Egyptian Informatics Journal (2017) 18, 55–66

Cairo University

Egyptian Informatics Journal

www.elsevier.com/locate/eij
www.sciencedirect.com

FULL-LENGTH ARTICLE

ABC optimized RBF network for classification of EEG signal for epileptic seizure identification

Sandeep Kumar Satapathy [a,*], Satchidananda Dehuri [b], Alok Kumar Jagadev [c]

[a] *Department of Computer Science & Engineering, ITER, S'O'A University, Bhubaneswar 751030, Odisha, India*
[b] *Department of Information & Communication Technology, Fakir Mohan University, Vyasa Vihar, Januganj, Balasore 756019, India*
[c] *School of Computer Engineering, KIIT University, Bhubaneswar 751024, Odisha, India*

Received 18 December 2015; revised 28 March 2016; accepted 3 May 2016
Available online 9 July 2016

KEYWORDS

Electroencephalography;
Radial basis function neural networks;
Artificial Bee Colony;
Discrete Wavelet Transform

Abstract The brain signals usually generate certain electrical signals that can be recorded and analyzed for detection in several brain disorder diseases. These small signals are expressly called as Electroencephalogram (EEG) signals. This research work analyzes the epileptic disorder in human brain through EEG signal analysis by integrating the best attributes of Artificial Bee Colony (ABC) and radial basis function networks (RBFNNs). We have used Discrete Wavelet Transform (DWT) technique for extraction of potential features from the signal. In our study, for classification of these signals, in this paper, the RBFNNs have been trained by a modified version of ABC algorithm. In the modified ABC, the onlooker bees are selected based on binary tournament unlike roulette wheel selection of ABC. Additionally, kernels such as Gaussian, Multi-quadric, and Inverse-multi-quadric are used for measuring the effectiveness of the method in numerous mixtures of healthy segments, seizure-free segments, and seizure segments. Our experimental outcomes confirm that RBFNN with inverse-multi-quadric kernel trained with modified ABC is significantly better than RBFNNs with other kernels trained by ABC and modified ABC.

1. Introduction

Electroencephalogram is a signal generated in human brain when there is an information flow among several neurons [1]. Human brain contains millions of neurons which are responsible for information flow. Due to this flow of information a human body acts accordingly. A neuron hits another neuron and this process continues for several neurons, due to which a very small amount of electric discharge is generated. This electric signal is quite small in amount and hence it is very difficult to measure the frequency. As a result, there are several electrodes placed on the scalp of those read this electric flow

* Corresponding author.
E-mail addresses: sandeepkumar04@gmail.com (S.K. Satapathy), satchi.lapa@gmail.com (S. Dehuri), alokkumar_j@gmail.com (A.K. Jagadev).

Peer review under responsibility of Faculty of Computers and Information, Cairo University.

ELSEVIER Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.eij.2016.05.001
1110-8665 © 2016 Production and hosting by Elsevier B.V. on behalf of Faculty of Computers and Information, Cairo University.

and frequency is recorded by a machine [2]. These signals are very transient in nature, because the frequency of these signals changes rapidly with respect to time. Therefore, the signal analysis [3] should be made carefully to analyze it properly. First of all, these signals must be recorded by some means before analyzing. There are 10–20 international standard for placement of electrodes on human scalp to record EEG signals. The electrodes are placed on different regions of scalp, such as Frontal, Parietal, Occipital, and Temporal. Through these electrodes the EEG signals can be recorded and processed in a machine [4,5]. Generally, these machines will draw a graph of the recorded signal, which can be later analyzed by a medical professional. Secondly, the analysis of these signals is quite necessary for certain applications in medical science. Usually, Fast Fourier Transform is used for continuous signals which are not suitable for analysis of EEG signal. Discrete Wavelet Transform is one of the most efficient methods for analysis of these kinds of signals [6]. It is a signal decomposition technique that uses two types of filters, low and high pass filter to divide the signal into low and high frequency bands. After this decomposition several features are extracted for each signal. This feature set can be used in further processing. After analyzing these signals we can discover valid and potentially useful information about human brain. In the sequel, this may help in identifying different types of human brain disorder diseases. Such a disease is known as Epilepsy [7]. It is a disorder in human brain activity due to abnormal EEG signal flow. The time period of this disorder attack is called as epileptic seizure. This seizure may occur for a small period of time during which a very high frequency of EEG is generated. The visual examination is strenuous; hence, numerous studies are made for development of semi-automatic seizure detection technique. Hence, the third and most important aspect of this research is the classification of EEG signal [8]. Classification is one of the fundamental tasks of data mining. It is a process of assigning an unlabeled pattern into a specific pre-defined class by constructing a model from the training patterns and subsequently validating in a test set. There can be two class or multi-class problem. For our study, we are going to classify the EEG signal into two groups that is either epileptic or normal [9–11]. Therefore, it is a two class problem. There are a lot of classification algorithms available, among which machine learning algorithms are most efficient and capable enough to classify data samples with high accuracy. Machine learning is a technique of constructing a model for doing a specific task by training the model with some previously known instances. This is just like a small kid learning different activities by observing the actions taking place near him. The different machine learning techniques used for classification are Artificial Neural Network [12–19], Support Vector Machine [2], and Radial Basis Function Neural Network [20], etc. All these techniques have their own applications in different areas. In [21] we have done an empirical analysis on application of these different machine learning techniques in classification of EEG signal for epileptic seizure identification. From this it has been concluded that SVM and PNN are very efficient. But the simplified architecture of RBF neural network grabs more attention for enhancing its accuracy in classification of EEG signal [20,22,23]. Compared to other techniques RBFNN has a simple architecture consisting of a single hidden layer along with input and output layers.

In this study, we have highly emphasized on the performance enhancement of RBFNN. For this, a novel algorithm for training RBFNN using Artificial Bee Colony (ABC) [20] has been proposed. Rest of the Sections in this paper is set out as follows. Section 2 describes about the materials and different methods used to carry out this research work and the proposed algorithm for RBFNN training. Section 3 describes our proposed work. In Section 4, the experimental studies have been carried out with an analysis of the outcomes. Section 5 concludes the article with lots of research issues.

2. Materials and methods

Data selection and preparation is one of the key subsections of this section. The basic approaches such as RBFNNs and ABC are discussed in Sections 2.3 and 2.4, respectively.

2.1. Data selection

For this research work, we have collected EEG data for epileptic seizure identification from publically available online resource. This is an openly available source of data for EEG used by many researchers for their research work. It is mainly categorized into five types set A, B, C, D and E. Each set contains 100 single channel EEG segments. Each segment is of 23.6 s duration. All these data have been prepared by removing artifacts due to eye or muscle movements. Sets A and B have been collected from healthy patients having eyes open and closed respectively. Sets C, D, and E have been collected from epileptic patients, but C and D are recorded in seizure-free activity, whereas set E contains seizure activity.

