



Seizure disorders recognition model from EEG signals using new probabilistic particle swarm optimizer and sequential differential evolution

Anuradha Thakare¹ · Ahmed M. Anter^{2,3} · Ajith Abraham⁴

Received: 13 December 2021 / Revised: 2 January 2023 / Accepted: 14 January 2023 /
Published online: 13 February 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Epilepsy is a chronic seizure state of an individual. The group of brain cells reflects abnormal electrical activity. Electroencephalography (EEG) is a popular tool that monitors brain activities and diagnoses neurological disorders. The classification of seizure and non-seizure data is a challenging task when dealing with complex transformed features with ultra-high dimensions and extracting the best possible features from EEG. This article presents a new hybrid approach to select the optimal features that involve particle swarm optimization (PSO) algorithm, newly developed probabilistic particle swarm optimization (PPSO) algorithm and sequential differential evolution (SDE) Algorithm. The EEG data of seizure patients have been used to evaluate the method. The features are extracted with Discrete Wavelength Transform. PSO, PPSO and SDE select the optimal features from the feature space of EEG. These optimal features are further evaluated for their performance with different classifiers. The performance of PSO, PPSO and SDE are compared. A broad study of the significance of bio-inspired algorithms for feature optimization for EEG signals is presented. Among all the classifiers, support vector machine (SVM) outperformed, giving an accuracy of 97.74% with PPSO and 98.34% with SDE at the 100th cycle. It demonstrated that optimal feature selection improved the performance of classifiers.

Keywords Epilepsy · Differential evolution · EEG · Discrete wavelet transform · Optimization · Classification · Feature selection

✉ Ahmed M. Anter
Ahmed_Anter@fcis.bsu.edu.eg

Anuradha Thakare
anuradha.thakare@pccoepune.org

Ajith Abraham
ajith.abraham@ieee.org

¹ Department of Computer Engineering, Pimpri Chinchwad College of Engineering, Pune, India

² Egypt-Japan University of Science and Technology (E-JUST), Alexandria 21934, Egypt

³ Faculty of Computers and Artificial Intelligence, Beni-Suef University, Benisuef 62511, Egypt

⁴ Department of Scientific Network for Innovation and Research Excellence, Machine Intelligence Research Labs, Auburn, WA, USA

1 Introduction

The brain disorder with abnormal electrical discharges is called Epilepsy. There are recurrent episodes of this abnormality which leads to seizures. The neurons communicate among themselves called electrical communication. Sometimes, this communication is disrupted called ‘Seizures’ (Harender & Sharma, 2017).

Timely seizure detection is most crucial. The electrical activities of the brain are measured by the Electroencephalogram (EEG). Also, EEG automatically analyzes the abnormality of brain activity.

In case of uncommon epileptic detection of seizures, there is the necessity of long-term EEG recording. The frequent seizure episodes need efficient detection methods which may help the people for their safety. There are many challenges associated with seizure detection. Automated systems are being developed to automatically detect the seizures with clear distinction among instances with seizures and without seizures.

The occurrence of epileptic seizures is random for time scale and the signals undergo the stepwise activities of signal pre-processing, extraction of the feature set and classification. These steps reflect the states of the brain (Thakare & Gore, 2019). Many researchers have defined the approaches to apply the techniques for improvement in the performance of seizure detection (Kalbhor & Harpale, 2016).

EEG classification with sparse Bayesian learning improves the performance of machine learning for the application of brain-computer interface (Jin et al., 2020).

Feature selection is an important method for data preprocessing and machine learning (ML) (Anter et al., 2020a, 2021a; Basha et al., 2021; ElSoud & Anter, 2016). Features may be linear or nonlinear in nature with high dimensional, highly correlated, and the vast number of features, which could largely decrease the detection accuracy (Li et al., 2020). So, it is desired to develop a feature selection method to identify the optimal number of features to ensure the continuity of the model work with the highest efficiency. Feature selection methods based on bioinspired optimization algorithms became more successful in different applications (Anter & Ali, 2020; Anter et al., 2020b, 2020c, 2020d; Azar et al. 2020). Therefore, this study aims to develop a new model based on probabilistic particle swarm optimization (PPSO) and Sequential Differential Evolution (SDE) for optimal subset of neuro-features. In addition, machine learning algorithms can be well trained with the support of evolutionary algorithms like Genetic Algorithms (GA) (Ríos et al., 2020). A distributed control approach for real-time data uses the framework of an adaptive ensemble classification (Sufang, 2020). An enhanced version of Differential Evolution uses roulette wheel selection (Ho-Huu et al., 2018). A well-known structural optimization along with frequency constraints is highly nonlinear. There may be many optimum solutions for complex optimization problems for example; the most efficient way to search the cascade controller’s parameter space is to use particle swarm optimization (PSO). An integrated mathematical model of a parallel robot platform’s hydraulic actuator was employed (Nedic et al., 2014). Moreover, a parallel robot platform can be found by efficiently searching the cascade controller’s parameters using a Swarm Intelligence technique like the firefly algorithm (Nedic et al., 2015).

The long-term EEG recording is needed for the detection of uncommon seizures. The frequent seizure episodes need efficient detection methods are required for frequent seizure episodes and safety of people. The occurrence of epileptic seizures is random for time scale and the signals undergo the stepwise activities of signal pre-processing, extraction of the feature set and classification.

In order to handle the challenges associated with seizure detection and classification, automated systems are being developed which automatically detect the seizures.

The classification of seizure and non-seizure data is a challenging task when dealing with complex transformed features with ultra-high dimensions and extracting the best possible features from EEG. Therefore, there is a need to design and develop a method to select optimal features. This article presents a new hybrid approach to select the optimal features.

