Research Article

Schizophrenia EEG Signal Classification Based on Swarm Intelligence Computing

Sunil Kumar Prabhakar, Harikumar Rajaguru, and Sun-Hee Kim

1Department of Brain and Cognitive Engineering, Korea University, Anam-dong, Seongbuk-gu, Seoul 02841, Republic of Korea
2Department of Electronics and Communication Engineering, Bannari Amman Institute of Technology, Sathyamangalam 638402, India

Correspondence should be addressed to Sunil Kumar Prabhakar; sunilprabhakar22@gmail.com and Sun-Hee Kim; sunheekim@korea.ac.kr

Received 12 August 2020; Revised 23 October 2020; Accepted 26 October 2020; Published 30 November 2020

Copyright © 2020 Sunil Kumar Prabhakar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

One of the serious mental disorders where people interpret reality in an abnormal state is schizophrenia. A combination of extremely disordered thinking, delusion, and hallucination is caused due to schizophrenia, and the daily functions of a person are severely disturbed because of this disorder. A wide range of problems are caused due to schizophrenia such as disturbed thinking and behaviour. In the field of human neuroscience, the analysis of brain activity is quite an important research area. For general cognitive activity analysis, electroencephalography (EEG) signals are widely used as a low-resolution diagnosis tool. The EEG signals are a great boon to understand the abnormality of the brain disorders, especially schizophrenia. In this work, schizophrenia EEG signal classification is performed wherein, initially, features such as Detrend Fluctuation Analysis (DFA), Hurst Exponent, Recurrence Quantification Analysis (RQA), Sample Entropy, Fractal Dimension (FD), Kolmogorov Complexity, Hjorth exponent, Lempel Ziv Complexity (LZC), and Largest Lyapunov Exponent (LLE) are extracted initially. The extracted features are, then, optimized for selecting the best features through four types of optimization algorithms here such as Artificial Flora (AF) optimization, Glowworm Search (GS) optimization, Black Hole (BH) optimization, and Monkey Search (MS) optimization, and finally, it is classified through certain classifiers. The best results show that, for normal cases, a classification accuracy of 87.54% is obtained when BH optimization is utilized with Support Vector Machine-Radial Basis Function (SVM-RBF) kernel, and for schizophrenia cases, a classification accuracy of 92.17% is obtained when BH optimization is utilized with SVM-RBF kernel.

1. Introduction

The reason for schizophrenia occurrence is generally not known, but researchers believe that it is a combination of the environment and brain genetics which contributes a lot to the development of this disorder [1]. The signs and symptoms usually involve disorganized speech, delusions, impaired functions of organs, and hallucinations [2]. Symptoms generally vary with type and severity depending on time, sometimes remission of symptoms can occur and sometimes the existing symptom can worsen to a great extent. The symptom of schizophrenia in teenagers is more or less the same as those in adults as they experience a drop in performance at school, lack of motivation, irritability, depressed mindset, withdrawal from friends and family, and also suffer from trouble sleeping [3]. People affected with schizophrenia generally lack awareness that their difficulty originates from a mental disorder which requires careful medical screening [4]. Suicidal thoughts and behaviour too are very common symptoms of schizophrenia. Certain naturally occurring brain channels such as neurotransmitters when altered or disturbed may contribute to schizophrenia [5]. Though the precise cause of schizophrenia is not known, the common risk factors are having a family history of schizophrenia, birth complications, exposure to toxic elements, malnutrition, and psychotropic drugs which alter the state of mind [6]. If schizophrenia is left untreated, then it can result in a plethora of problems such as anxiety disorder, depression, alcohol abuse, social isolation, aggressive behaviour leading to victimization, social isolation,
and financial issues followed due to various health and medical problems [7]. There is definitely no sure method to prevent schizophrenia, but considering and taking the treatment plan effectively can help in preventing relapses [8]. Characterized by relapsing episodes of psychosis, it is a serious disorder/mental illness. To determine the neural dynamics of human cognition, EEG recording acts as a sensitive tool as it can provide a millisecond-level resolution [9]. EEG data are complex, and at the same time, they are dimensional too as they are dependent on an event of time series [10]. As EEG signals provide electrical activity of the brain, it is easy to analyze the schizophrenia patient data.

Some of the most common works related to schizophrenia EEG signal analysis and classification reported in the literature are as follows. An EEG-dependent nonlinearity analysis technique for schizophrenia diagnosis was conducted by Zhao et al. [11]. The abnormal EEG complexity in patients with schizophrenia and depression was performed by Li et al. [12]. Fractal dimension was used by Raghavendra et al. [13] to analyze the complexity of EEG in schizophrenia patients. The complexity measures and entropy for EEG signal classification of both schizophrenia and control participants were performed by Sabeti et al. [14]. A preliminary report on the reduced nonlinear complexity and chaos during sleep in the first episode schizophrenia was given by Keshavan et al. [15]. A machine learning-based diagnosis of schizophrenia using combined sensor-level and source-level EEG features was proposed by Shim et al. [16]. A preliminary data analysis of comparing the EEG nonlinearity in deficit and nondeficit schizophrenia patients was conducted by Cerquera et al. [17]. The nonlinear analysis of EEG in schizophrenia patients with persistent auditory hallucination was performed by Lee et al. [18]. For a schizophrenia patient, the estimate of the first positive Lyapunov exponent of the EEG signal was performed by Kim et al. [19]. A magnetoencephalography (MEG) study on the LZC in schizophrenia patients was conducted by Fernandez et al. [20]. A multiscale entropy analysis for the abnormal EEG complexity signals was performed by Takahashi et al. [21]. As a diagnostic test for schizophrenia, the spectral EEG abnormality was analyzed by Boutros et al. [22]. An in-depth analysis of the utility of quantitative EEG in unmedicated schizophrenia was conducted by Kim et al. [23]. For schizophrenic and healthy adults, the machine learning identification of EEG features which helps in predicting working memory performance was done by Johannesen et al. [24]. A data-driven methodology for resting-state EEG signal classification of schizophrenia with control participants using random matrix theory was developed by Liu et al. [25]. Deep-learning methods along with random forest and voting classifiers were performed by Chu et al. for the individual recognition in schizophrenia using resting-state EEG streams [26]. Convolutional Neural Networks (CNNs) along with the Pearson Correlation Coefficient (PCC) to represent the EEG channel relationships were used to classify the schizophrenic and healthy patients using EEG signals by Naira and Alamo [27]. Random Forest Machine learning algorithm was used to identify and diagnose the schizophrenia EEG signals by Zhang [28].