2.2. Discrete Wavelet Transform

Basically, all types of signals are analyzed in time domain with their amplitudes. Signals such as EEG and ECG, are generally collection of amplitudes with respect to time. If we plot this data it can give a shape from which the pathological condition of a patient can be observed. If there is any significant deviation in shape it can be shown and observed properly by visualizing the graph. But sometimes it is necessary to get the frequency content of a signal for proper and accurate analysis of a signal. It can be done by using any transformation technique such as Fourier Transform. But again the disadvantage of this is, it is not so effective for transient signals such as EEG as EEG signals have very uncertain and rapidly changing frequency. So, it is very difficult to analyze effectively. As a result, we need some other transformation technique such as Wavelet Transformation for analysis of EEG signals. This is just a new perspective for analysis and processing of data. The basic idea behind this technique is to use a scale for analysis. This wavelet transform can be divided into two categories such as Continuous Wavelet Transform (CWT), and Discrete Wavelet Transform (DWT). CWT was first developed as an alternative to Short Time Fourier Transform (STFT). Here, the product of the signal with a function that is wavelet function was calculated. This transform was then calculated for different time domain. It is defined and is given as in Eq. (1):

$$\mathrm{CWT}(a,b) = \int_{\infty}^{\infty} x(t) * \varphi_{a,b}^{\nabla}(t)dt \tag{1}$$

where $x(t)$ represents the original signal. a, b represents the scaling factor and translation along the time axis respectively. The ∇ symbol denotes the complex conjugation and $\varphi^{\nabla}_{a,b}$ is calculated by scaling the wavelet at time b and scale a (as given in Eq. (2)).

$$\varphi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\varphi\left(\frac{t-b}{a}\right) \tag{2}$$

where $\varphi_{a,b}(t)$ represents the mother wavelet. In CWT, it is assumed that the scaling and translation parameter a and b change continuously. But the main disadvantage of CWT is, the calculation of wavelet coefficients for every possible scale can result in a large amount of data. It can be overcome by the help of DWT. It analyzes the signal at different frequency band by decomposing the signal into a set of high and low pass filters called as *approximation* and *detailed coefficients*. These coefficients can be calculated by using the wavelet toolbox available in MATLAB. Using the predefined functions available inside this toolbox, we can easily extract the features of EEG signal. Some of the snapshots of wavelet GUI toolbox of MATLAB are given in Figs. 1–4. From the data available at [27], a rectangular window of length 256 discrete data was selected to form a single EEG segment. The wavelet coefficients have been computed using Daubechies of order four. This technique was found to be more suitable because of its smoothing features which are more appropriate to detect changes in EEG signal. For our work, the original signal has been decomposed as four detailed coefficients (d_1, d_2, d_3, d_4) and four approximation coefficients (a_1, a_2, a_3, a_4). For simplicity, all the approximation coefficients are ignored except the one in the last step i.e. a_4. Hence, the signal is decomposed into five segments by using DWT. In this work, for four detailed coefficients we get 247 coefficients (129 + 66 + 34 + 18) and eighteen for approximation coefficients. Several statistical features have been extracted. But for this study, four important features were taken into considerations:

I. Maximum of wavelet coefficients in each sub-band.
II. Minimum of wavelet coefficients in each sub-band.
III. Mean of wavelet coefficients in each sub-band.
IV. Standard deviation of wavelet coefficients in each sub-band.

Therefore, for five coefficients all total twenty features have been extracted and the dataset has been constructed.

2.3. Radial basis function neural networks

RBFNN is one of the simplest form of Neural Network consisting of exactly three layers namely input, hidden, and output layer (as shown in Fig. 5). The restriction of only three layers makes it simplest and somehow efficient neural network architecture. The idea of RBFNN has been derived from function approximation. An RBF network positions one or more RBF neurons in the space described by the predictor variables. This space has as many dimensions as there are predictor variables. The Euclidean distance is computed from the point being evaluated to the center of each neuron. The radial basis function is so named because the radius distance is the argument to the function. Output of RBFNNN depends on the distance of the input from a given stored vector. For our work, N number of input neurons, m number of hidden neurons and one output neuron are taken. There are several kernel functions used in RBFNN, such as Gaussian, Multi-quadric, and Inverse Multi-quadric. Each of the functions has its own benefits depending on the data domain they are used in. Based on

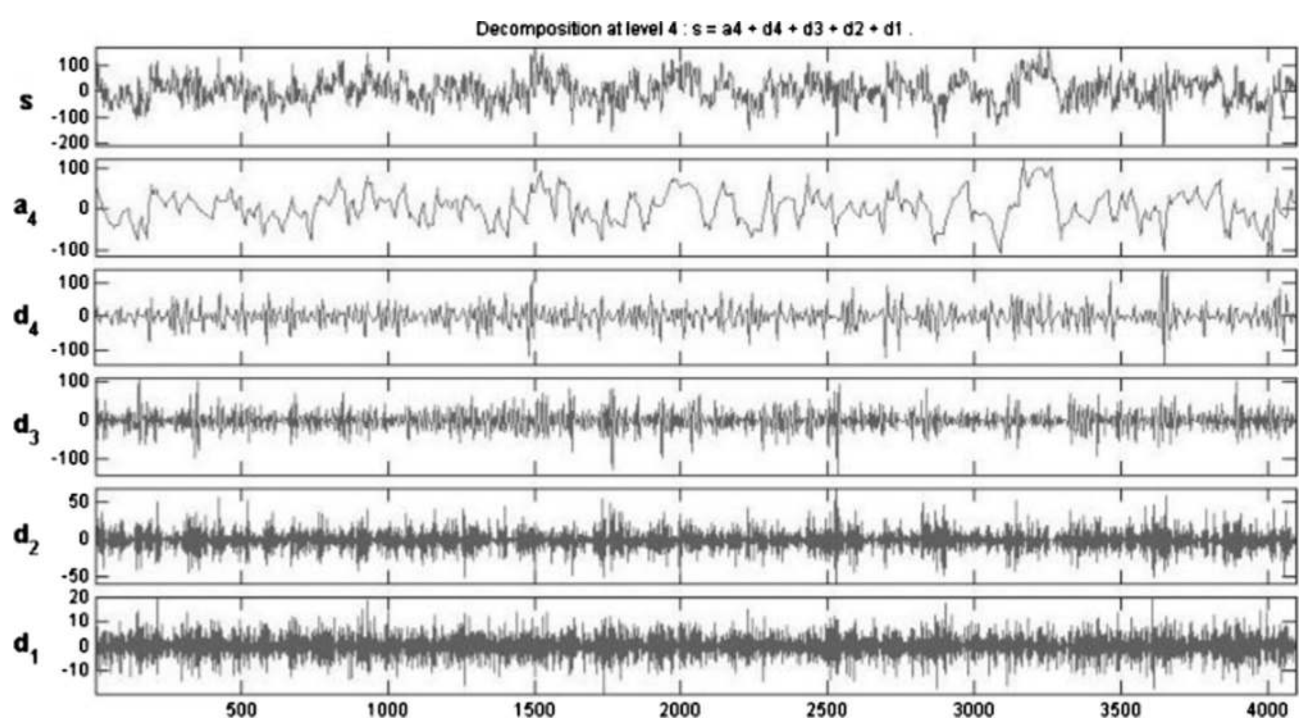


Figure 1 Single channel EEG signal decomposition of set A using db-2 up to level 4.