The rest of this paper is structured as follows: Sect. 2 introduces the related work to the epileptic EEG data. Section 3 discusses the proposed new hybrid method is designed to generate optimal subsets of features. In Sect. 4, experimental results and analysis are presented. Finally, the conclusion and remarks are presented in the last section of conclusions.

2 Related research

The most widely accepted, authorized data repository is with the University Hospital of Bonn.

The research literature discusses several schemes for the analysis of EEG data. Concerning selected methods, the performance parameters reflect the variations in the values. Accuracy, sensitivity, and specificity are mostly used parameters.

The detection of seizures and related analysis with a deep convolutional neural network automatically detects the presence of seizures. It is time-consuming with EEG signals (Acharya et al., 2018). A deep convolutional neural network with thirteen layers is presented for the detection of the three classes which are normal, parietal and seizure namely.

One of the limitations in clinical applications with limited annotated data is the large training data size. In seizure detection research, the group of supervised classification methods is recommended for accurate detection of activities related to epilepsy and real-time EEG data. From various localized channels, EEG facilitates the records with multi-frequency brain signals which are non-stationary.

EEG signals have complex features. In the classification task, the non-stationary feature of EEG needs to be classified. This is an important aspect of the analysis. The tools are linear and nonlinear and used for the analysis of data. With stationary assumptions, tools can be functional to nonlinear signals. This leads to the loss of information in many cases. Nonlinear tools are available which are majorly suited for the signals which are not stationary and not linear (Quintero-Rincón et al., 2018).

In the process of decomposition of EEG data into several brands, the non-stationary EEG signals for analysis of frequency have several transforms. Fourier Transform (FT) is mostly not suitable for signals which are not stationary. Therefore, Short Time Fourier Transform is recommended. It has one constraint and that is the window of finite length. This gives a poor-quality resolution. The solution to this problem is Wavelet transform (WT). This is for the decomposition of the EEG data (Quintero-Rincón et al., 2018).

The EEG signals decomposition leads to better extraction of the feature. Discrete Wavelet Transform (DWT) is the most promising technique which can be applied to the EEG data for decomposition. The features are similar to averaging the power, standard deviation and mean absolute value are extracted. These features are useful for further analysis with K nearest neighbor (Quintero-Rincón et al., 2018).

Discrete Wavelet Transform (DWT) and Continuous Wavelet Transform (CWT) can be applied to EEG data. The core difficulty of CWT is redundancy. Therefore, for the decomposition, DWT is used. DWT is a not redundant and extremely efficient transform. At different levels, this is used for signal decomposition using wavelets (Quintero-Rincón et al., 2018).

In the classification of EEG signals, the major and important role is with problem algorithm and feature selection. They help in extracting the correct and meaningful features of the objective of the problem.

A genetic algorithm is a well-known optimization technique for the selection of optimal features. Genetic algorithm with classifiers has performed the task of classification with better accuracy (Sathish et al., 2017).

The research on wavelet transformation of wavelets for the abstraction of features and seizure detection was presented. The observation states that the most performing method is wavelet transform. It produces the results efficiently with accuracy and sensitivity (Thakare & Gore, 2019).

A model which is hybrid for the detection of epileptic seizure (Subasi et al., 2019) was established. This work uses methods for optimization like Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). They search for optimal parameters for the task of EEG classification. The DWT is used for the analysis of EEG signals. The calculation of various statistical values of features includes averaging the power, the value of the mean, and standard deviation.

The analysis of frequency and its process uses the Fourier transform for the same. Since nature-wise, EEG signal is not stationary and not linear too; mostly Fourier Transform is not preferred one. Short-Time Fourier Transform (STFT) is a solution to over this (Deshprabhu & Shenvi, 2015).

Intelligent classifiers (Garde et al., 2018) discuss the performance of several classifiers for the classification of EEG data. It is observed that analysis of EEG data before the classification task leads to better performance.

Hybrid swarm intelligence for clustering with multiple objectives (Thakare & Kharche, 2017) was proposed which include the development of the hybrid algorithm for clustering tasks. The hybridization of PSO with traditional clustering algorithms is discussed.

The well-known technique for the classification of patterns is the Support vector machine (Ganji et al., 2019). Fuzzy Neural Network learning algorithms mostly focus on minimizing the risk. SVM gives excellent generalization performance by simultaneously minimizing both empirical and expected risks.

The differential evolution algorithm optimizes the weight of the base learner (Zhang et al., 2017) with an ensemble approach. Several weak learners are combined to improve the classification performance. In feature selection with differential evolution, one of the important steps is the elimination of noisy features (Hancer, 2020). This discovers meaningful knowledge. Differential evolution (DE) is one of the efficient methods for numerical optimization with an evolutionary approach (Shen et al., 2016). In classification, the multi-objective features (Xue et al., 2014) are selected with conflicting objectives that minimize the features by maximizing the accuracy. The cooperative micro-DE is developed (Parsopoulos, 2009) for optimization of a large amount of data. Hybridization of Differential Evolution of self-adaptive nature (Zhao et al., 2011) optimizes large-scale data. The approaches of improvement in DE concerning optimization of high-dimensional function; are discussed in the research article (Brest & Maučec, 2011; Wang et al., 2011).

Differential evolution (DE) is a scheme for identifying global optimal patterns that use a neighborhood-based mutation operator (Das et al., 2009) over continuous spaces but has slow convergence. A group of DE and all variants use the concept of the neighborhood which calculates the neighborhood of each member. The proposed work is about balancing the examination and utilization abilities of DE. It avoids granting burdens of evaluations of function.