A higher order pattern discovery was used to classify the schizophrenia EEG by Sheng et al. [29]. The complexity of the EEG signals in schizophrenia syndromes was analyzed by Kutepov et al. [30]. A fractal-based classification of EEG signals for both schizophrenic and healthy patients was performed by Namazi et al. [31]. A new approach for EEG signal classification using Linear Discriminant Analysis (LDA) and Adaboost was found for schizophrenic and control participants by Sabeti et al. [32]. In this work, with the advent of some features, optimization techniques and classifiers, the schizophrenia EEG signal classification is performed, and this attempt is the first of its kind in schizophrenia EEG signal classification. The organization of the work is as follows. Section 2 explains the materials and methods, Section 3 explains about the feature extraction techniques, Section 4 explains about the swarm intelligence computing techniques, Section 5 explains about the classification techniques, Section 6 explains the results and discussion followed by the conclusion in Section 7.

2. Materials and Methods

The EEG dataset for 14 healthy subjects and 14 Schizophrenic subjects was collected from the Institute of Psychiatry and Neurology, Warsaw, Poland, and the details are explained clearly in [33]. For the subjects, 7 males with an average age of 27.9+3.3 years and 7 females with an average age of 28.3+4.1 years were selected. The standard 10–20 International system was used for the acquisition of the EEG data. The patients’ data were obtained in a relaxed state and with their eyes closed. The segmentation of the acquired EEG signals was performed where it is considered to be stationary. A very simplified pictorial representation of the work is given in Figure 1.

Over the duration of 15 minutes, 19 channel EEG signals are obtained. Every channel of EEG signals comprises of 2,25,000 samples which are, then, divided into groups of 5000 sample segments. Therefore, the data matrix of [5000 x 45] is framed per channel. For the preprocessing of EEG signals, Independent Component Analysis (ICA) was utilized in this work.

3. Feature Extraction Techniques

To describe a very large amount of data, feature extraction is necessary as it involves in mitigating the number of resources to a certain limit. Once the preprocessing of the EEG signals is conducted, the following features are extracted from the EEG signals as follows:

(i) DFA: to trace the self-similarity characteristics of the EEG signal, DFA [34] is used.

(ii) Hurst exponent: the self-similarity and the predictability estimation of the EEG signal are expressed by the Hurst exponent [35]. If the magnitude value of the Hurst exponent is greater, then it denotes that the EEG signal is pretty smooth and less complicated.
(iii) RQA: to measure the complexity of the EEG signals, the total number of times of recurrences is evaluated by RQA [36].

(iv) Entropy: to evaluate the irregularity and uncertainty present in the EEG signal, entropy features are used. When the complexity and the variability of the EEG signal increases, then the entropy of the EEG signals is higher. Sample entropy [37], Shannon entropy [38], approximate entropy [39], Renyi entropy [40], permutation entropy [41], Kolmogorov–Sinai entropy [42], and fuzzy entropy [43] are the types of entropy usually used for the analysis, and in this work, only sample entropy has been extracted.

(v) FD: to compare and analyze the complexity of details in the EEG, FD is used, thereby enabling the detection of EEG signal patterns [44]. The Fractal Dimension of a signal is expressed through the Katz method as follows:

\[ D = \log_{10}(H) \log_{10}(d) \]

where between the successive points, \( H \) represents the sum of distances and \( d \) represents the diameter estimated.

(vi) Kolmogorov complexity: the characteristics of the EEG signal are easily explained by this parameter [45]. If the signals are more random, then the description length is also longer.

(vii) Hjorth: to quantify the morphology of the EEG signal, the important Hjorth parameters utilized here are complexity, mobility, and activity [46].

(viii) LZC: to assess the repetitiveness of the EEG signals, LZC is used [47]. If the LZC values are higher, then it shows that the signal is more repetitive.

(ix) LLE: to assess the degree of chaos present in the EEG signals, an estimate of it is located by the LLE [48]. If the complexity of the signals is high, then the value of LLE is also high.

The feature extraction is initiated using DFA at first among the nine features. The attained feature matrix per method per channel is in the form of \([5000 \times 10]\). Then, four types of optimization producer such as AF optimization, GS Optimization, BH optimization, and MS optimization are utilized to further extract a better represented feature column matrix as \([5000 \times 1]\). This procedure is repeated for all the channels among the subjects. This feature extraction method is repeated for all the other eight features such as the Hurst exponent, RQA, entropy, fractal dimension, Kolmogorov complexity, Hjorth, and LLE, and thus, the feature extraction is performed for each data segment as such.

4. Swarm Intelligence Computing Techniques

To determine and understand a certain subset of initial features, feature selection is required. The features which are selected will usually have the useful and most relevant information from the original data so that, using the reduced form of representation, the desired task can be performed easily. The following four optimization techniques are utilized in this work.

4.1. Artificial Flora Algorithm. Developed by Cheng et al. [49], four basic elements are comprised in this algorithm such as original plant, offspring plant, location of the plant, and the distance of propagation initially. The plants that are used to spread seeds are called original plants. The seeds of original plants are called offspring plants, and at that moment, they cannot spread seeds. Plant location is the specific location of the plant, and the distance of propagation refers to how long a seed can spread. The three major patterns are present here such as evolution behaviour, spread behaviour, and select behaviour. The probability that the evolution of the plant adapts to the environment behaviour is called evolution behaviour. The movement of seeds is referred by the spreading behaviour, and select behaviour refers to the survival or death of the flora due to the environment. The main rules here are as follows:

**Rule 1.** Due to the environmental or any external factors, a random distribution of a species in a region is performed; in that region, no such species were found earlier, and so, it becomes the primitive original plant.

**Rule 2.** As the environment changes, the plants will adopt to live in the main environment. Therefore, a complete inheritance to the parent plant does not depend on the proper distance of offspring plants.

**Rule 3.** When the seeds are spread around the original plant autonomously, the range is a circle where radius is the maximum propagation distance. Anywhere in the circle, the offspring plants can be distributed.
Rule 4. Plants have varied survival probability because of environmental factors such as climate and temperature. The survival probability is related to the fitness of the plants. Here, the adaptability of the plants to the environment can be referred by fitness. Or in other words, the survival probability of a plant in a specific position is termed as fitness. If the fitness is higher, then the probability of survival is greater.

Rule 5. The survival probability will be lower if the distance is further from the original plants because the basic difference between the current environment and the previous environment will be much greater.

Rule 6. If the seeds are spread in an external margin, then the spread distance cannot cross the maximum area of limit because of other constraints.

4.1.1. Evolution Behaviour. The seeds are spread by the original plant around in a circle which is nothing but the propagation distance and is evolved for the respective propagation distances of the parent plant and the grandparent plant and is represented as

\[ d_h = d_{1h} \times \text{rand}(0,1) \times c_1 + d_{2h} \times \text{rand}(0,1) \times c_2, \]  

(2)

where \( d_{1h} \) is the propagation distance of the grandparent plant, \( d_{2h} \) is the propagation distance of the parent plant, \( c_1 \) and \( c_2 \) are the learning coefficients, and \( \text{rand}(0,1) \) denotes the independent uniformly distributed number in \((0,1)\).

