values

Input parameter	Value
BFO parameters	
Di-mention of search space (D)	200
Number of bacterium or population (n)	100
Number of chemotaxis steps (Nc)	100
Number of reproduction steps (Nr)	100
Number of iterations	25
PSO parameters	
Inertial weight (w)	1
Acceleration coefficient (C1)	2
Acceleration coefficient (C2)	2

Table 3 Performance Evaluation Parameters

Parameter	MIT-BIH record (16265)
Mean	2.8592e−005
SD	0.0286
PSD	2.9748
MSE	0.0077

Table 4 Comparison of statistics of the convergence time

Researcher	Iterations	Convergence time
Zhang et al. (2010)	100	GA—0.065s PSO—0.023s
Manasrah and Ba Ali (2016)	500	GA—0.869s; PSO—0.761s; GAPSO—0.987s
Shanmugasundaram et al. (2019)	–	GAPSO—178.32s GABFO—161.86s BFOPSO—257.13s
Proposed (BFPSO)	100	GA—0.062s PSO—0.043s BFPSO—0.024s

power spectral density (PSD) is value of 2.9748 and mean square error (MSE) value of 0.0077 on given ECG signal of 200 samples taken from MIT-BIH database.

4.2.1 Comparison on convergence time

The “Convergence time” is defined as the time of converging at a global optimum. Thus, the runs when particles converge at the local optimum are discarded, and the computation time of runs when particles converge at the global optimum are averaged. A comparison of the convergence time among GA, PSO, and BFPSO is shown in Table 4. The convergence time of BFPSO is less than those of GA, PSO and BFO.

4.3 Classification results

The present research work proposed three efficient approaches for ECG classification. The XWT coefficients, WT coefficients and WT + BFPSO coefficients (spectral estimation) are used for feature extraction: These features are classified using KNN, neural networks (NN) and SVM classifiers. The detection accuracy of AF using WT coefficients is found to be 99.1%. The detection accuracy of MI with SVM classifier is 98.9%. The detection accuracy of BBB with SVM classifier is 99.3% as depicted in Table 5. It is clear that the BFPSO with SVM technique for the classification of ECG features have given fast and accurate results compared to the other techniques in the literature survey. We have selected 9193 normal beats used for classification, 3778

Table 5 Classification of WT, XWT and WT + BFPSO features using KNN, NN and SVM classifiers

Classifier	AF (%)			MI (%)			BBB (%)		
	Sen	Spe	acc	sen	spe	acc	sen	spe	acc
XWT + KNN	72.1	70.6	70.2	79.5	78.3	77.5	86.1	83.3	85.5
XWT + NN	80.3	85.1	85.5	81.3	82.5	81.4	91.6	84.3	89.9
XWT + SVM	83.2	87.3	88.1	90.5	87.3	92.2	93.3	91.3	94.5
WT + KNN	61.1	60.6	60.2	69.5	68.3	67.5	86.1	83.3	85.5
WT + NN	85.3	80.1	87.5	86.3	76.5	79.4	80.6	87.3	86.9
WT + SVM	91.1	90.6	**99.1**	87.6	85.1	87.1	94.6	95.5	96.1
WT + BFPSO + KNN	61.1	60.6	60.2	85.5	88.3	87.5	56.1	53.3	56.1
WT + BFPSO + NN	80.3	85.1	87.5	90.3	89.5	90.4	89.6	87.3	89.9
WT + BFPSO + SVM	92.2	88.3	91.1	94.5	92.3	**98.9**	93.3	92.2	**99.3**