The intelligent epileptic seizure model was proposed based on global search whale optimization algorithm and local search genetic algorithm for optimal subset neuro-features selection. In addition the metaheuristic algorithms was applied on extreme learning machine using differential evolutionary (DE) for weights and biases adaption (Anter et al., 2021b).

The examination of features for the detection of patterns misused is focused on intrusion detection systems (Chebrolu et al., 2005). The performance of feature selection algorithms is compared with various classifiers. It is found that the selection of significant input features is important to design a lightweight system. The hybrid architecture is designed for real-world intrusion detection.

The process of selection of features on bi-objective Differential Evolution (Das et al., 2017) is one of the most demanding tasks found in Data Mining and pattern recognition. An improved binary differential evolution simultaneously optimizes the feature selection criteria. The main contribution of this paper are:

- The new hybrid method, Probabilistic Particle Swarm Optimization (PPSO) is designed to generate optimal subset of features.
- Feature are extracted from EEG signals using the coefficient and the energy of the coefficients of the wavelet sub-bands
- The state-of-the-art review of Significance of bio-inspired algorithms for feature optimization for EEG signals is studied and presented.
- The future direction of proposed framework for large-scale data is discussed.

The comparative study of different methods are presented to detect the epileptic seizure disorder in Table 1.

3 Proposed model

The proposed new hybrid method is designed to generate optimal subsets of features. The optimal features are essential for the classification algorithm to work efficiently. This is due to high computations of complex transformed features which always results in a growing mass of feature sets.

The proposed PPSO has the capability of simple computation and rapid convergence as a stochastic search scheme. Algorithms for stochastic optimization are highly problem dependent. It is possible to propose computational complexity, but it won't tell you how much better one algorithm is than another. That depends on what you hope to accomplish. We Implemented PPSO algorithms and the time it takes to reach to satisfactory answer is observed.

The computational burden is majorly specified through optimization problem than algorithm. Selecting optimal feature set using probabilistic function is achieved in the said research. An algorithm is executed for a number of cycles and performance is observed. The objective function parameters change the complexity of an optimization problem and as a result the employed algorithm. The time complexity of PPSO is $O(N^2)$.

The topological neighbors of a particle are tracked by PPSO in order of best position. A new particle velocity is updated at each time step using the individual and global best positions. When PPSO is included in the traditional probabilistic function, the computation is simple and barely increases the computational strain. Premature convergence is also mitigated by the adaptability of PPSO to change the ratio of local to global exploration of the problem space. The number of iterations necessary to get a good outcome varied depending

Table 1 Comparative Analysis of some Existing Methods

Authors	Year	Strength	Limitation
Ru et al. (2022)	2022	An epilepsy detection method in noisy environment. A significant method that overcomes delay in diagnosis	Lack of epilepsy information is caused by filtering. Some epilepsy cannot be detected
Subasi et al. (2019)	2019	A hybrid model to determine the optimum parameters for classification of EEG data	GA and PSO is used to optimize the parameters of only classifier SVM
Harender and Sharma (2017)	2017	Classifier perform for separate and combined statistical features	It uses a smaller number of features to detect seizure disorder
Quintero-Rincón et al., (2018)	2018	It requires only estimating and classifying two scalar parameters	Due to the very high dynamics of epileptic signals, defining the sliding time window is difficult
Acharya et al. (2018)	2018	A 13-layer CNN model is used with highest performance accuracy	This deep network can be difficult to converge. Lack of huge database
Anter et al. (2021b)	2022	The proposed model is based on naive Bayes, genetic algorithm, and whale optimizer (NB-GWOA) for optimal feature selection and ELM guided by DE to provide optimal parameters for an efficient leaning solution	The features are represented by all-time points of EEG signals without feature selection methods
Sathish et al. (2017)	2017	Genetic algorithm can be fine-tuned to produce better result all the time by changing the fitness functions	To train feed forward neural network (FFNN) and support vector machine (SVM) for every iteration is complex and time consuming
Tzimourta et al. (2019)	2019	this is the first methodology validated on both Bonn and Freiburg databases	Discrete Wavelet Transform (DWT) is applied to decompose the signal into sub-bands of interest depending on the sampling frequency

on the objective function. Too few iterations ended the search prematurely, whereas too many iterations resulted in unnecessarily increased computing complexity and required more time.

The proposed method is divided into two phases. In phase I, the features are extracted from the EEG signals using wavelet sub-bands to form the feature vector. In phase II, the optimal feature subsets of EEG signals which are non-stationary are extracted with Sequential Differential Evolution (SDE) methodology and compared with the optimal feature subsets extracted with new Probabilistic Particle Swarm Optimization (PPSO) algorithm.