The normal grandparent distance of propagation is expressed as

\[ d'_{1h} = d_{2h}. \]

(3)

The standard deviation between the respective position of the original and offspring plant is the new parent propagation distance and is given as

\[ d_{2h} = \sqrt{\frac{\sum_{i=1}^{N} (L_{i,h} - L'_{i,h})^2}{N}}. \]

(4)

4.1.2. Spreading Behaviour. The original flora with \( N \) solutions is randomly generated by the AF algorithm if there are \( N \) plants in the flora. The matrix \( L_{i,h} \) is used to express the original plant position where \( i \) is the dimension and \( h \) is the total number of plants in the flora.

\[ L_{i,h} = \text{rand}(0,1) \times d \times 2 - d, \]

(5)

where the maximum limit area is represented by \( d \), and \( \text{rand}(0,1) \) is an array of random numbers that have a uniform distribution between \((0,1)\). With the help of the following propagation function, the position of the offspring plant is generated as follows:

\[ L'_{i,h,om} = D_{i,h,om} + L_{i,h}. \]

(6)

where the number of seeds that one plant can propagate is expressed as \( m \), \( L'_{i,h,om} \) denotes the offspring position, \( L_{i,h} \) is the original plant position, and \( D_{i,h,om} \) is random number with the Gaussian distribution with zero mean and variance \( d_h \). If the survival of the offspring plant is not guaranteed, then a new plant is generated as shown in the above-mentioned equation.

4.1.3. Select Behaviour. The survival probability is used to assess whether the offsprings are alive or not and is represented as

\[ l = \sqrt{\frac{F(L'_{i,h,om})}{F_{\text{max}}}} \times P_y^{(h_{\text{om}}-1)}, \]

(7)

where \( P_y^{(h_{\text{om}}-1)} \) is \( P_y \) to the power of \((h \times m - 1)\) and \( P_y \) is the selective problem. This value has to be between 0 and 1, and in our experiment, it is 0.5. When the offspring plant is farther from the original plant, then fitness is lower. \( P_y \) can assess the exploration ability of this algorithm. To get a local optima solution, \( P_y \) should be larger. The maximum fitness in the flora is determined by the \( F_{\text{max}} \), and the fitness of the \( h_{\text{th}} \) solution is determined by \( F(L'_{i,h,om}) \). To decide whether the offspring plant is alive or not, the roulette wheel selection method is used in our work. The procedure is explained in Pseudocode 1.

**Pseudocode 1**

Input: times: maximum run time

\( B \): maximum branching number

\( N \): number of original plants

\( l \): survival probability of offspring plants

\( t = 0 \): population Initialization.

Define related parameters

Evaluate the fitness value of \( N \) individuals and compute the best solution

While \((t < \text{times})\)

For \( i = 1: N \times B \)

Propagation distance is evolved by new plants

Original plants spread their offspring

If \( \text{rand}(0,1) > l \)

Offspring is alive

Else

Offspring is dead

End if

End for

Evaluate novel solutions and randomly select \( N \) plants as new original plants

If new solution is better, then old plant is replaced by new plant

Find the current best solution

\( t = t + 1 \)

End while

Output: optimized solution
4.2. Glowworm Swarm Optimization Algorithm. In this algorithm, initially, the position location and information on data exchange can be carried out by most glowworms by means of sending out rhythmed short beams [50]. Around a particular searching scope, the main intention of GS optimization is to find out the flaring neighbor. The glowworms move always from the first position to a best position and finally into a more extreme value point. The attraction of the glowworm individuals is highly related to its brightness. The attractiveness of a particular individual glow worm is inversely proportional to the distance between the two individuals, so it implies that it has a direct proportion to brightness. Each position of the individual glow worm accounts for the objective function value. The individual search scope is defined by the dynamic decision domain. The individual movement is later updated step by step in the procedure given below.

4.2.1. Procedure

(i) Parameter initialization: 'n' individuals are initially placed in a random fashion around a feasible region. \( f_0 \) denotes the fluorescein value, \( q_0 \) indicates the dynamic decision domain, \( st \) indicates step, \( i_n \) expresses domain threshold, the update coefficient of domain is expressed as \( \beta, q_{dn} \) denotes the maximum search radius, and the iteration number is expressed as \( n \).

(ii) The objective function value \( H(y_j(n)) \) is transformed to \( f_j(n) \) as

\[
  f_j(n) = (1 - \rho)f_j(n - 1) + \gamma H(y_j(n)),
\]

where \( y_j(n) \) expresses for the individual position \( j \) at 'n' instant of time.

(iii) In each \( q_d(n) \), the higher fluorescein value is selected, thereby forming a set of neighborhood \( I_j(n) \). Therefore,

\[
  I_j(n) = \left\{ h : \|y_h(n) - y_j(n)\| < q_d(n); f_j(n) \leq f_h(n) \right\}.
\]

(iv) The probability of a particular individual \( j \) may progress forward as \( h \) and is expressed as

\[
  p_{jh}(n) = \frac{f_j(n)}{\sum_{m \in I_j(n)} f_m(n) - f_j(n)},
\]

where \( h \) is chosen by \( p_{jh}(n) \).

(v) The updation of the position of individual \( j \) is expressed as

\[
  y_j(n + 1) = y_j(n) + \text{st} \left( \frac{y_h(n) - y_j(n)}{\|y_h(n) - y_j(n)\|} \right).
\]

(vi) The updation of the dynamic decision domain is expressed as

\[
  q_d(n + 1) = \min\{q_d, \max\{0, q_d(n), \beta \left( i_n - |f_j(n)| \right) \}\}.
\]

4.3. Black Hole Algorithm. The algorithm is inspired from the black hole phenomenon [51]. In this method, at each iteration, the best candidate is chosen as the black hole and the other normal candidates are chosen as normal stars. The black hole creation is one of the original candidates of the entire population, and it is not random. Depending on one’s current location and a randomly generated number, the movement of the candidate towards the black hole is ascertained. The step-by-step details of the algorithm is as follows:

(i) It is a very famous population-based algorithm. Therefore, in the search space of some function, some randomly generated population of candidate solutions (stars) is placed.

(ii) The fitness value evaluation of the population is conducted after the initialization process.

(iii) The best candidate in the population is the one which has the highest fitness value, and it is assumed to be the initial black hole.

(iv) While this process is going on, the rest of the candidates do the normal stars.

(v) The specialty of the black hole is that it has the capacity to absorb the surrounding stars around it.

(vi) Once the stars and the black hole are initialized, the absorption of the stars by the black hole takes place, and therefore, the movement of all other stars is now towards or around the black hole.

(vii) The formulation of the capability of the absorption of stars by the black hole is ascertained as follows:

\[
  Y_j(t + 1) = Y_j(t) + \text{rand}(Y_d - Y_j(t)).
\]

Here, \( Y_j(t) \) and \( Y_j(t + 1) \) are the locations of the star \( j \) in iteration \( t \) and \( t + 1 \). \( Y_d \) is the location of the black hole in the entire search space. Random number in the interval of \([0, 1]\) is determined by rand.