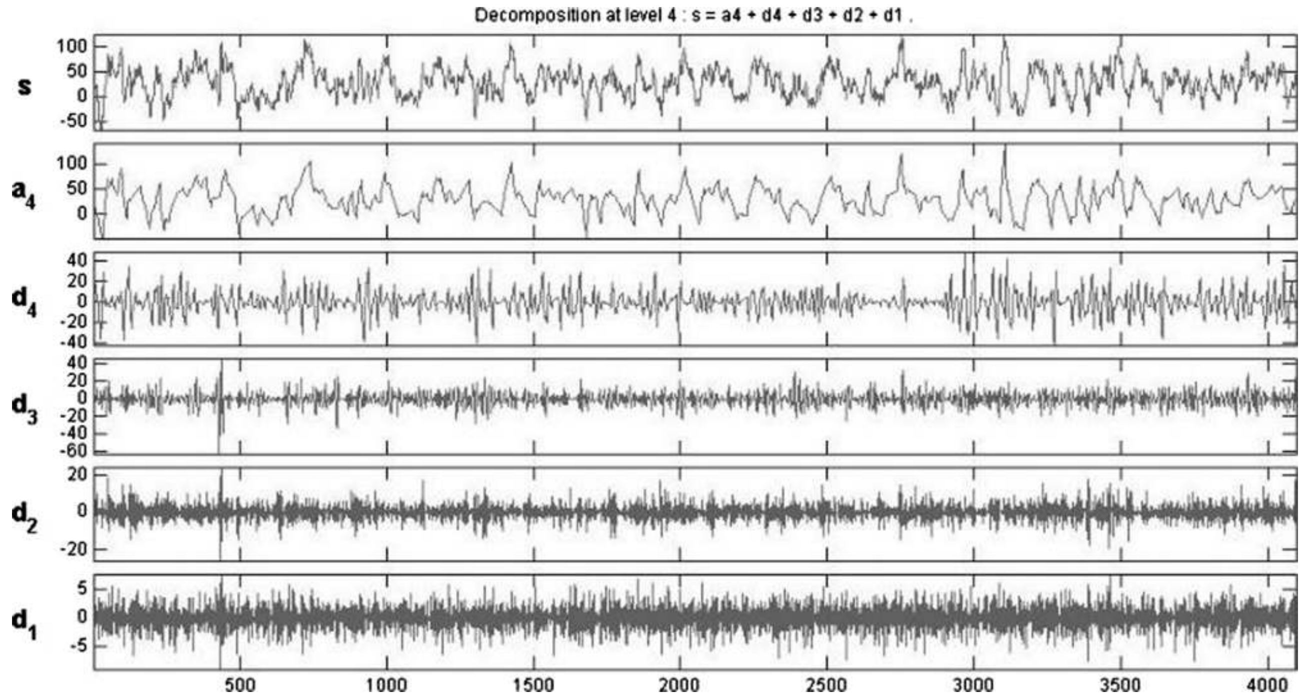


Figure 2 Single channel EEG signal decomposition of set D using db-2 up to level 4.

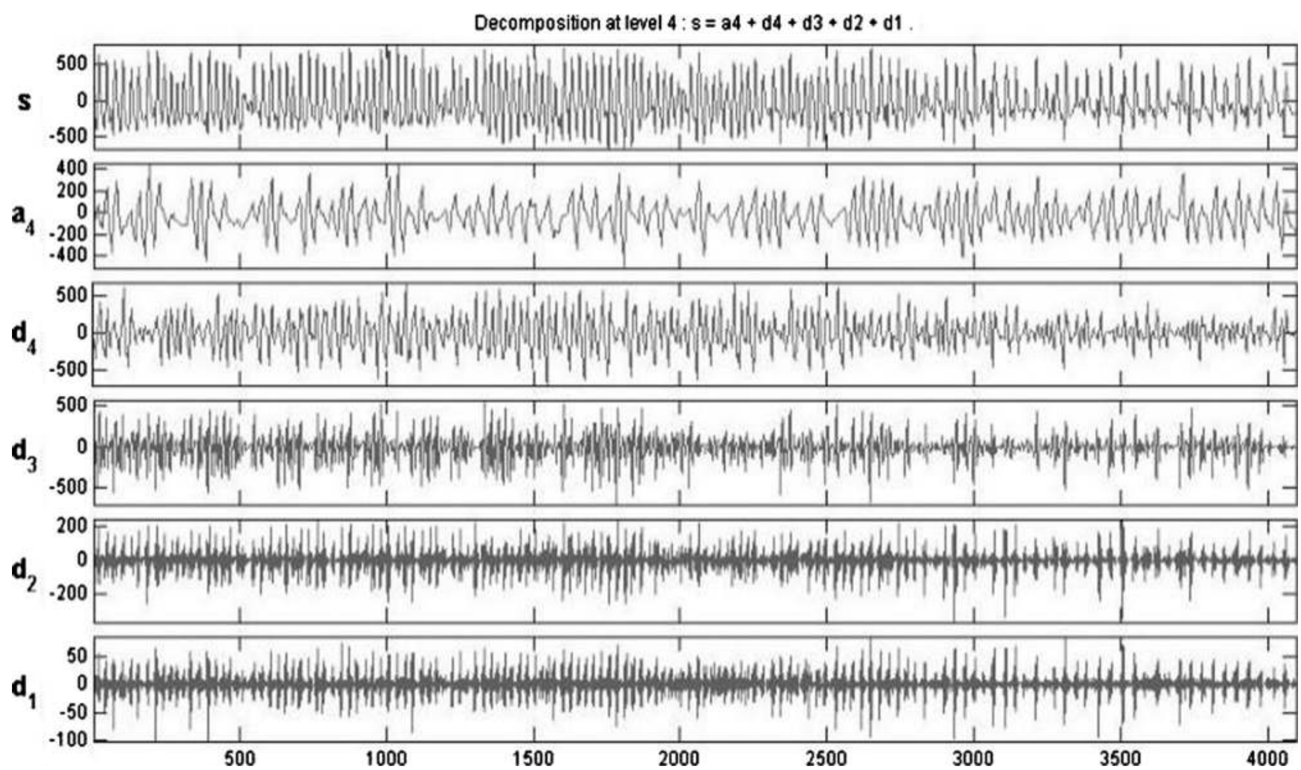


Figure 3 Single channel EEG signal decomposition of set E using db-2 up to level 4.

the recommendation of our previous research, we used to verify the performance of Gaussian, Multi-quadric, Inverse Multi-quadric basis function in RBFNNs for identification of epileptic seizure, but it was found that the performance of Inverse Multi-quadric is pretty higher than the performances of the other two.

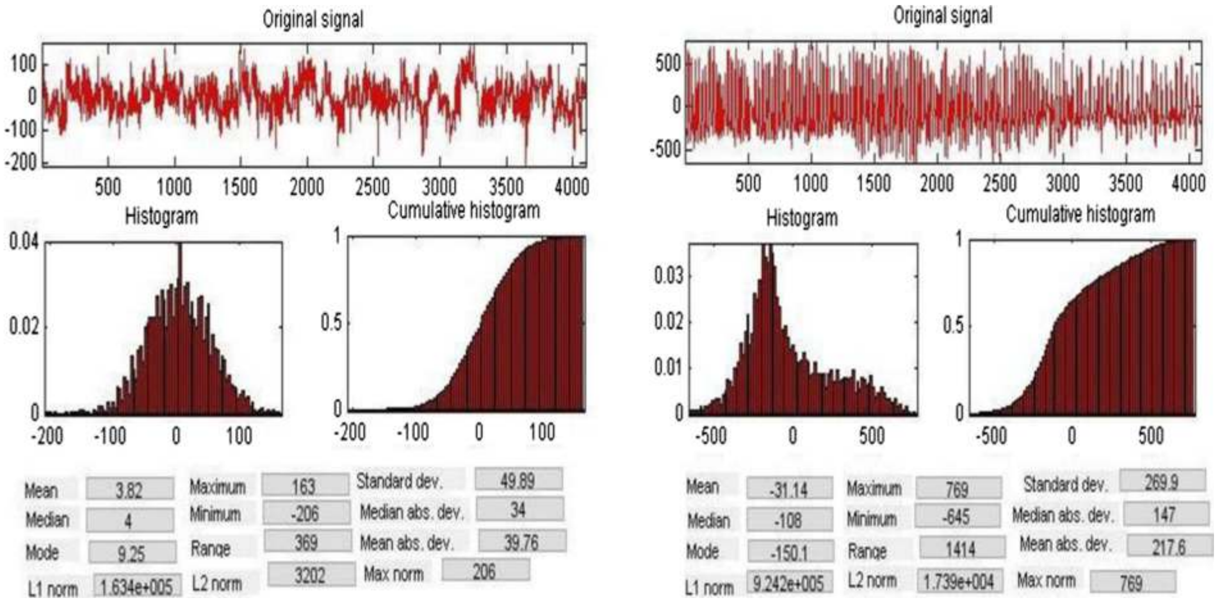


Figure 4 Statistical feature extraction from signals after decomposition.