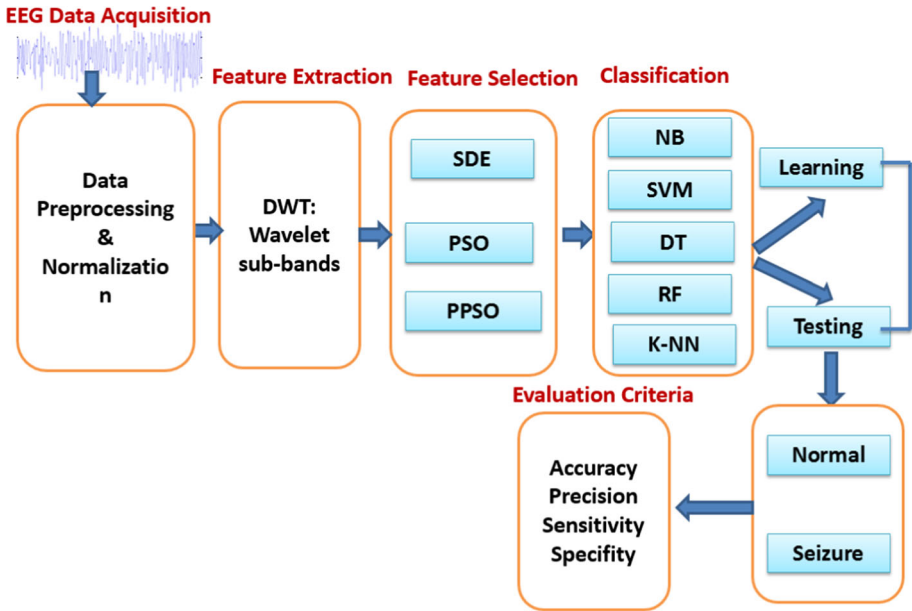


Fig. 1 Workflow of the proposed model. Here, DWT is Discrete Wavelet Transform, SDE is Sequential Differential Evolution, PSO is Particle Swarm Optimization, PPSO is Probabilistic Particle Swarm Optimization, NB is Naïve Bayes Algorithm, SVM is Support Vector Machine, DT is Decision Tree algorithm, RF is Random Forest Algorithm, K-NN is K Nearest Neighbour Algorithm

Generally, the EEG signal computes different frequency bands and are namely the alpha, the beta, the delta, the gamma, and the theta. The sub-bands are decomposed, and the calculated parameters are called features.

In phase III, the performance of SDE and PPSO is evaluated with different classification algorithms which classify seizure and non-seizure EEG signals. Figure 1 represents the workflow of the proposed model.

3.1 Data acquisition and signal decomposition:

In this subsection, the data acquisition process is discussed. The acquisition of EEG data is from the data repository of the Epileptology Section of the University of Bonn (Wang et al., 2011). There are a total of 500 individuals. The five subsets *A* to *E* have 100 1-channel EEG segments. Each file has brain action recorded for the time; 23 s. There are 4097 data points prepared from Time Series samples. Each individual in this dataset has 4097 data points. These are recorded for the time; 23 s.

In a second, each chunk has 178 data points. Therefore, row-wise; it has 23 multiplied by 500, which is equal to 11,500 pieces of information, and column-wise; in a second, 178 data points. There are two sets, *A* and *B* which include the segments, are extracted from the face of EEG recordings (Garde et al., 2018; Tzamourta et al., 2019). This data is about five health-wise fit volunteers. They were in an awake state with the situations of opened eyes (*A*) and closed eyes (*B*).

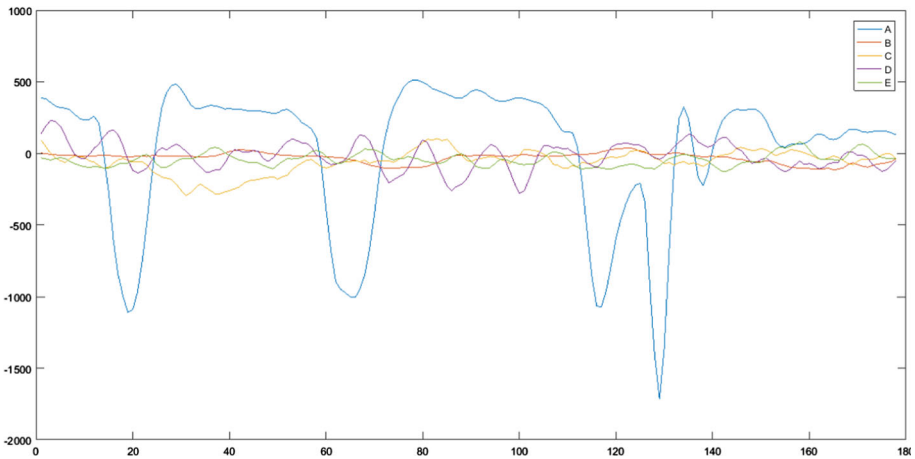


Fig. 2 Different EEG signals for seizure and seizure-free activity. Here set A is situations of opened eyes and set B is closed eyes, sets C and D are Seizure-free activities, set E is tainted with seizure

The EEG repository has a diagnosis of pre-surgical cases. These are taken from five dissimilar patients. Hence, we have other three sets (C, D and E). Seizure-free activity is in sets C and D, whereas; set E is tainted with seizure. Figure 2 shows different EEG signals for the five sets seizure and seizure-free activity.

DWT produces estimated and complete coefficients. The signals are accepted through high pass and low pass filters. This is continued until the preferred occurrence is obtained. In DWT, the signals are decomposed into high and low-frequency sub-band.

Until decomposition of the entire signal, it is divided into sub-bands of high and low frequency. Approximation coefficients are the low frequencies coefficients and detail coefficients are with high frequencies. The decomposition of EEG data into various bands leads to its frequency range like delta, theta, alpha, beta, and gamma.

3.2 Feature extraction and feature selection

In literature, extraction takes place with the statistical and non-statistical domains. In this research, the features are computed for an alpha, beta and gamma band. These six features are evaluated, and the mean absolute value of the sub-band is with coefficients of wavelet. For each coefficient, a variance is taken of the wavelet in the sub-bands. A sub-band has both the standard deviation as the average power of the wavelet coefficients and the energy of the coefficients of the wavelet.