(viii) The black hole absorbs the candidate solution (each star representing it) that crosses the horizon of the black hole. Using the following equation, the radius of the horizon in the BH algorithm is computed as follows:
\[ R = \left[ \frac{G_d}{\sum_{j=1}^{S} G_j} \right], \]  

(14)

where \( G_d \) represents the black hole fitness value, \( G_j \) is the fitness value of the star \( j \), \( S \) denotes the number of stars which indicates the size of population/candidate solutions, and \( R \) denotes the black hole radius.

The candidate is usually collapsed when the distance between the best candidate (black hole) and the candidate solution is less than \( R \). In such a case, the value of a new candidate would be carried out, and it is randomly distributed in the search space giving the optimized and the best values.

4.4. Monkey Search Algorithm. Based on the mountain climbing procedure of monkeys, one of the recently developed metaheuristic algorithms is MS algorithm [52]. Initially, for the monkeys, the population size is defined. Each monkey’s specific position is denoted by a vector \( a_j \), and it is performed randomly. With the help of a step-by-step climbing process, the monkey positions are changed. One of the famous recursive optimization algorithms is MS algorithm [51]. The objective function value is improved by the climb process in MS algorithm. The objective function value is improved by the climbing process. Once a monkey arrives on top of a mountain after the climbing process, then it will search for higher points than its current position. A monkey with its sharp and keen eyesight would easily shift its base to a higher point by jumping. The maximum distance watched by a monkey is determined by the eyesight of a monkey. The updation of the position is carried out. Then, the monkeys start to search the novel search domains by utilizing the current position as a pivot. This process is termed as a somersault procedure, and it helps the monkeys to get a new position. The evaluation of the objective function values is conducted, and if the number of iterations is satisfactory, then the process will be stopped. The main process of this algorithm is as follows:

(1) Representation of a solution: the population size of monkey \( M \) is defined initially. For any monkey, \( j \in \{1, 2, \ldots, M\} \) defines its position \( a_j = (a_{j1}, a_{j2}, \ldots, a_{jn}) \) which gives the optimization problem solutions with the \( n \) dimensions.

(2) Initialization: in a random manner, the initial solution of the monkeys is generated.

(3) Climbing process:

(i) A vector is generated randomly \( \Delta a_j = (\Delta a_{j1}, \Delta a_{j2}, \ldots, \Delta a_{jn}) \), where \( \Delta a_{jk} \) is set as \( s \), where \( s \) denotes the step length of climb process.

\[ g'_{jk}(a_j) = \frac{g(a_{j} + \Delta a_j) - g(a_{j} - \Delta a_j)}{2\Delta a_{jk}}, \]  

(15)

(\( k = 1, 2, \ldots, n \), and at point \( a_j \), the vector is termed as the pseudogradient of the objective function)

(ii) Assign \( z_k = a_{jk} + \alpha \text{sign}(g_{jk}(a_j)), k = 1, 2, \ldots, n \) and \( z = (z_1, z_2, \ldots, z_n) \)

(iii) Assume \( a_j - z \) has given \( z \) is feasible or else keep as such \( a_j \)

(iv) Unless the maximum number of allowed iterations \( N_q \) has reached or there is a little change in the objective function of the iteration, the steps (i) to (iii) of the climbing process are repeated

(4) Watch jump process:

(i) The real number \( z_j \) is generated randomly from \( (a_{jk} - e, a_{jk} + e), k = 1, 2, \ldots, n \), where \( e \) represents the eyesight of the monkey which denotes the maximum distance that can be witnessed by a monkey.

(ii) Assume \( a_j \leftarrow z \) given that \( g(z) > g(a_{j}) \) and \( z \) is feasible. Unless a certain number of watch times has been obtained or until an appropriate point \( z \) is found out, the steps are repeated.

(iii) The climb process is repeated by employing \( z \) as an initial position.

(5) Somersault process:

(i) From the somersault interval \([q, d]\), a real number \( \theta \) is generated randomly.

(ii) Set \( z_k = a_{jk} + \theta(v_k - a_{jk}), v_k = (1/M) \sum_{j=1}^{M} a_{jk}, k = 1, 2, \ldots, n \).

(iii) Set \( a_{j} - z \) if \( z = (z_1, z_2, \ldots, z_n) \) is quite feasible. Unless a feasible solution \( z \) is found out, steps (i) and (iii) of this somersault process are repeated.

(6) Termination.

Unless the stopping criterion is met, the above-mentioned steps are repeated. As the stopping criteria, the number of iterations are used.

Thus, using these metaheuristic algorithms, the optimal and best solutions are explored and exploited from the initial random solutions.

5. Classification Techniques

The optimized values or the best selected feature values through swarm intelligence computing techniques are finally fed to the classifiers for classification.

5.1. ANN. Here, a three-layer perceptron comprising of an input layer, hidden layer, and an output layer was utilized [53]. The total number of hidden neurons is expressed by the following equation:

\[ H = [(A + B)d], \]  

(16)

where the number of input neurons is expressed as \( A \), the number of output neurons is expressed as \( B \), and \( d \) is a
constant and its range is from \( d \in [0, 1] \). For the ANN classifier, the total number of hidden neurons used here is 50.

5.2. QDA. The ratio of the between class variance is maximized, and the ratio of the within class variance is minimized by QDA [54]. Between classes, it also allows quadratic decision boundaries so that the classifier can perform the classification more accurately providing good classification accuracy. In this QDA, no shrinkage or regularization is used.

5.3. SVM. Due to its good scalability and its high classification performance, SVM is used [55]. Creating a hyperplane to maximize the margins between the classes that can be obtained by mitigating the cost function, so that the maximum classification accuracy is attained is the main idea of SVM. The hyperplanes which are represented by the vectors are known as support vectors. By minimizing the cost function, the optimal solution that maximizes the distance between the hyperplane and the nearest training point is obtained by the SVM as follows:

\[
\text{Minimize}, \quad \frac{1}{2}||w||^2 + C \sum_{j=1}^{n} \xi_j,
\]

subject to, \( z_j(w^T x_j + f) \geq 1 - \xi_j, \quad \xi_j \geq 0, \)

where \( w^T, x_j \in R^2 \) and \( f \in R_t, ||w||^2 = w^T w, \)

The tradeoff between the margin and the error is denoted as \( C \). The measure of the training data is expressed as \( \xi \). The class label for the \( j^{th} \) sample is denoted as \( z_j \). SVM can be utilized as a both linear and nonlinear classifier. Various types of kernel functions are utilized to make SVM as a nonlinear classifier. The types of kernels generally used are Polynomial, Radial Basis Function (RBF), and Sigmoid kernels. Here, in our work, only SVM-RBF kernel is used, and this nonlinear SVM is used to get higher classification accuracy.

5.4. Logistic Regression. Between an independent variable and a response variable, the relationship is assessed by Logistic Regression (LR) [56]. An observation is classified into one of the two classes using Logistic Regression in a very simple manner. When the variables are nominal or binary, it can be used. Similar to the Bayesian group, the data are comprehensively analyzed after the discretization process for the continuous variables is performed.