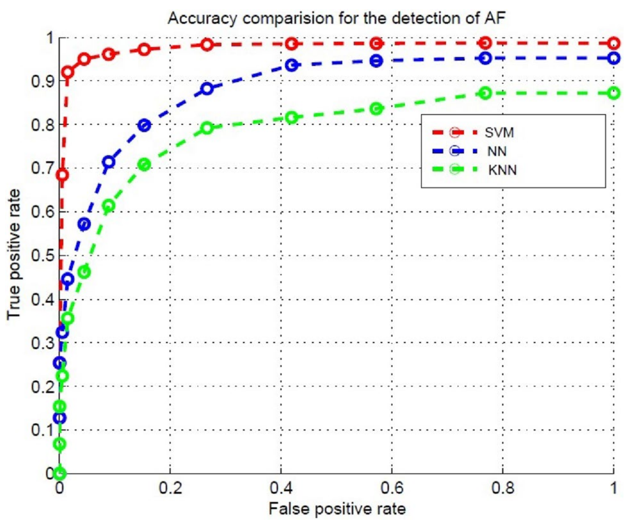

Fig. 17 Detection of AF using three classifiers

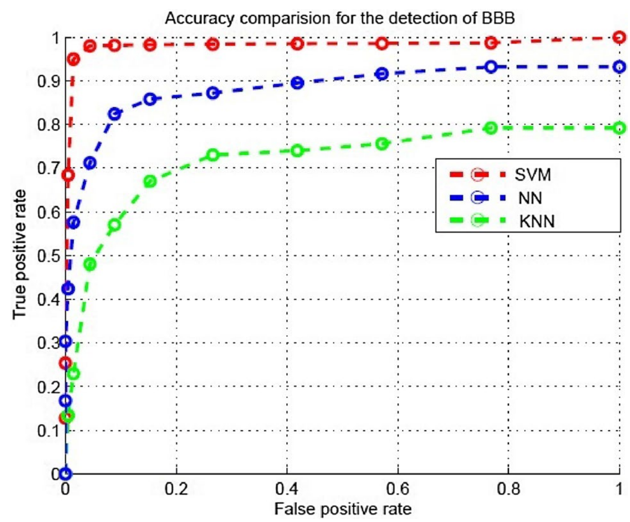

Fig. 18 Detection of BBB using three classifiers

RBBB and LBBB beats and 6068 AF and MI beats user for classification. Hence the total number of beats used for classification are 19,039. Out of which 18,800 are correctly classified and 239 beats are mis-classified.

The specificity is defined as the fraction of correctly classified abnormal beats to the total number of abnormal beats. The sensitivity of an arrhythmia is defined as the fraction of correctly identified normal beats to the total normal beats. The overall accuracy is the division of the total ECG beats correctly classified to the total number beats used for the classification. The performance of feature extraction methods are compared with three classifiers: KNN, NN and SVM. The experimental results state that the proposed BF-PSO optimized features with SVM classifier has greater accuracy for the detection of AF,MI and BBB than other the classifiers.

In this research work, an effort was made to increases classification rate of diseases (AF, BBB and MI). The overall achievement is estimated by the parameters specificity, sensitivity and accuracy in terms of ROC performance curves

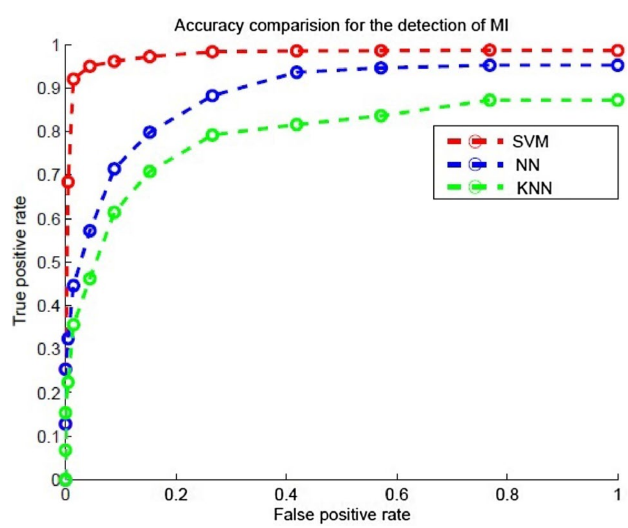

Fig. 19 Detection of MI using three classifiers

shown in Figs. 17, 18 and 19. Table 5 shows the performance of different feature extraction techniques in terms of sensitivity, specificity and accuracy.

5 Conclusion

The present research work proposed BFPSO approach was used for feature optimization. These features are classified using KNN, NN and SVM classifiers. The detection accuracy for AF, MI with SVM classifier are 99.1%, 98.9% and 99.3%. It is clear that SVM have given fast and accurate results compared to the other techniques in the literature. Hence, it has been briefly concluded that the hybrid BF-PSO algorithm is more efficient and effective algorithms to obtain the optimize ECG features.

The future implementations of the work can be as follows:

- This work can be implemented with very large databases in order to increase the detection capability of heart diseases.
- Effective feature extraction and classification techniques can be introduced in the wireless wearable devices for continuous real time monitoring of heart patients.
- The future scope of this research is to use effective training techniques. These effective training techniques can be integrated with machine learning techniques to provide significant classification results with better accuracy. Moreover the training time can be reduced with the help of these training techniques.
- Deep learning methods can be used to improve the classification accuracy.