The information of frequency of the signal is calculated with Mean. The information regarding the frequency contents of the signal is received from average power. The magnitude of the variation is evaluated by computing the standard deviation in the signal frequency. The approximation A4 and details D1-D4 are derived from these features as shown in Fig. 3. The resulting feature vectors classify EEG data. Here, four levels are selected. To analyze the EEG recordings, the Daubechies wavelets are selected and experimented with. The extraction of coefficients of wavelet cannot be considered to input the classifier as there will be an increase in the classifier's dimensionality.

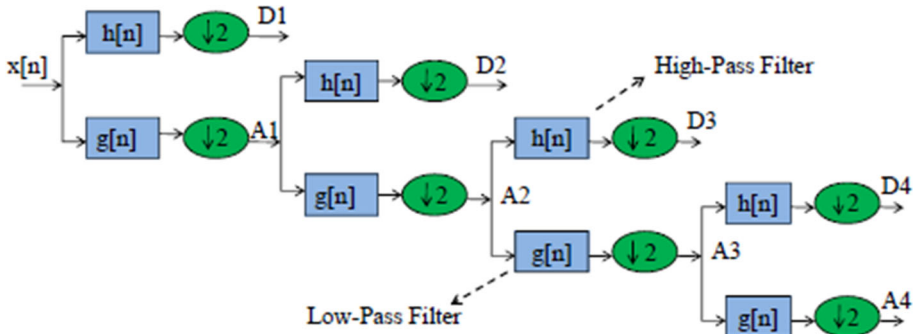


Fig. 3 Four level wavelet decomposition of the EEG

The DWT transform has been successfully used for the research on seizure dataset. In this research, the data is preprocessed, and noise is removed using gaussian filter before applying DWT. Therefore, filtering process is not required to remove noise.

The features considered for selection may vary in classification accuracy. In the proposed work, SDE, PSO and PPSO select the prominent feature vectors.

3.3 Feature analysis using analysis of variance (ANOVA)

The one-way ANOVA test has been performed with one independent variable called attention and dependent variable, predefinedlabel. Statistics shows, degree of freedom (df) is 1.0 that represents maximum number of logically independent values, these values have the freedom to vary, in the data sample. The F statistic shows the variability among the samples and within samples is 0.03113.

Similarly, the one-way ANOVA test has been performed for categorical and numerical features (all features). The df and F values are calculated. Below Table 2 shows the statistics for 1 way ANOVA for all features.

3.4 Probabilistic particle swarm optimization(PPSO) for optimal feature selection

The parameters of the classification algorithm can be discovered using optimization methods. PPSO is a newly developed hybrid algorithm. It uses the probabilistic function to calculate the probability that an unknown pattern belongs to a particular class. The fitness function of PSO is the probability value calculated (Thakare & Kharche, 2017). The basic logic of PSO is modified for the classification task. PSO and PPSO both fall in the category of nature-inspired computing. PSO is popular for optimization whereas PPSO is developed and tested on various datasets for optimization results.

The Naive Bayes technique is used in PPSO to predict the class of the test dataset. When assuming independence holds and less training data is required, a naive Bayes classifier outperforms other classifiers such as logistic regression. The Naive-Bayes method, at its most basic level, assesses the probability of a feature’s connection to a target variable and then selects the feature with the highest probability. This concept is applied to the construction of the Probabilistic Particle Swarm Optimization method, which is used to choose the best characteristics. Probability is determined and fed into the PSO algorithm’s fitness function

Table 2 The statistics for 1 way ANOVA for all EEG features

	sum_sq	df	F	PR(> F)
C{predefinedlabel}	3.686578e + 04	1.0	107.509701	4.326723e-25
SubjectID	2.081054e + 04	1.0	60.688647	7.199484e-15
Video ID	5.191777e + 04	1.0	151.405002	1.340862e-34
Med iation	1.337088e + 08	1.0	3899.278742	0.000000e + 00
Raw	2.252243e + 03	1.0	6.568095	1.039363e-02
Delta	2.596543e + 04	1.0	75.721591	3.662415e-18
Theta	9.760643e + 03	1.0	28.464440	9.706193e-08
Alphal	2.113429e + 04	1.0	61.632787	4.467087e-15
Alpha2	1.450451 e + 04	1.0	42.298734	8.125640e-11
Betal	3.454187e + 03	1.0	10.073260	1.507927e-03
Beta2	1.258355e + 04	1.0	36.696737	1.41882Be-09
Gammal	1.886511 e + 04	1.0	55.015312	1.271280e-13
Gamma2	2.127858e + 03	1.0	6.205358	1.274899e-02
Residual	4.388175e + 06	12,797.0	NaN	NaN

as an input. Figure 4 represents the workflow of probabilistic particle swarm optimization algorithm and Table 3 represents pseudo-code of probabilistic particle swarm optimization algorithm.

3.5 Sequential differential evolution (SDE) for optimal feature selection

Differential Evolution (DE) is a technique of evolutionary computation to optimize the parameters of a problem and concerning a given measure of quality, tries to improve the candidate solutions in a search space through the number of generations. This method is proven to be the best method for optimizing the large search space. Like other evolutionary methods, SDE also results in the best possible solution or near-optimal solution.

The gradient of the problem which is being optimized is not the part of DE. It is represented as a real-valued function that is multidimensional. This states that the problem of finding optimal need not be differentiable. This is the need for the traditional methodology of optimization methods which include gradient descent and quasi-newton methods. DE is applicable on noisy, continuous, or change over time etc. optimization problems (Price et al., 2006; Rocca et al., 2011).