5.5. FLDA. The main intention of the Fischers Linear Discriminant Analysis (FLDA) is to trace a direction in the feature space along which the specific distance of the means relative to the within-class scatter explained by the within-class scatter matrix \( S_W \) reaches a maximum [57]. When the within-class scatter matrix reaches a maximum, the class separability is maximized. By maximizing the following criterion with the between-class scatter matrix, this goal can be achieved.

\[
J(W) = \frac{W^T S_W W}{W^T S_W W}
\]

To maximize the criterion, the direction \( w \) is expressed as \( W = S_W^{-1} (m_1 - m_2) \), where \( m_1 \) and \( m_2 \) are the means for the two classes. For the two classes, the FLDA acts as a sub-optimal classifier when their respective distributions are Gaussian in nature.

5.6. KNN. One of the famous nonparametric algorithms utilized for both classification and regression is KNN [58]. On the underlying data distribution, KNN does not make any assumption. There is also no explicit training phase available here. During the testing phase, the utilization of the training data is carried out where the measurement between the training instance and test instance is performed. The prediction of the class of the test instance is performed by utilizing the majority voting of the any of the K-nearest training instances. The value of K is assumed to be 4 in our work.

6. Results and Discussion

It is classified with 70% of training and 30% of testing methodology and the performance of it is computed. The experiment was repeated five times to check whether we get the similar results every time when the analysis is done. The mathematical formulae for computing the Performance Index (PI), Sensitivity, Specificity, Accuracy, Good Detection Rate (GDR), and Mean Square Error (MSE) rate is mentioned in literature, and using the same, the values are computed and exhibited. PC indicates Perfect Classification, FA indicates False Alarm, and MC indicates Missed Classification, respectively.

The sensitivity is expressed as

\[
\text{Sensitivity} = \frac{\text{PC}}{\text{PC} + \text{FA}} \times 100.
\]

Specificity is expressed as

\[
\text{Specificity} = \frac{\text{PC}}{\text{PC} + \text{MC}} \times 100.
\]

Accuracy is expressed as

\[
\text{Accuracy} = \frac{\text{Sensitivity} + \text{Specificity}}{2}.
\]

Performance Index is expressed as

\[
\text{PI} = \left( \frac{\text{PC} - \text{MC} - \text{FA}}{\text{PC}} \right) \times 100.
\]

Good Detection Rate (GDR) is expressed as

\[
\text{GDR} = \left( \frac{\text{PC} - \text{MC}}{\text{PC} + \text{FA}} \right) \times 100.
\]

The Mean Square Error (MSE) is expressed as follows:
where $O_j$ indicates the observed value at a specific time, $T_j$ denotes the target value at model $j$; $j = 1$ to 19, and $N$ is the total number of observations per channel for a single optimization method, and in our case, it is 45000. The training of the classifiers was implemented with a zero-training error of MSE.

Table 1 shows the average statistical parameters such as mean, variance, skewness, and kurtosis of the nine extracted features through four-optimization process for the normal cases. The higher value of mean indicates the peaked position of the feature selected. Lower value of mean indicates that there exists the peak and valley points in the features. As in the case of AF optimization in Kolmogorov complexity features, peaked value of mean is attained. The variance parameter shows the energy component of the feature. Here also in Kolmogorov complexity in AF optimization method arrived higher value. Skewness depicts the skewed features of the data points, and all the features in Table 1 indicates the same. The flatness is indicated by the higher kurtosis values. In the case of Hurst exponent, Fractal dimension and Hjorth features at AF optimization show higher value of kurtosis.

Table 2 demonstrates the average statistical parameters such as mean, variance, skewness, and kurtosis of the nine extracted features through four-optimization process for the schizophrenia cases. The higher value of mean indicates the peaked position of the feature selected. Lower value of mean indicates that there exist the peak and valley points in the features, as all the optimization methods made the mean parameter as a smaller one among the nine features. Skewness depicts the skewed features of the data points and all the features in Table 2 indicate the same. The flatness is indicated by the higher kurtosis values. In the case of Kolmogorov complexity at AF optimization and BH optimization, it shows higher value of Kurtosis.

The correlation among the normal and schizophrenia cases can be established by calculating the Canonical Correlation Analysis (CCA) as shown in Table 3. If the CCA value is greater than 0.5 that indicates the two classes are highly correlated and for lower value it is vice versa. As shown in Table 3 the CCA is calculated for the nine features among the normal schizophrenia cases. As observed from Table 3, the lower value of CCA indicates the Nil correlation among the features across the classes.

Table 4 exhibits the average PCC with different features for normal and schizophrenia cases. The values in Table 2 indicates the nonlinear relation among the features in the same class of the data. Therefore, the uncorrelated and nonlinear features have to be optimized by the optimization process.

Table 5 shows the CCA of various optimization techniques with different features for normal and schizophrenia cases. It is observed from Table 5 that the low value of CCA is definitely indicating the Nil correlation among the features of the two classes.

Table 6 depicts the Average PCC at various optimization techniques with nine different features for normal and schizophrenia cases. As shown in Table 6, low value of CCA in the normal cases indicates the presence of nonlinearity in the features. The negative value of PCC in the schizophrenia cases mention about the inverse relation among the features as well as the optimization methods.

Table 7 shows the consolidated results of accuracy among the classifiers at various optimization techniques with different features of normal cases. As indicated in Table 7, ANN classifier is ebbed at low value of accuracy in the three types of optimization methods such as AF, GS, and BH. The poor performance of ANN is due to over learning and the exhibition of false alarm in the classifier outputs. FLDA classifier arrived at low accuracy in the case of MS optimization method. SVM classifier is outperforming all the classifiers in terms of higher accuracy value of 87.54% in BH optimization method.

Table 8 denotes the consolidated results of accuracy among the classifiers at various optimization techniques with different features of schizophrenia cases. As observed in Table 8, LR classifier is ebbed at low value of accuracy of 78.64% in the black hole optimization method. The poor performance of LR is due to the rigid parametric values and the exhibition of missed classification in the classifier outputs. SVM classifier is outperforming all the classifiers in terms of higher accuracy value of 92.17% in BH optimization method. Evenly nature of the performance shows higher accuracy values among the classifiers for MS method.

Table 9 represents the average perfect classification among the classifiers at various optimization techniques with different features of normal cases. As observed in Table 9, LR classifier is ebbed at low value of perfect classification of 54.3% in the AF optimization method. The poor performance of LR is due to the nonadaptive parametric values and the exhibition of false alarm in the classifier outputs. SVM classifier is outperforming all the classifiers in terms of higher perfect classification value of 75.093% in BH optimization method. Evenly nature of the performance shows higher perfect classification values among the classifiers for MS method.