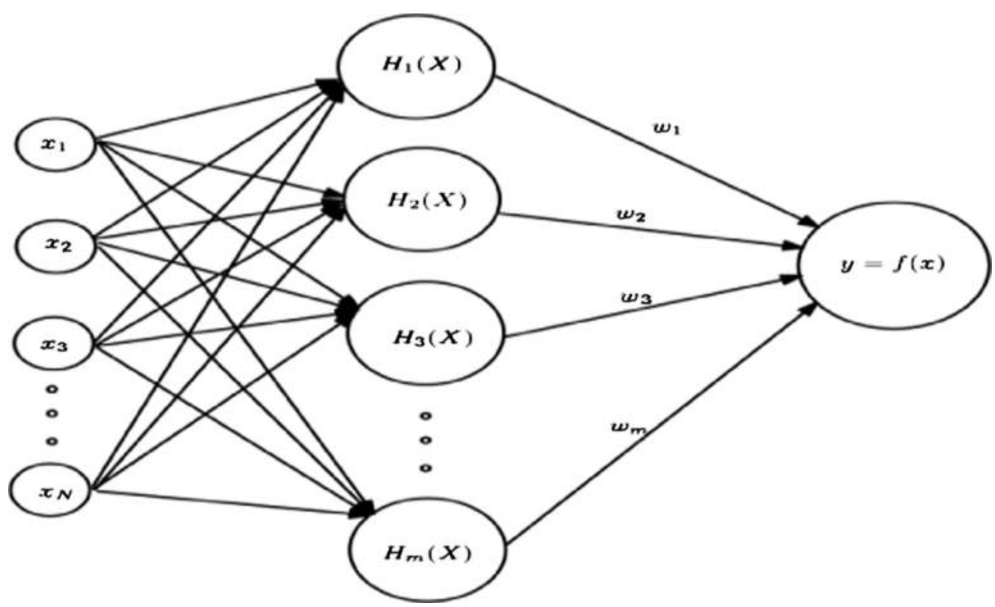


Figure 5 Architecture of RBFNNN.

Table 1 Parameter description for RBFNNN.

Parameter symbol	Description	Dimension
N	Number of input vectors	200 or 300
D	Desired output vector	200×1 or 300×1
M	Number of hidden neurons	40
W	Weight vector	40x1
N	Number of input neurons	20
X	Input vector	1x20
C	Center matrix	40x20
Σ	Spread vector	40x1

The different symbols and dimensions used in the above figure are as follows (as given in Table 1):

The above symbols can be further described as follows:

Input Vector $(X) = \{x_1, x_2, \ldots x_N\}$
Hidden Neurons $= \{H_1(X), H_2(X), H_3(X), \ldots H_m(X)\}$
Weight Vector $(W) = \{w_1, w, \ldots w_m\}$

$$\text{Center Matrix } (C) = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1N} \\ C_{21} & & & C_{2N} \\ \vdots & & & \vdots \\ C_{m1} & C_{m2} & \cdots & C_{mN} \end{bmatrix}$$

Spread Vector $(\sigma) = \{\sigma_1, \sigma_2, \ldots \sigma_m\}$

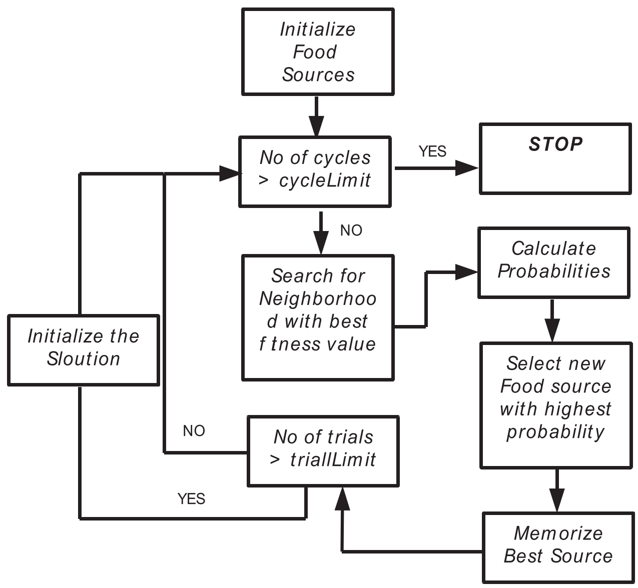


Figure 6 Working procedure of ABC algorithm.

Hence the output of RBFNNN can be defined as

$$\text{RBF Output } (y) = \sum_{j=1}^{m} w_j * H_j(x) \tag{3}$$

where $H_j(x)$ can be any one of the following:

$$\text{Gaussian Function, } H_j(x) = \exp\left(\frac{-\|X - C_j\|^2}{\sigma^2}\right) \tag{4}$$

$$\text{Multi-quadric function, } H_j(x) = \sqrt{\left((X - C_j)^2 + \sigma^2\right)} \tag{5}$$

$$\text{Inverse multi-quadric function, } H_j(x) = \frac{1}{\sqrt{((X - C_j)^2 + \sigma^2)}} \tag{6}$$

$$\text{where, } (X - C_j)^2 = -\sum_{k=1}^{N} (x_k - c_{jk})^2 \tag{7}$$

To measure the performance of training algorithms, error is calculated by finding the difference between desired output and actual output. Hence, the Mean Square Error Function can be defined as (given in Eq. (8)),

$$\text{MSE}(c, \sigma, w) = \frac{1}{n}\left(\sum_{i=1}^{n} d_j - \sum_{j=1}^{m} w_j * H_j(x)\right)^2 \tag{8}$$

2.3.1. Learning of RBFNN

Learning or training of a network is a process by which it adapts to the environment by adjusting few parameters. For RBFNN, to get the desired output for a given input there are mainly three adjustable parameters, such as Center, Spread, and Weight. There are several learning algorithms proposed by several researchers among which Gradient Descent approach is the most common. This is a first order derivative based optimization algorithm for finding local minimum of a function. According to Eq. (8), the error can be calculated by finding the difference between desired and actual output. Then the partial derivative of this error with respect to weight and center can be calculated to adjust the parameter with minimizing the error. The formula of gradient descent is given as follows (as shown in Eq. (9)):

$$w_i = w_i - \eta \frac{\partial E}{\partial w_i}, \quad c_{ij} = c_{ij} - \eta \frac{\partial E}{\partial c_{ij}} \tag{9}$$

where η is the learning parameter or step size. We have performed several experimental evaluations by considering different η values between 0.5 and 1.0. The detailed results are given in the next section. There are also several other learning techniques such as Particle Swarm Optimization (PSO) [24], Differential Evolution (DE) [25], and Genetic Algorithm (GA) [26]. Basically, RBF networks are used in many applications because of its architectural simplicity and requirement of less number of adjustable parameters. Therefore, to employ the RBFNNN in the relevance of EEG classification, we require some supplementary techniques for improving its performance. This can be done by integrating optimization techniques with the training methods. There are several optimization techniques available such as PSO, ABC, and GA. Yet again, we opt for ABC optimization technique owing to its requirement for less number of adjustable parameters and its capability to produce global optimal solutions. In this study, we have proposed a new innovative training algorithm for RBFNN based Artificial Bee Colony (ABC) optimization algorithm. The different parameters such as center, spread and weight are trained by using ABC optimization algorithm. This is biologically inspired algorithm from the behavior of artificial bees. It has been explained in the next section.