With the given population of aspirant solutions, DE works on the optimality of a problem and creates new aspirants or candidate solutions. This is done by combining existing with simple formulae and finally considering only that candidate solution that has the best score or highest fitness for the objective of the optimization problem. Therefore usually, the problem of optimization is considered as a black box that provides a quality measure for a given candidate solution. This is the reason; it doesn't require gradient.

The Sequential Differential Evolution Algorithm extracts the optimal features from feature space. These optimal features are further evaluated for their performance with different classifiers.

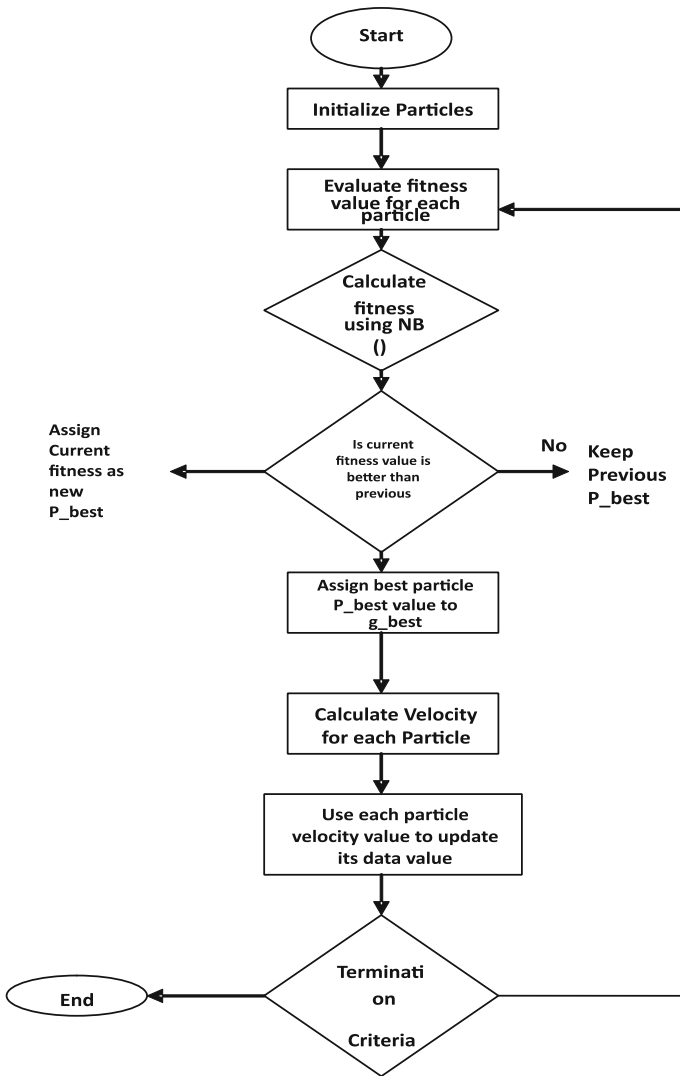


Fig. 4 The workflow of probabilistic particle swarm optimization algorithm

4 Population formation

DE’s nearby all adaptable development have a paired vector populations. Both the populations enclose N_{pop} T-dimensional vectors of the real-valued parameter. Suppose the population is denoted by Q_x and is a collection of the vector components, $Y_{m,g}$. These vectors are acceptable as initial points and compared to all the available vectors in the population.

$$Q_x, g = (Y_m, g), \quad m = 0, 1, \dots, N_{pop} - 1, \quad g_t = 0, 1, \dots, g_t \text{ max}, \quad (1)$$

$$Y_i, g = (Y_n, i, g), \quad n = 0, 1, T - 1 \quad (2)$$

Table 3 Pseudo-code of PPSO algorithm

<p>Calculating Fitness using Naive Bayes from sklearn.naive bayes Create an instance of Naive Bayes classifier = NB()</p> <p>Objective function definition: def f_per_particle(m, alpha):</p> <p>Calculate for the objective function per particle all_features = 30</p> <p>Use binary mask to obtain the subset of the features if np.count_nonzero(m) == 0: // subset of the features from the binary mask X_subset = X else: X_subset = X[:,m==1]</p> <p>Execute and store result in C classifier.fit(X_subset, y) C = (classifier.predict(X_subset) == y).mean() return j</p> <p>PSO procedure: def f(x, alpha=0.88): n_particles = x.shape[0] j = [f_per_particle(x[i], alpha) for i in range(n_particles)]</p>
--

Here, we need to work with arrays and modular arithmetic so; indices starting with 0 are preferred. A generation of specific vector, g represents the index, and it starts from 0 to gt_{max} with the increment of 1. I is a population index that is assigned to all vectors. This has values from 0 to $N_{pop} - 1$. Vectors along with their set of parameters are indexed from 0 to $(T - 1)$.

The operator mutation with DE is usually random. To construct an intermediary population, $Q_{v,g}$, of N_{pop} mutant vectors, $R_{m,g}$, randomly vectors are chosen once initialized.

$$Qv, g = (Rm, g), m = 0, 1, \dots, N_{pop} - 1, gt = 0, 1, \dots, gt_{max} \quad (3)$$

$$Rm, g = (Rn, m, g), n = 0, 1, T - 1 \quad (4)$$

The vectors in the present population are crossed with a mutant produced. In this way, a trial population, Qu , of N_{pop} trial vectors, $u_{i,g}$ is generated.

$$Qu, g = (u_m, g), m = 0, 1, \dots, N_{pop} - 1, gt = 0, 1, \dots, gt_{max} \quad (5)$$

$$u_m, g = (u_n, m, g), n = 0, 1, T - 1 \quad (6)$$

During crossover or recombination, the trial vectors mostly overwrite the mutant population when recombination happened. Therefore, a single array is capable of holding both the populations.