Table 10 denotes the average perfect classification among the classifiers at various optimization techniques with different features of schizophrenia cases. As shown in Table 10, LR classifier is reached at low value of perfect classification of 57.29% in the BH optimization method. The poor performance of LR is due to the nonadaptive parametric values and the exhibition of missed classification in the classifier outputs. SVM classifier is outperforming all the classifiers in terms of higher perfect classification value of 84.34% in BH optimization method. Evenly nature of the performance shows higher perfect classification values among the classifiers for MS method.
Artificial flora

- Mean: $2.320835$ ± $0.55226$ 06
- Variance: $0.090556$ ± $0.444369$ 00
- Skewness: $1.83474$ ± $0.014862$ 03
- Kurtosis: $7.92829$ ± $2.70178$ 35

Glowworm swarm

- Mean: $3.60521$ ± $1.49249$ 86
- Variance: $6.1955$ ± $0.304512$ 37
- Skewness: $2.7591$ ± $0.0952$ 92
- Kurtosis: $10.378$ ± $0.3245$ 78

Black hole

- Mean: $0.04486$ ± $1.64093$ 90
- Variance: $0.00005$ ± $0.000033$ 67
- Skewness: $1.83474$ ± $0.18256$ 99
- Kurtosis: $4.37384$ ± $0.12708$ 10

Monkey search

- Mean: $0.360178$ ± $0.64032$ 94
- Variance: $0.01044$ ± $0.449776$ 03
- Skewness: $0.232422$ ± $0.362725$ 00
- Kurtosis: $0.06264$ ± $13.54124$ 87

Table 2: Average parameters at various optimization techniques with different features for schizophrenia cases.

<table>
<thead>
<tr>
<th>Optimization methods</th>
<th>Parameters</th>
<th>DFA</th>
<th>Hurst</th>
<th>RQA</th>
<th>Sample entropy</th>
<th>Fractal dimension</th>
<th>Kolmogorov complexity</th>
<th>Hjorth</th>
<th>LZC</th>
<th>LLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial flora</td>
<td>Mean</td>
<td>$1.947$ ± $0.797$ 97</td>
<td>$0.895$ ± $1.0434$ 43</td>
<td>$0.9192$ ± $0.9182$</td>
<td>$0.9182$ ± $0.922$ 92</td>
<td>$0.922$ ± $0.7272$ 77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>$6.1955$ ± $0.00198$ 76</td>
<td>$8.213E$ ± $3.00456$ 28</td>
<td>$7.331E$ ± $0.00011$ 01</td>
<td>$4.82E$ ± $0.393E$ 06</td>
<td>$3.393E$ ± $0.0382$ 85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>$2.7591$ ± $0.0952$ 92</td>
<td>$0.1699$ ± $3.4882$ 22</td>
<td>$7.9156E$ ± $-31.6315$</td>
<td>$4.8058$ ± $1.162$ 62</td>
<td>$0.6692$ ± $0.4216$ 85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurtosis</td>
<td>$10.378$ ± $-0.3245$ 78</td>
<td>$-0.000067$ ± $16.5549$</td>
<td>$139.9124$ ± $998.7822$</td>
<td>$25.231$ ± $2.795$ 98</td>
<td>$0.3459$ ± $0.5087$ 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glowworm swarm</td>
<td>Mean</td>
<td>$3.60521$ ± $0.45140$ 57</td>
<td>$0.528535$ ± $2.05044$ 43</td>
<td>$1.109916$ ± $2.762614$</td>
<td>$0.775$ ± $0.874$ 36</td>
<td>$0.4336$ ± $0.4363$ 69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>$6.1955$ ± $0.00198$ 76</td>
<td>$8.213E$ ± $0.30456$ 28</td>
<td>$7.331E$ ± $0.00011$ 01</td>
<td>$4.82E$ ± $0.393E$ 06</td>
<td>$3.393E$ ± $0.0382$ 85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>$2.7591$ ± $0.0952$ 92</td>
<td>$0.1699$ ± $3.4882$ 22</td>
<td>$7.9156E$ ± $-31.6315$</td>
<td>$4.8058$ ± $1.162$ 62</td>
<td>$0.6692$ ± $0.4216$ 85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurtosis</td>
<td>$10.378$ ± $-0.3245$ 78</td>
<td>$-0.000067$ ± $16.5549$</td>
<td>$139.9124$ ± $998.7822$</td>
<td>$25.231$ ± $2.795$ 98</td>
<td>$0.3459$ ± $0.5087$ 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black hole</td>
<td>Mean</td>
<td>$0.04486$ ± $1.64093$ 90</td>
<td>$0.38062$ ± $0.08666$ 74</td>
<td>$0.471683$ ± $0.193926$</td>
<td>$0.246$ ± $0.209$ 09</td>
<td>$1.40907$ ± $0.140973$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>$0.01454$ ± $0.828733$ 87</td>
<td>$0.000476$ ± $0.020484$ 24</td>
<td>$0.054866$ ± $0.00054$ 23</td>
<td>$1.095E$ ± $6.7E$ 06</td>
<td>$0.102673$ ± $1.65326$ 73</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>$-0.911996$ ± $2.338573$ 86</td>
<td>$0.602759$ ± $-0.25059$ 28</td>
<td>$2.625459$ ± $31.6492$</td>
<td>$-2.561$ ± $9.701$ 21</td>
<td>$1.24327$ ± $15.6235$ 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurtosis</td>
<td>$1.689095$ ± $8.526093$ 75</td>
<td>$8.936210$ ± $1.381618$ 63</td>
<td>$8.043591$ ± $103.234$</td>
<td>$5.236$ ± $15.6235$ 26</td>
<td>$1.35525$ ± $15.6235$ 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monkey search</td>
<td>Mean</td>
<td>$0.360178$ ± $0.64032$ 94</td>
<td>$0.142771$ ± $0.139355$ 45</td>
<td>$0.000866$ ± $0.540895$</td>
<td>$0.933$ ± $0.000151$ 47</td>
<td>$2.272099$ ± $14.98741$ 93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>$0.01044$ ± $0.449776$ 03</td>
<td>$0.000665$ ± $0.001762$ 24</td>
<td>$2.15E$ ± $0.000703$ 23</td>
<td>$1.9E$ ± $85.4282$ 84</td>
<td>$0.43814$ ± $11.6151$ 51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>$0.232422$ ± $2.863725$ 00</td>
<td>$0.441075$ ± $0.154336$ 35</td>
<td>$0.980327$ ± $-0.20562$ 23</td>
<td>$0.5348$ ± $0.799874$ 95</td>
<td>$11.6151$ ± $0.43814$ 51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Average CCA with different features for normal and schizophrenia cases.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>DFA</th>
<th>Hurst</th>
<th>RQA</th>
<th>Sample entropy</th>
<th>Fractal dimension</th>
<th>Kolmogorov complexity</th>
<th>Hjorth</th>
<th>LZC</th>
<th>LLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCA</td>
<td>$0.14231$ ± $0.05738$ 76</td>
<td>$0.0876$ ± $0.14356$ 54</td>
<td>$0.06889$ ± $0.15469$ 74</td>
<td>$0.06334$ ± $0.12817$ 17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Average PCC with different features for normal and schizophrenia cases.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>DFA</th>
<th>Hurst</th>
<th>RQA</th>
<th>Sample entropy</th>
<th>Fractal dimension</th>
<th>Kolmogorov complexity</th>
<th>Hjorth</th>
<th>LZC</th>
<th>LLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>$0.021077$ ± $0.014018$ 77</td>
<td>$0.006344$ ± $0.047908$ 43</td>
<td>$0.016744$ ± $0.001745$ 54</td>
<td>$0.023395$ ± $0.00045$ 26</td>
<td>$0.000145$ ± $0.049885$ 75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia cases</td>
<td>$0.053388$ ± $0.012278$ 78</td>
<td>$0.068795$ ± $0.030514$ 20</td>
<td>$0.032015$ ± $0.107226$ 43</td>
<td>$0.000210$ ± $0.062845$ 17</td>
<td>$0.000376$ ± $0.043985$ 85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5: CCA at various optimization techniques with different features for normal and schizophrenia cases.