2.4. Artificial Bee Colony algorithm

Artificial Bee Colony (ABC) is a swarm intelligence technique developed by Dervis Karaboga in the year of 2005. Its main aim is to optimize numerical problems. It has been motivated from the foraging behavior of honeybees. The basic nature or intelligence of a honeybee can be used for solving many real life problems. Honeybees are one of the interesting swarms

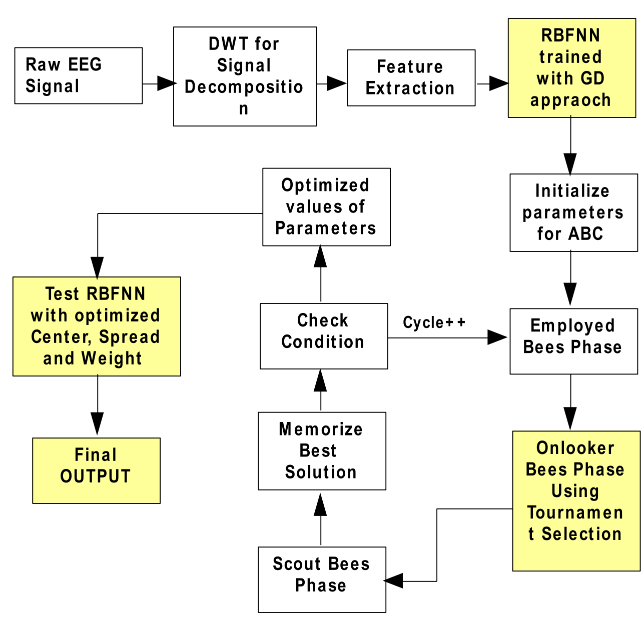


Figure 7 Model of classification using RBFNNN with ABC algorithm.

in nature. They have the skills like photographic memories, space-age sensory, and navigation systems. Honeybees are social insects that live in colonies (as shown in Fig. 6).

In ABC algorithm there are mainly two types of bees, such as employed and unemployed bees. Unemployed bees can be again categorized as onlookers' bees and scouts bees. The initializations of food sources are done using the following formula:

$$x_{mi} = l_i + rand(0,1) * (u_i - l_i) \tag{10}$$

The employed bees search for a new food source having more nectar within the neighborhood of the food source in their memory. The neighbor food sources can be selected by using the following formula (as shown in Eq. (11)):

$$v_{mi} = x_{mi} + \varphi_{mi}(x_{mi} - x_{ki}) \tag{11}$$

where m = number of solutions, i, k = number of parameters to optimize, and φ_{mi} is a random number. After selecting the neighborhoods, their fitness can be calculated using a fitness function. The fitness value of a solution can be calculated as follows (as shown in Eq. (12)):

$$fit_m(\vec{x}_m) = \begin{cases} \frac{1}{1+f_m(\vec{x}_m)}, & \text{if } f_m(\vec{x}_m) \geqslant 0 \\ 1 + \text{abs}(f_m(\vec{x}_m)), & \text{if } f_m(\vec{x}_m) < 0 \end{cases} \tag{12}$$

The bees which are waiting in the dancing area for taking decision on selecting a food source are called as onlooker bees. They select a food source depending on the probability of fitness values provided by employed bees according to the following formula (as shown in Eq. (13)):

$$P_i = \frac{fit_i(\vec{x}_i)}{\sum_{i=1}^{FS} fit_i(\vec{x}_i)} \tag{13}$$

Onlooker bees visit the food source that they select and identify a nearby modified source. They evaluate and choose between the original and new source. The employed bees whose sources were abandoned become scouts and go in search of new food sources. The scout discovers a new food source by employing Eq. (11), where *rand* is a random number between 0 and 1. The algorithm avoids getting into local optimum by having the scouts perform a random global search for new food sources.

3. Our proposed method

Our study work, mainly focuses on classifying epileptic seizure patients vs. non-seizure patients by suitably trained RBFNNs. The trained RBFNN is developed by combining the best attributes of gradient descent trained RBFNN and modified ABC. Initially, we adopt gradient descent approach to train the RBFNN and then the trained parameters such as centers, spreads, weights, are feeding as the seed points of the ABC and modified ABC. The optimized parameters set up the final architecture of RBFNN to assign a class label to sample with no class label. The detailed flowchart of the proposed model is given in Fig. 7.

Once the raw EEG signal is collected from the source, it should first be analyzed to discover the hidden characteristics or features of these signals. This can be carried out using DWT technique, which decomposes the signal into several levels, thus extracts different statistical features. The EEG signal comes with 5 different sets (A, B, C, D, and E). Therefore, the experimental work is divided into three parts. First is consisting of A & E, the second is set D & E, and the third is a collection of A & D with E. These datasets are now ready for the classification work.

Here, we have taken three prominent kernels (discussed in Section 2.3) for the nodes of the hidden layer of RBFNN. These three kernels play the pivotal role in addition to the novel training algorithms while classifying the EEG signals. By considering each individual kernel, RBFNN has been trained with Gradient descent approach and then successively trained with the ABC. Then these intermediate values of different parameters of RBFNN will be used to initialize the solution vector for ABC algorithm. By using this algorithm, the optimal values of center, width, and weight will be calculated. Here the objective function is taken as the Mean Square Error as given in Eq. (8). With an objective of minimizing the error, ABC starts initializing the solutions (consists of three parameters, such as center, width, and weight) and then repeats the loop with required steps up to several runs/the limit, and the parameters will be optimized.

The ABC algorithm will proceed in three different phases, Employed bees, Onlooker bees, and Scout bees. To make the selection of onlooker bees easy and competitive, we replace the roulette wheel selection mechanism by binary tournament. The inspiration of adopting this mechanism in ABC came from selection mechanism of genetic algorithms, in which randomly selected pair of bees will compete among each other to be selected depending on their fitness value. The detailed pseudo-code of proposed method is given below.

Pseudo code for proposed method
Input: Preprocessed EEG dataset for epileptic seizure Identification.
Output: Class label prediction

Step1: Initialize and setup the parameters for RBFNNN
Step2: Load the training sample for RBFNNN and train the network
Step3: Initialize and setup parameters for ABC algorithm with Gradient Descent learning approach up to certain number of iterations.
Step4: Initialize solution (Center, Width and Weight) for each food source in ABC algorithm to values of Center, width and weight values found in Step2.
Step5: Find the fitness value using Eq. (12) with the objective function defined in Eq. (8).
Step6: Set cycle=0 and for each employed bee set trials = 0
Step7: Repeat Step8 to 15 until cycle < cycleLimit
  Step8: For each employed bee find the neighbor bee by using Eq. (11)
  Step9: Find the fitness value and make a greedy selection.
  Step10: If the solution is selected set trials = 0, Otherwise trial ++
  Step11: Select an onlooker bee using tournament selection
  Step12: Find the fitness value and make a greedy selection.
  Step13: If the solution is selected set trials = 0, Otherwise trial ++
  Step14: Memorize the best solution found so far
  Step15: If for any solution trials = trialLimit
    Step16: Abandon the solution (Scout bee)
    Step17: Randomly initialize the solution and go to Step7.
Step18: Set the RBF network by taking the optimized parameters (Center, Width and Weight) found in above steps.
Step19: Test the network with EEG test samples.

4. Experimental study

For this study, five sets of EEG signals for Epileptic seizure identification have been collected from publicly available source [27]. There are three combinations of these sets taken for experimental study, that is set A & E (Experiment 1), set D & E (Experiment 2) and set A + D & E (Experiment 3). All these three datasets are first taken for classification using RBFNN with Gradient Descent Learning algorithm. This algorithm is evaluated by taking different values of learning parameter (η) from 0.5 to 1.0. By different experimental evaluation we found that for EEG dataset the Gaussian and Inverse-multi-quadric basis functions outperform as compared to Multi-quadric function. After that a deep research has been done to enhance the performance of RBFNN using ABC algorithm.