5 Initialization

When the population is initialized, the initial settings of parameters are important. One should specify for each parameter the upper and lower bounds before the initialization of the parameters. The initialization vectors are two-dimensional, and these values are collected in these vectors. The upper and lower bounds will be specified after initialization. In a prescribed range, a random number generator is used to assign a value to each parameter of the vector.

6 Mutation

The mutation is an operator which evolves the various target regions in the population. DE performs mutation and recombination to produce a newly generated population of Np trial vectors. This is referred to as a *Differential mutation*. This has highly scaled, sampled randomly, vector difference to a resultant third vector. This creates the mutant vector which is the blend of these three vectors which are randomly chosen.

The formula can be evolved by using scale factor, *base vector*, *target vector* and *difference vector*. A population is evolved at a rate that is controlled by a positive real number called the scale factor which has no upper limit. The effective values are greater than the *base vector* and *target vector*. There are various ways to determine the *base vector* index. Here, we assumed that it is just a vector index that is randomly chosen and is completely dissimilar to the *target vector* index.

The dissimilar vector indices are distinct from each other. These are also distinct from base and target vector indices. For each mutation, these different vectors are randomly selected.

7 Crossover

The crossover strategy is also called the recombination technique. To harmonize the mutation, which is differential in nature, the uniform crossover strategy is employed with the DE model. Crossover works with the parameter values of vectors which are taken from two dissimilar vectors in nature. DE performs crossover with all vectors with a mutant vector. A user can define the value of crossover probability to control the values of parameters. These are the values copied from the mutant. It is important to establish the various sources contributing to a given parameter. The output of uniform crossover and uniform random number generator are compared.

8 Selection

In each run, there is a comparison of the trial vector with the target vector with the objective function value. If this value of the trial vector is minor or equal to the values of the target vector, there is a substitute for the target vector. In a successive generation, the trailing vector

becomes the target vector. DE compares every one trial vector with the target vector. The parameters are inherited from these vectors. DE is more powerful and can tightly put together crossover and selection over the existing evolutionary methods.

Upon successful installation of the new population, these processes of mutation, crossover and selection are repeated. This cycle is executed until the termination criteria or specific optimum is located. For example, one can write a source code to compare the number of generations with the present maximum, g_{max} .

8.1 Classification of optimal feature subsets

The decomposition of sub-bands of parameters calculates the features. The wavelet coefficients are considered for calculating the features. The feature selection process optimizes the feature subset which is employed as input. The classifier used these features for classification. Naïve Bayes, Support Vector Machine, Random Forest, k-Nearest Neighbor and Decision Tree algorithms are used for this research.

The features are extracted at various decomposition levels. The seizure abnormality is detected with classifiers and accuracy is measured.

8.1.1 Naïve Bayes algorithm (NB)

This algorithm easily and fastly predicts the class of test data set. In terms of training data, it works better compared to logistic regression and others. This algorithm calculates the relation of the feature with the target variable. This gives a notion to match the probability. The selected feature set has the highest probability (Anter et al., 2021b).

NB is a probability-based method of classification. It uses Bayesian theory. In search space, each feature of a given class is self-determining. Less training data for classification is required.

8.1.2 Support vector machine (SVM)

SVM is an algorithm that is supervised in nature. The initial training is with the tagged EEG feature set. These tags are with seizure or without a seizure. The samples of the training phase build a model to assign the input samples category-wise. This is considered a new set.

The points indicate training samples and are separated as a category. This constructs a hyperplane and sometimes a set with hyperplanes. There are supportive vectors nearby the hyper plane. SVM looks for promising hyperplanes. An objective is to raise the margin linking the support vectors. SVM encourages high-dimensional data and produces good results (Ho-Huu et al., 2018).

8.1.3 Random forest (RF)

The Random Forest is a widely used method. As per research studies, for complex data RF is the better choice for the classification task. If RF is trained systematically, it can optimize and predict various types of conditions of patients. RF is capable of classifying various degrees of predictions in epilepsy and works with multiple conditions. This model is never much responsive to the deletion of partial features. Therefore, the rule of classification becomes more robust (Rajesh et al., 2015; Wang et al., 2019).

8.1.4 K-nearest neighbor (k-NN)

K-Nearest Neighbor is a method for classification which is pretty straightforward and non-linear. Even with large training data, this algorithm performs well. This algorithm is simple as it won't require the separate training and testing phases (Quintero-Rincón et al., 2018; Sathish et al., 2017).

8.1.5 Decision tree classifier (DT)

Among all learning algorithms, one of the opted algorithms for machine learning is a Decision tree. In the decision tree, nodes represent attributes or features, a decision rule is represented by branch or links and the outcome is towards each leaf. It is good in classifying observed data. Inverted tree representation is an outcome of the Decision Tree technique (Castro, 2006; Rajesh et al., 2015).

9 Result and analysis

This section discusses the result and analysis with several machine learning algorithms and performance measures. A wavelet-based method, Discrete Wavelet Transform (DWT) with four-level decomposition automatically detects the epileptic activity by extracting the major features which contribute to detection. Daubechies (db4) is selected for this work which calculates Mean, Variance, Standard deviation, Average power and Energy in each sub-band. These bands are alpha, beta and gamma bands. These bands reflect abnormalities of the brain activity and the availability of seizure disorder which are considered as major extracted feature vectors.