<table>
<thead>
<tr>
<th>Optimization methods</th>
<th>CCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial flora</td>
<td>0.047944</td>
</tr>
<tr>
<td>Glowworm swarm</td>
<td>0.088456</td>
</tr>
<tr>
<td>Black hole</td>
<td>0.060256</td>
</tr>
<tr>
<td>Monkey search</td>
<td>0.089556</td>
</tr>
</tbody>
</table>

Table 6: Average PCC at various optimization techniques with different features for normal and schizophrenia cases.

<table>
<thead>
<tr>
<th>Optimization methods</th>
<th>Normal PCC</th>
<th>Schizophrenia PCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial flora</td>
<td>0.005745</td>
<td>−0.01573</td>
</tr>
<tr>
<td>Glowworm swarm</td>
<td>0.04604</td>
<td>−0.07745</td>
</tr>
<tr>
<td>Black hole</td>
<td>0.040539</td>
<td>−0.01178</td>
</tr>
<tr>
<td>Monkey search</td>
<td>0.08175</td>
<td>−0.1422</td>
</tr>
</tbody>
</table>

Table 7: Consolidated results of accuracy (%) among the classifiers at various optimization techniques with different features for normal cases.

<table>
<thead>
<tr>
<th>Optimization methods</th>
<th>ANN</th>
<th>QDA</th>
<th>SVM</th>
<th>LR</th>
<th>FLDA</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial flora</td>
<td>77.42581</td>
<td>79.08989</td>
<td>85.40703</td>
<td>77.15035</td>
<td>80.07678</td>
<td>82.61267</td>
</tr>
<tr>
<td>Glowworm swarm</td>
<td>77.35069</td>
<td>79.35989</td>
<td>87.14997</td>
<td>83.3216</td>
<td>80.18607</td>
<td>81.87646</td>
</tr>
<tr>
<td>Black hole</td>
<td>77.25071</td>
<td>78.82409</td>
<td>87.54716</td>
<td>77.3041</td>
<td>80.27763</td>
<td>79.3533</td>
</tr>
<tr>
<td>Monkey search</td>
<td>81.47733</td>
<td>79.93699</td>
<td>85.08729</td>
<td>84.48008</td>
<td>79.67025</td>
<td>81.07509</td>
</tr>
</tbody>
</table>

Table 8: Consolidated results of accuracy (%) among the classifiers at various optimization techniques with different features for schizophrenia cases.

<table>
<thead>
<tr>
<th>Optimization methods</th>
<th>ANN</th>
<th>QDA</th>
<th>SVM</th>
<th>LR</th>
<th>FLDA</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial flora</td>
<td>83.33014</td>
<td>79.66433</td>
<td>90.60692</td>
<td>79.11691</td>
<td>81.08746</td>
<td>79.24898</td>
</tr>
<tr>
<td>Glowworm swarm</td>
<td>83.95856</td>
<td>79.89032</td>
<td>89.62917</td>
<td>81.28533</td>
<td>80.95178</td>
<td>82.65979</td>
</tr>
<tr>
<td>Black hole</td>
<td>85.2518</td>
<td>79.5597</td>
<td>92.17549</td>
<td>78.64844</td>
<td>83.37886</td>
<td>79.64616</td>
</tr>
<tr>
<td>Monkey search</td>
<td>86.06456</td>
<td>82.1462</td>
<td>91.37198</td>
<td>82.34527</td>
<td>81.96818</td>
<td>83.82426</td>
</tr>
</tbody>
</table>

Table 9: Average perfect classification (%) among the classifiers at various optimization techniques with different features for normal cases.

<table>
<thead>
<tr>
<th>Optimization methods</th>
<th>ANN</th>
<th>QDA</th>
<th>SVM</th>
<th>LR</th>
<th>FLDA</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial flora</td>
<td>54.85161</td>
<td>58.17979</td>
<td>70.81364</td>
<td>54.3007</td>
<td>60.15023</td>
<td>65.22285</td>
</tr>
<tr>
<td>Glowworm swarm</td>
<td>54.70138</td>
<td>58.07196</td>
<td>74.2991</td>
<td>66.63904</td>
<td>60.36965</td>
<td>63.75292</td>
</tr>
<tr>
<td>Black hole</td>
<td>54.50142</td>
<td>57.64817</td>
<td>75.09224</td>
<td>54.6082</td>
<td>60.55359</td>
<td>58.7066</td>
</tr>
<tr>
<td>Monkey search</td>
<td>62.95465</td>
<td>59.87399</td>
<td>70.17104</td>
<td>68.95316</td>
<td>59.34051</td>
<td>62.15018</td>
</tr>
</tbody>
</table>

Table 10: Average perfect classification (%) among the classifiers at various optimization techniques with different features for schizophrenia cases.

<table>
<thead>
<tr>
<th>Optimization methods</th>
<th>ANN</th>
<th>QDA</th>
<th>SVM</th>
<th>LR</th>
<th>FLDA</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial flora</td>
<td>66.66027</td>
<td>59.32866</td>
<td>81.20947</td>
<td>58.23381</td>
<td>62.17283</td>
<td>58.49795</td>
</tr>
<tr>
<td>Glowworm swarm</td>
<td>67.91504</td>
<td>59.7798</td>
<td>79.25667</td>
<td>62.57066</td>
<td>61.90357</td>
<td>65.31875</td>
</tr>
<tr>
<td>Black hole</td>
<td>70.49922</td>
<td>59.11939</td>
<td>84.34931</td>
<td>57.29688</td>
<td>66.75354</td>
<td>59.29232</td>
</tr>
<tr>
<td>Monkey search</td>
<td>72.1262</td>
<td>64.2897</td>
<td>82.74271</td>
<td>64.68721</td>
<td>63.9228</td>
<td>67.64852</td>
</tr>
</tbody>
</table>
method. The performance of PI parameter for the ANN and QDA classifier among the four-optimization method is poor, and this is due to the more missed classification and false alarms of the classifier outputs.