References

Acharya UR, Fujita H, Oh SL, Hagiwara Y, Tan JH, Adam M (2017) Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. Inf Sci 415:190–198

Angelov P, Kasabov N (2006) Evolving intelligent systems, eIS. IEEE SMC eNewsLetter 15:1–13

Anguluri R, Abraham A, Snasel V (2011) A Hybrid Bacterial foraging-PSO algorithm based tuning of optimal FOPI speed controller. Acta Montan Slovaca 16(1):55

Babaoglu İ, Findik O, Ülker E (2010) A comparison of feature selection models utilizing binary particle swarm optimization and genetic algorithm in determining coronary artery disease using support vector machine. Expert Syst Appl 37(4):3177–3183

Biswas A, Dasgupta S, Das S, Abraham A (2007) Synergy of PSO and bacterial foraging optimization—a comparative study on numerical benchmarks. In: Corchado E, Corchado JM, Abraham A (eds) Innovations in hybrid intelligent systems, vol 44. Springer, Berlin, Heidelberg, pp 255–263

Ceylan R, Özbay Y (2011) Wavelet neural network for classification of bundle branch blocks. Proc World Congr Eng 2(4):1–5

Das S, Biswas A, Dasgupta S, Abraham A (2009) Bacterial foraging optimization algorithm: theoretical foundations, analysis, and applications. In: Abraham A, Hassanien AE, Siarry P, Engelbrecht A (eds) Foundations of computational intelligence, vol 3. Springer, Berlin, Heidelberg, pp 23–55

Debbarma S, Saikia LC, Sinha N (2014) Automatic generation control using two degree of freedom fractional order PID controller. Int J Electr Power Energy Syst 58:120–129

El-Wakeel AS, Ellissy AEEKM, Abdel-hamed AM (2015) A hybrid bacterial foraging-particle swarm optimization technique for optimal tuning of proportional-integral-derivative controller of a permanent magnet brushless DC motor. Electric Power Compon Syst 43(3):309–319

Engin M (2004) ECG beat classification using neuro-fuzzy network. Pattern Recogn Lett 25(15):1715–1722

Faust O, Acharya UR, Adeli H, Adeli A (2015) Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis. Seizure 26:56–64

Fong S, Wong R, Vasilakos AV (2016) Accelerated PSO swarm search feature selection for data stream mining big data. IEEE Trans Serv Comput 9(1):33–45

Gandomi AH, Yang XS, Alavi AH (2013) Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems. Eng Comput 29(1):17–35

Garcia MP, Montiel O, Castillo O, Sepúlveda R, Melin P (2009) Path planning for autonomous mobile robot navigation with ant colony optimization and fuzzy cost function evaluation. Appl Soft Comput 9(3):1102–1110

Hramov AE, Koronovskii AA, Makarov VA, Pavlov AN, Sitnikova E (2015) Mathematical methods of signal processing in neuroscience. In: Wavelets in neuroscience. Springer, Berlin, Heidelberg, pp 1–13

Kaveh A, Chung W (2013) Automated classification of coronary atherosclerosis using single lead ECG. In: 2013 IEEE conference on wireless sensor (ICWISE). IEEE, pp 108–113. https://doi.org/10.1109/ICWISE.2013.6728790

Kim WS, Jin SH, Park YK, Choi HM (2007) A study on development of multi-parametric measure of heart rate variability diagnosing cardiovascular disease. In: World congress on medical physics and biomedical engineering 2006. Springer, Berlin, Heidelberg, pp 3480–3483

Kora P (2017) ECG based myocardial infarction detection using hybrid firefly algorithm. Comput Methods Programs Biomed 152:141–148

Kora P, Kalva SR (2015) Hybrid bacterial foraging and particle swarm optimization for detecting bundle branch block. SpringerPlus 4(1):481

Kumar A (2014) Changing trends of cardiovascular risk factors among Indians: a review of emerging risks. Asian Pac J Trop Biomed 4(12):1001–1008

Kumar KS, Jayabarathi T (2012) Power system reconfiguration and loss minimization for an distribution systems using bacterial foraging optimization algorithm. Int J Electr Power Energy Syst 36(1):13–17

Kumar M, Pachori RB, Acharya UR (2017) Characterization of coronary artery disease using flexible analytic wavelet transform applied on ECG signals. Biomed Signal Process Control 31:301–308

Lee ZJ, Su SF, Chuang CC, Liu KH (2008) Genetic algorithm with ant colony optimization (GA-ACO) for multiple sequence alignment. Appl Soft Comput 8(1):55–78

Lee J, Reyes BA, McManus DD, Maitas O, Chon KH (2013) Atrial fibrillation detection using an iPhone 4S. IEEE Trans Biomed Eng 60(1):203–206