4.1. Environment and parameter setup

For DWT of EEG signal we have used the MATLAB toolbox for wavelet transform. After this all other programming codes for entire experimental work have been designed using Java platform (JDK 1.8 with Eclipse Luna IDE). For ABC algorithm there are several parameters that have been set initially as given below:

a. Colony size = 40 (that is number of employed bees + onlooker bees)
b. No. of food sources = 20 (colony size/2)
c. maxLimit = 100 (number of times a food source can be improved)
d. maxCycle = 50 (number of cycles for foraging)

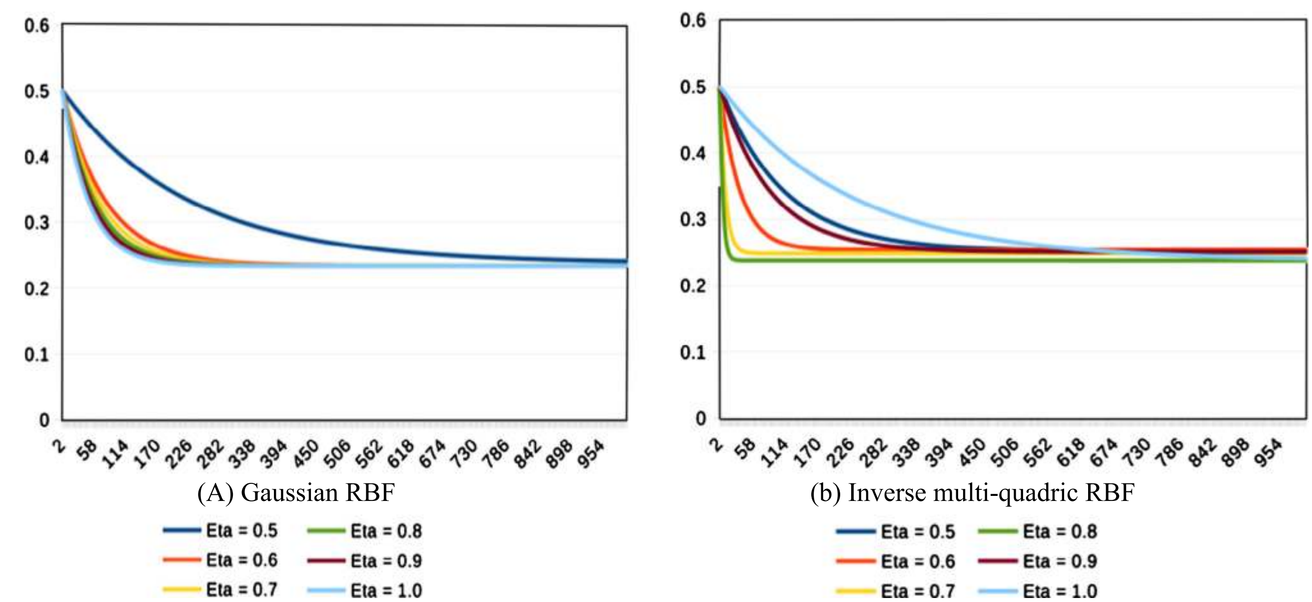


Figure 8 Experiment 1 (A & E) MSE graph for gradient descent approach with varying η value.

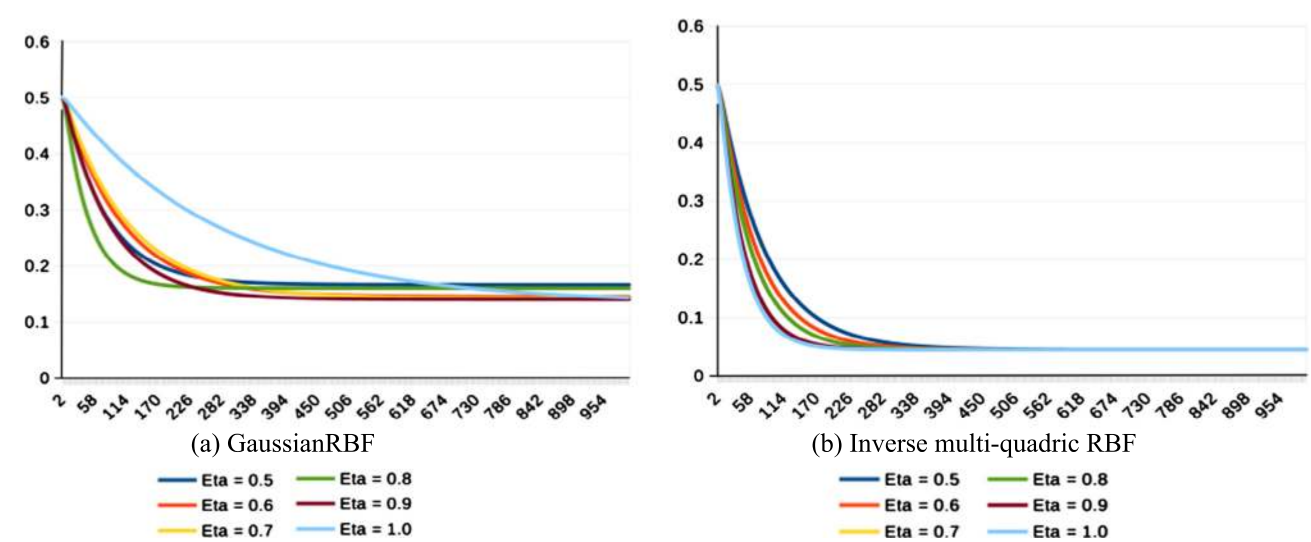
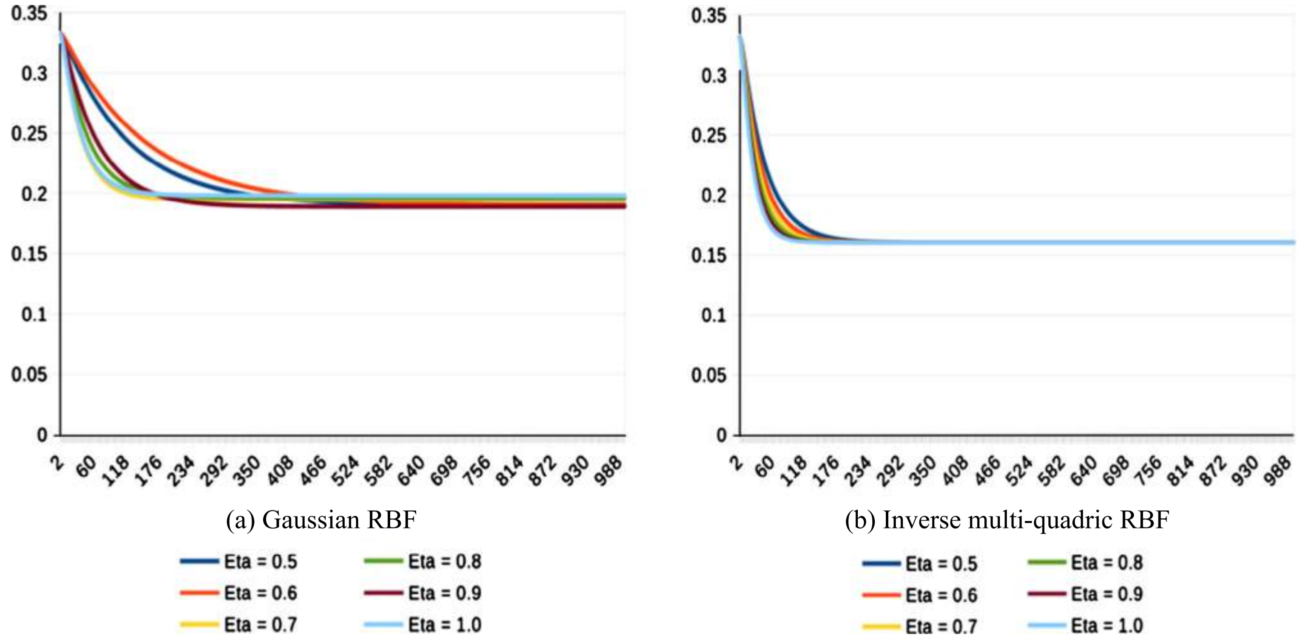
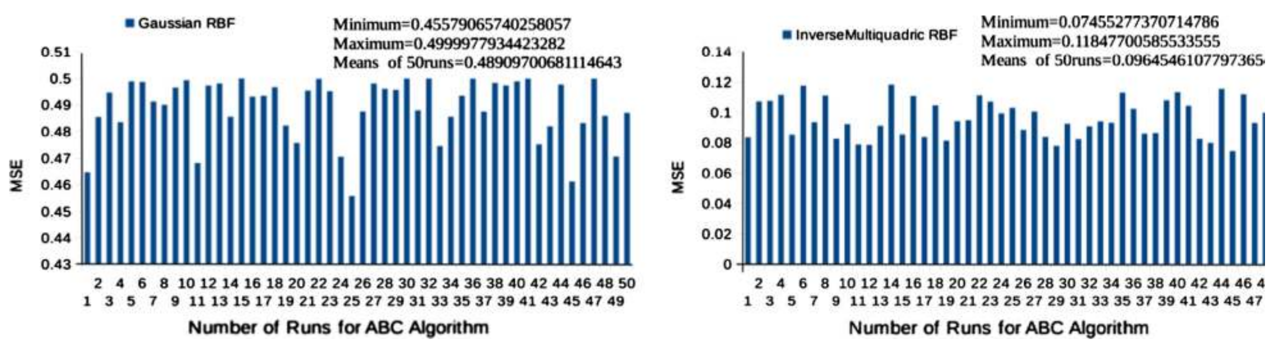