Table 12 signifies the average Performance Index among the classifiers at various optimization techniques with different features of schizophrenia cases. As shown in Table 12, LR classifier is reached at low value of PI of 24.1% in the BH optimization method. The poor performance of LR is due to the exhibition of missed classification in the classifier outputs. SVM classifier is outperforming all the classifiers in terms of higher PC, PI, and GDR values and lower error rate value of 27.4%. For a significant accuracy and error rate, KNN classifier closely follows the performance of the SVM classifier.

Table 14 depicts the average performance of parameters among the classifiers at various optimization techniques with different features of schizophrenia cases. As shown in Table 14, LR classifier is reached at low value of PI of 22.01 and PC of 56.7%. FLDA denotes low GDR value of 37.67%. SVM classifier is outperforming all the classifiers in terms of higher PC, PI, and GDR values and lower error rate value of 37.4%. For a significant accuracy and error rate, ANN classifier closely follows the performance of the SVM classifier.

### Table 11: Average performance index (%) among the classifiers at various optimization techniques with different features for normal cases.

<table>
<thead>
<tr>
<th>Optimization methods</th>
<th>ANN</th>
<th>QDA</th>
<th>SVM</th>
<th>LR</th>
<th>FLDA</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial flora</td>
<td>17.53914</td>
<td>26.58122</td>
<td>57.531</td>
<td>15.63139</td>
<td>31.68778</td>
<td>43.07008</td>
</tr>
<tr>
<td>Glowworm swarm</td>
<td>17.00598</td>
<td>26.44067</td>
<td>63.01535</td>
<td>46.80996</td>
<td>31.16374</td>
<td>38.30346</td>
</tr>
<tr>
<td>Black hole</td>
<td>16.29544</td>
<td>25.40426</td>
<td>64.60384</td>
<td>16.62192</td>
<td>29.83996</td>
<td>27.25943</td>
</tr>
<tr>
<td>Monkey search</td>
<td>37.21904</td>
<td>30.80569</td>
<td>53.90694</td>
<td>51.9737</td>
<td>30.50127</td>
<td>34.76854</td>
</tr>
</tbody>
</table>

### Table 12: Average performance index (%) among the classifiers at various optimization techniques with different features for schizophrenia cases.

<table>
<thead>
<tr>
<th>Optimization methods</th>
<th>ANN</th>
<th>QDA</th>
<th>SVM</th>
<th>LR</th>
<th>FLDA</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial flora</td>
<td>47.33871</td>
<td>30.02955</td>
<td>76.75093</td>
<td>26.06912</td>
<td>34.70983</td>
<td>26.38648</td>
</tr>
<tr>
<td>Glowworm swarm</td>
<td>50.81356</td>
<td>29.57847</td>
<td>72.88667</td>
<td>34.9601</td>
<td>33.86657</td>
<td>41.54275</td>
</tr>
<tr>
<td>Black hole</td>
<td>54.83947</td>
<td>27.8239</td>
<td>81.59472</td>
<td>21.0137</td>
<td>45.32586</td>
<td>28.93259</td>
</tr>
<tr>
<td>Monkey search</td>
<td>59.19286</td>
<td>41.3174</td>
<td>78.69396</td>
<td>40.9822</td>
<td>42.56467</td>
<td>49.06799</td>
</tr>
</tbody>
</table>

### Table 13: Average performance of parameters among the classifiers at various optimization techniques with different features for normal cases.

<table>
<thead>
<tr>
<th>Performance parameters</th>
<th>ANN</th>
<th>QDA</th>
<th>SVM</th>
<th>LR</th>
<th>FLDA</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfect classification</td>
<td>56.75227</td>
<td>58.44348</td>
<td>72.59401</td>
<td>61.12578</td>
<td>60.1035</td>
<td>62.45814</td>
</tr>
<tr>
<td>Performance index</td>
<td>22.0149</td>
<td>27.30796</td>
<td>59.76428</td>
<td>32.79244</td>
<td>30.79819</td>
<td>35.85038</td>
</tr>
<tr>
<td>Accuracy</td>
<td>78.37614</td>
<td>79.22174</td>
<td>86.29786</td>
<td>80.56403</td>
<td>80.05268</td>
<td>81.22938</td>
</tr>
<tr>
<td>GDR</td>
<td>54.51164</td>
<td>58.44344</td>
<td>72.43905</td>
<td>59.11251</td>
<td>37.6745</td>
<td>55.96857</td>
</tr>
<tr>
<td>Error rate</td>
<td>43.24759</td>
<td>41.55656</td>
<td>27.4062</td>
<td>38.87433</td>
<td>39.89642</td>
<td>37.54199</td>
</tr>
</tbody>
</table>

### Table 14: Average performance of parameters among the classifiers at various optimization techniques with different features for schizophrenia cases.

<table>
<thead>
<tr>
<th>Performance parameters</th>
<th>ANN</th>
<th>QDA</th>
<th>SVM</th>
<th>LR</th>
<th>FLDA</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfect classification</td>
<td>69.30018</td>
<td>60.62939</td>
<td>81.88954</td>
<td>60.69714</td>
<td>63.68819</td>
<td>62.68939</td>
</tr>
<tr>
<td>Performance index</td>
<td>53.04615</td>
<td>32.18734</td>
<td>77.48157</td>
<td>31.5282</td>
<td>39.11673</td>
<td>36.48245</td>
</tr>
<tr>
<td>Accuracy</td>
<td>84.65127</td>
<td>80.31514</td>
<td>90.94589</td>
<td>80.34899</td>
<td>81.84498</td>
<td>81.3448</td>
</tr>
<tr>
<td>GDR</td>
<td>59.85653</td>
<td>60.18572</td>
<td>80.80855</td>
<td>41.90007</td>
<td>45.78074</td>
<td>52.10981</td>
</tr>
<tr>
<td>Error rate</td>
<td>30.7002</td>
<td>39.37056</td>
<td>18.11075</td>
<td>39.30309</td>
<td>36.31199</td>
<td>37.31102</td>
</tr>
</tbody>
</table>
7. Conclusions and Future Work

The disorders in the areas of the lobes of the brain can lead to schizophrenia. As the lobes of the brain are important for information processing and memory management activities, a huge damage occurs to it due to schizophrenia. The diagnosis, classification, and analysis of schizophrenia spectrum disorders are quite challenging. To incorporate the latest scientific techniques to clinical diagnosis, the scientific community is working very hard. For the brain state interpretation and diagnosis, EEG is emerged as a highly useful and beneficial tool. The proposed method in this work explores and utilizes a plethora of features with four different types of optimization techniques before proceeding to classification. The best results show that, for normal cases, a classification accuracy of 87.54% is obtained when BH optimization is utilized with SVM-RBF Kernel, and for schizophrenia cases, a classification accuracy of 92.17% is obtained when BH optimization is utilized with SVM-RBF kernel. We plan to incorporate other optimization mechanisms with deep learning techniques for schizophrenia EEG signal classification in future work.

Data Availability

Data will be provided to genuine researchers upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (no. NRF-2018R1A2B6006046).

References