Lehtinen R, Hoist O, Turjanmaa V, Edenbrandt L, Pahlm O, Malmivuo J (1998) Artificial neural network for the exercise electrocardiographic detection of coronary artery disease. In: Proceedings of the 2nd international conference on bioelectromagnetism (Cat. No. 98TH8269). IEEE, pp 57–58

Lewenstein K (2001) Radial basis function neural network approach for the diagnosis of coronary artery disease based on the standard electrocardiogram exercise test. Med Biol Eng Comput 39(3):362–367

Majhi R, Panda G, Majhi B, Sahoo G (2009) Efficient prediction of stock market indices using adaptive bacterial foraging optimization (ABFO) and BFO based techniques. Expert Syst Appl 36(6):10097–10104

Manasrah AM, Ba Ali H (2018) Workflow scheduling using hybrid ga-pso algorithm in cloud computing. Wirel Commun Mobile Comput. https://doi.org/10.1155/2018/1934784

Moeini R, Babaei M (2017) Constrained improved particle swarm optimization algorithm for optimal operation of large scale reservoir: proposing three approaches. Evol Syst 8(4):287–301

Ordóñez-De León B, Aceves-Fernandez MA, Fernandez-Fraga SM, Ramos-Arreguín JM, Gorrostieta-Hurtado E (2019) An improved particle swarm optimization (PSO): method to enhance modeling of airborne particulate matter (PM10). Evol Syst. https://doi.org/10.1007/s12530-019-09263-y

Padmavathi K, Krishna KSR (2014) Myocardial infarction detection using magnitude squared coherence and support vector machine. In: 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom), 7-8 November 2014, Greater Noida, India. IEEE, pp 382–385. https://doi.org/10.1109/MedCom.2014.7006037

Rajasekhar A, Jatoth RK, Abraham A, Snasel V (2011) A novel hybrid ABF-PSO algorithm based tuning of optimal FOPI speed controller for PMSM drive. In: 2011 12th International Carpathian Control Conference (ICCC). IEEE, pp 320–325. https://doi.org/10.1109/CarpathianCC.2011.5945872

Raju M, Gupta MK, Bhanot N, Sharma VS (2018) A hybrid PSO-BFO evolutionary algorithm for optimization of fused deposition modelling process parameters. J Intell Manuf. https://doi.org/10.1007/s10845-018-1420-0

Roeva O, Fidanova S, Paprzycki M (2016) InterCriteria analysis of ACO and GA hybrid algorithms. In: Fidanova S (ed) Recent advances in computational optimization, Studies in Computational Intelligence, vol 610. Springer, Cham, pp 107–126

Schreck DM, Ng L, Schreck BS, Bosco SF, Allegra JR, Zacharias D, Wortzel JV (1988) Detection of coronary artery disease from the normal resting ECG using nonlinear mathematical transformation. Ann Emerg Med 17(2):132–134

Shanmugasundaram K, Mohmed ASA, Ruhaiyem NIR, Mizher MAA, Choo AM, Abdullah SNHS, Razi MJM (2019) Hybrid improved bacterial swarm optimization algorithm for hand-based multimodal biometric authentication system. J ICT 18(2):123–141

Sridhar C, Acharya UR, Fujita H, Bairy GM (2016) Automated diagnosis of coronary artery disease using nonlinear features extracted from ECG signals. In: 2016 IEEE international conference on systems, man, and cybernetics (SMC). IEEE, pp 545–549. https://doi.org/10.1109/SMC.2016.7844296

Tam JH, Ong ZC, Ismail Z, Ang BC, Khoo SY, Li WL (2018) Inverse identification of elastic properties of composite materials using hybrid GA-ACO-PSO algorithm. Inverse Probl Sci Eng 26(10):1432–1463

Vetterli M, Kovačević J, Goyal VK (2014) Foundations of signal processing. Cambridge University Press, Cambridge

Xizhi Z (2008) The application of wavelet transform in digital image processing. In: 2008 International Conference on MultiMedia and Information Technology. IEEE, pp 326–329. https://doi.org/10.1109/MMIT.2008.134

Yang XS (2010) Firefly algorithm, stochastic test functions and design optimisation. arXiv:1003.1409

Younus ZS, Mohamad D, Saba T, Alkawaz MH, Rehman A, Al-Rodhaan M, Al-Dhelaan A (2015) Content-based image retrieval using PSO and k-means clustering algorithm. Arab J Geosci 8(8):6211–6224

Zhang Y, Jun Y, Wei G, Wu L (2010) Find multi-objective paths in stochastic networks via chaotic immune PSO. Expert Syst Appl 37(3):1911–1919