Figure 9 Experiment 2 (D & E) MSE graph for gradient descent approach with varying η value.



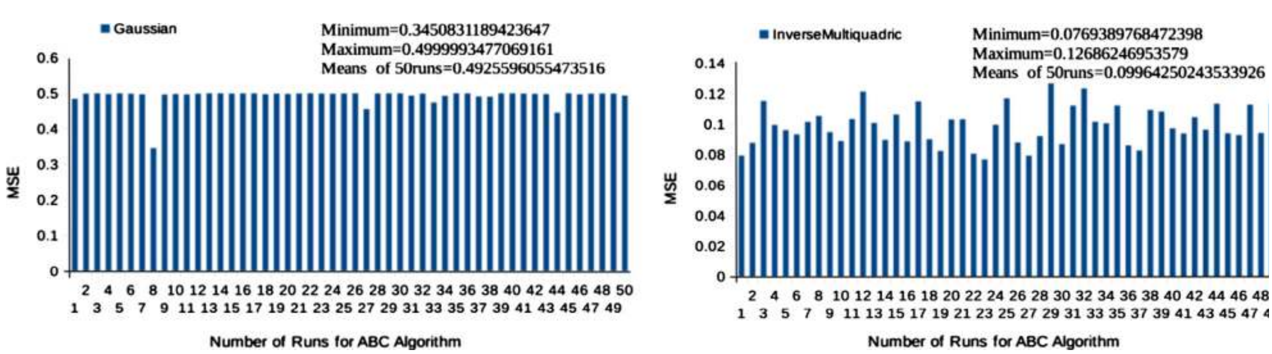
(a) Gaussian RBF (b) Inverse multi-quadric RBF

Figure 10 Experiment 3 (A + D & E) MSE graph for gradient descent approach with varying η value.



(a) GaussianRBF (b) Inverse-multi-quadric RBF

Figure 11 Experiment 1 (set A & E) MSE graph for ABC trained RBF.



(a) GaussianRBF (b) Inverse-multi-quadric RBF

Figure 12 Experiment 2 (set D & E) MSE graph for ABC trained RBFNNN.

e. Number of parameters to optimize = 880 (number of center parameters + number of spread parameters + number of weight parameters)
f. $lb = -1$, $ub = +1$ (*lb*-lower bound and *ub*-upper bound for parameters)
g. Fitness function, $f(c,\sigma,w)=\frac{1}{n}\left(\sum_{i=1}^{n}d_j-\sum_{j=1}^{m}w_j*H_j(x)\right)^2$, where for Gaussian function $H_j(x)$ is given in Eq. (4), Multi-quadric function $H_j(x)$ is given in Eq. (5) and Inverse multi-quadric function $H_j(x)$ is given in Eq. (6).

Table 2 Performance Comparison between GD learning and ABC learning with Inverse-multi-quadric RBFNN.

Experiments	Specificity		Sensitivity		Accuracy	
	RBFNN with GD	RBFNN with ABC	RBFNN with GD	RBFNN with ABC	RBFNN with GD	RBFNN with ABC
Set A & E	84.1	85.7	65.7	66.4	71.5	72.5
Set D & E	100.0	100.0	80.0	88.5	87.5	93.5
Set A + D & E	76.8	95.6	63.2	66.0	73.3	81.6

Table 3 Performance comparison between ABC learning and modified ABC learning with Inverse-multi-quadric RBFNN.

Experiments	Specificity		Sensitivity		Accuracy	
	RBFNN with ABC	RBFNN with MABC	RBFNN with ABC	RBFNN with MABC	RBFNN with ABC	RBFNN with MABC
Set A & E	85.7	88.1	66.4	66.0	72.5	73.5
Set D & E	100.0	100.0	88.5	97.0	93.5	98.5
Set A + D & E	95.6	96.0	66.0	68.4	81.6	83.0

Table 4 Performance comparison between GD learning and ABC learning with Inverse-multi-quadric RBFNN with 10-fold cross validation.

Experiments	Average specificity		Average sensitivity		Average accuracy	
	RBFNN with GD	RBFNN with ABC	RBFNN with GD	RBFNN with ABC	RBFNN with GD	RBFNN with ABC
Set A & E	84.0	84.2	65.2	65.2	71.0	71.5
Set D & E	100.0	100.0	79.5	87.5	86.0	92.5
Set A + D & E	75.2	93.4	62.0	65.5	71.4	80.5

Table 5 Performance comparison between ABC learning and modified ABC learning with Inverse-multi-quadric RBFNN with 10-fold cross validation.

Experiments	Average specificity		Average sensitivity		Average accuracy	
	RBFNN with ABC	RBFNN with MABC	RBFNN with ABC	RBFNN with MABC	RBFNN with ABC	RBFNN with MABC
Set A & E	84.2	87.5	65.2	65.6	71.5	72.6
Set D & E	100.0	100.0	87.5	96.4	92.5	98.0
Set A + D & E	93.4	95.6	65.5	68.0	80.5	82.3

For RBFNN, there are mainly three types of basis functions that are used like, Gaussian, Multi-quadric and Inverse multi-quadric. But due to high performance of Gaussian and Inverse multi-quadric, the Multi-quadric function has been ignored.

Classification results of the classifiers were collected by a confusion matrix. In a confusion matrix, each cell contains the number of exemplars classified for the corresponding combination of desired and actual network outputs. The test performance of the methods was determined by the computation of the following statistical parameters for different experiments.

Experiment 1 (set A and E):

$$\text{Specificity} = \frac{EE}{EE + AE}, \tag{14}$$

$$\text{Sensitivity} = \frac{AA}{AA + EA}, \quad \text{and} \tag{15}$$

The accuracy of the model is defined as

$$\text{Accuracy} = \frac{AA + EE}{EE + AA + EA + AE}. \tag{16}$$

where AA: the count of cases that belong to the A class and are predicted as A (true positives); AE: the count of cases that belong to the E class and are predicted as A (false positives); EE: the count of cases that belong to the E class and are predicted as E (true negatives); and EA: the count of cases that belong to the A class and are predicted as E (false negatives). Similarly, the meaning of A and E is defined as follows: A: EEG signals recorded from healthy volunteers with eyes open, E: EEG signals recorded from epilepsy patients during epileptic seizures.

Similarly, the performance metrics of other experiments have been defined like Eqs. (14)–(16). However, the notations are different. The results have been validated using k-fold cross validation. Here, k value is chosen as 10. So, the whole dataset is divided into 10 unique subsets i.e. in each cycle of classification process, one set is taken for testing purpose and rest of the sets are taken for training purpose. As a result, total 10 cycles of classification task have been performed and the performance metrics were computed. Thus, the average of these metrics was taken as the final performance results. It was observed that there was a very minute difference between the best performance results and average performance results obtained through cross validation.

4.2. Result and analysis

Figs. 8–10 show the MSE graph for RBFNNN with (a) Gaussian RBF and (b) Inverse multi-quadric RBF with varying learning parameter for experiment numbers 1, 2 and 3 respectively. From these experiments it has been concluded that for Inverse multi-quadric RBF there is no effect of the learning parameter. For Gaussian RBF as the value of learning parameter increases, the mean square error quickly tends to its minima and for certain value of learning parameter it gives minimum MSE.

Now, RBFNN has been trained using ABC algorithm and for performance evaluation the Mean Square Error graph has been plotted for different runs of ABC. Figs. 11 and 12 show the variation in MSE for 50 runs of ABC algorithm with (a) Gaussian RBF and (b) Inverse multi-quadric RBF. It is being clearly observed that using ABC training algorithm the performance of RBFNNN with Inverse multi-quadric function has been enhanced. The MSE has been successfully reduced to 0.07 (approximately), after training the network with ABC optimization algorithm.

Table 2 shows the comparison of *sensitivity*, *specificity*, and *accuracy* between two training approaches, Gradient descent and Artificial Bee Colony of RBFNN. Clearly, it shows that the performance of RBFNN trained with ABC algorithm is better than the traditional Gradient Descent approach in all three experiments. Table 3 shows the comparison of performance of general ABC and our modified ABC. These results were taken from the best performances of the classifiers. Tables 4 and 5 show the results taken from the 10-fold cross validated classifier. Evidently, there is not much of difference in these results. Though, there is some improvement in performance for experiments 1 and 3, we can find a huge improvement in experiment 3. From this experimental evaluation, it is clearly proved that modified ABC algorithm can classify EEG data for epilepsy identification with highest accuracy.

5. Conclusion

Our approach is primarily based on Artificial Bee Colony algorithm, which is a new and robust algorithm used for training the RBFNN. However, we noticed that after adopting the binary tournament selection in the onlooker bees phase our novel approach for classifying epileptic seizure vs. non-epileptic seizure in three distinct experiments was performing significantly better than GD and ABC trained RBFNN. The performance of ABC trained RBFNN algorithm is compared with Gradient Descent approach trained RBFNN which is mostly used by the researchers. Finally, our concluding remark says ABC can be applied successfully to enhance the performance of RBF network for classification of EEG signal for detecting epileptic seizures. Moreover, the pre-processing of EEG signal is done by using DWT which is also an important requirement before going for the classification task. Our future research will be on this subject to study the class imbalancement problem in the analysis of EGG signal.

References

[1] Niedermeyer E, da Silva F Lopes. Electroencephalography: basic principles, clinical applications, and related fields. 5th ed. London: Lippincott Williams and Wilkins; 2005.

[2] Majumdar K. Human scalp EEG processing: various soft computing approaches. Appl Soft Comput 2011;11(8):4433–47.

[3] Sanei S, Chambers JA. EEG signal processing. New York: Wiley; 2007.

[4] Lehnertz K. Non-linear time series analysis of intracranial EEG recordings in patients with epilepsy – an overview. Int J Psychophysiol 1999;34(1):45–52.

[5] Alicata FM, Stefanini C, Elia M, Ferri R, Del Gracco S, Musumeci SA. Chaotic behavior of EEG slow-wave activity during sleep. Electroencephalogr Clin Neurophysiol 1996;99(6):539–43.

[6] Ocak H. Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. Expert Syst Appl 2009;36(2):2027–36.

[7] Witte H, Iasemidis LD, Litt B. Special issue on epileptic seizure prediction. IEEE Trans Biomed Eng 2003;50(5):537–8.

[8] Acharya UR, Sree SV, Swapna G, Martis R Joy, Suri JS. Automated EEG analysis of epilepsy: a review. Knowl Based Syst 2013;45:147–65.

[9] Finley KH, Dynes JB. Electroencephalographic studies in epilepsy: a critical analysis. Brain 1942;65:256–65.

[10] Gotman J. Automatic recognition of epileptic seizure in the EEG. Electroencephalogr Clin Neurophys 1982;54(5):530–40.

[11] Elo P, Saarinen J, Varri A, Nieminen H, Kaski K. Classification of epileptic EEG by using self-organizing maps. In: Aleksander I, Taylor J, editors. Artificial neural networks, vol. 2. p. 1147–50.

[12] Gabor AJ, Seyal M. Automated interictal EEG spike detection using artificial neural networks. Electroencephalogr Clin Neurophysiol 1992;83(5):271–80.

[13] Jando G, Siegel RM, Horvath Z, Buzsaki G. Pattern recognition of the electroencephalogram by artificial neural networks. Electroencephalogr Clin Neurophysiol 1993;86(2):100–9.

[14] Webber WRS, Litt B, Wilson K, Lesser RP. Practical detection of epileptiform discharges (Eds) in the EEG using an artificial neural network: a comparison of raw and parameterized EEG data. Electroencephalogr Clin Neurophysiol 1994;91(3):194–204.

[15] Pradhan N, Sadasivan PK, Arunodaya GR. Detection of seizure activity in EEG by an artificial neural network: a preliminary study. Comput Biomed Res 1996;29(4):303–13.

[16] Varsta M, Heikkonen J, Del J, Millan R. Epileptic activity detection in EEG with neural networks. In: Proc of 3rd int conf eng applications of neural networks. p. 179–86.

[17] Gabor AJ. Seizure detection using a self-organizing neural network: validation and comparison with other detection strategies. Electroencephalogr Clin Neurophysiol 1998;107(1):27–32.

[18] Walczack S, Nowack WJ. An artificial neural network approach to diagnosing epilepsy using lateralized bursts of theta EEGs. J Med Syst 2001;25(1):9–20.
[19] Nigam VP, Graupe D. A neural-network-based detection of epilepsy. Neurol Res 2004;26(1):55–60.
[20] Mishra BSP, Dehuri S, Wang GN. A state-of-the-art review of artificial bee colony in the optimization of single and multiple criteria. Int J Appl Metaheurist Comput 2013;4(1):32–49.
[21] Satapathy SK, Jagadev AK, Dehuri S. An empirical analysis of training algorithms of neural networks: a case study of EEG signal classification using java framework. Adv Intell Syst Comput 2015;309:151–60.
[22] Robert C, Gaudy JF, Limoge A. Electroencephalogram processing using neural networks. Clin Neurophysiol 2000;113(5):694–701.
[23] Aslan K, Bozdemir H, Sahin C, Ogulata SN, Erol R. A radial basis function neural network model for classification of epilepsy using EEG signals. J Med Syst 2008;32(5):403–8.
[24] Dehuri S, Roy R, Cho SB, Ghosh A. An improved swarm optimized functional link artificial neural network (ISO-FLANN) for classification. J Syst Softw 2012;85(6):1333–45.
[25] Dash ChSK, Dash AP, Dehuri S, Cho SB, Wang GN. DE + RBFNNs based classification: a special attention to removal of inconsistency and irrelevant features. Eng Appl AI 2013;26(10):2315–26.
[26] Dehuri S, Cho SB. Evolutionary optimized features in functional link neural network for classification. Expert Syst Appl 2010;37(6):4379–91.
[27] Andrzejak RG, Lehnertz K, Rieke C, Mormann F, David P, Elger CE. Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: dependence on recording region and brain state. Phys Rev Lett 2001;64:061907-1–7-8.