DWT with four-level decompositions extracts the feature vectors are further optimized with three methods namely, Particle Swarm Optimization (PSO), new Probabilistic Particle Swarm Optimization (PPSO) and Sequential Differential Evolution (SDE). The resulting features are selected and feature vectors are generated.

The machine learning algorithms classify the features extracted with DWT and selected with PSO, PPSO and SDE into two groups describing features accountable for seizures and features not contributing seizure description.

In the said research, the performance of various Machine Learning algorithms for classification of feature set with and without optimization are compared. The experimented machine learning algorithms are Naïve Bayes Algorithm (NB), Support Vector Machine (SVM), Decision Tree algorithm (DT), Random Forest Algorithm (RF) and K Nearest Neighbour Algorithm (KNN). Initially, the feature vectors are extracted with Discrete wavelet transform. The classification performance of the Support Vector Machine supersedes other classifiers. Then the optimal feature set are selected using three algorithms: Particle Swarm Optimization (PSO), newly proposed Probabilistic Particle Swarm Optimization (PPSO) and Sequential Differential Evolution (SDE). The classification of these optimal feature set is compared using same set of classification algorithms. The results of all the feature optimization and selection methods are recorded for number of cycles and the statistical results are presented.

The classification accuracy is evaluated with the performance measures accuracy; sensitivity and specificity.

The performance of the proposed method is evaluated using measures for accuracy, sensitivity and specificity because confusion matrix, is also known as an error matrix. It allows visualization of the performance of the proposed algorithm and comparison with other algorithms. Each row of the matrix represents the instances in a predicted class while each column represents the instances in an actual class (or vice versa).

		Predicted	
		Yes	No
Actual	Yes	TP	FN
	No	FP	TN

Accuracy is calculated for training and validation. It gives correct classification rate using following Eq. (7). Here, True Positive (TP) is positive tuples that were correctly labeled. True Negative (TN) is negative tuples that were correctly labeled. False Negative (FN) is positive tuples that were mislabeled as negative and False Positive (FP) is Negative tuples that were incorrectly labeled as positive.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} * 100 \tag{7}$$

Sensitivity is the true positive rate or recall that measures the proportion of actual positives that are correctly identified and specificity is true negative rate which measures the proportion of actual negatives that are correctly identified. Specificity and sensitivity are calculated using Eqs. (8) and (9).

$$Sensitivity \frac{TP}{P} \tag{8}$$

$$Sensitivity \frac{TN}{P} \tag{9}$$

Table 4 depicts the results of the classification of feature vectors with DWT. These results are tabulated to compare the classification accuracy with and without optimization of the feature set.

The Classifiers are evaluated for their performance. It is tabulated without using the process of feature selection. Discrete wavelet transform extracts the features from the EEG data. The classification performance of the Support Vector Machine supersedes other classifiers with

Table 4 Classification of Feature Extracted with DWT

Algorithms	Classification of feature vectors		
	Accuracy	Sensitivity	Specificity
NB	95.13	96.69	88.79
SVM	95.86	95.74	96.48
RF	89.20	87.21	88.34
k-NNN	89.60	88.47	100
DT	92.10	92.53	91.82

Table 5 Comparison of performance of feature selection algorithm at 20th cycle

Feat. Selection Algo	Classifica. Algo	Accuracy	Sensitivity	Specificity
PSO	NB	95.48	96.21	92.27
	SVM	97.26	97.84	94.87
	RF	92.44	92.18	90.10
	k-NN	97.13	97.99	93.65
	DT	94.72	94.59	90.83
PPSO	NB	95.09	96.14	90.55
	SVM	97.57	98.26	94.75
	RF	93.23	92.45	90.74
	k-NN	97.43	98.21	94.32
	DT	95.12	95.45	92.69
SDE	NB	94.32	95.47	90.12
	SVM	98.01	98.20	94.39
	RF	96.10	95.66	94.48
	k-NN	95.77	96.52	92.60
	DT	97.67	97.52	93.33

accuracy as a measure. These resultants are documented to make a comparison with the classification results of optimized feature vectors. In this section, further the results of feature selection methods, PSO, PPSO and SDE are compared with classification algorithms. The results of all the feature optimization and selection methods are recorded for cycles 20, 40, 60, 80 and 100. The SVM classifier outperformed giving accuracy of 97.53% with PSO, 97.74% with PPSO and 98.34% with SDE at 100th cycle, converged state.

The statistical results are presented in Tables 5, 6, 7, 8 and 9.

10 Observations on results

The experimentation is done on the dataset described in subsection data acquisition and decomposition. The results of the experimentation are presented in Tables 2, 3 and 4. The algorithms are converged when the number of generations exceeds 100. Here the results of all the feature optimization and selection methods are recorded for cycles 20, 40, 60, 80 and 100. The performance is closely monitored and statistical results are tabulated in Tables 5, 6, 7, 8 and 9 respectively. Following are the detailed discussions.

- i. Among all the classifiers, Random Forest and KNN outperform when features are extracted from the dataset with DWT, as depicted in Table 4. Here the intention is to understand how classifiers perform with feature extraction.
- ii. The features extracted are now optimized with the evolutionary optimization techniques, PSO, PPSO and SDE. Here, the intention is to optimize the feature set and select the feature vectors. The performance is evaluated with several classifiers. The accuracy of classifying every instance as either seizure or non-seizure class is well tabulated in Tables 5, 6, 7, 8 and 9. It has been observed that the classification accuracy is improved with an optimized feature set than accuracy with features extracted with DWT.

Table 6 Comparison of performance of feature selection algorithm at 40th cycle

Feat. Selection Algo	Classific. Algo	Accuracy	Sensitivity	Specificity
PSO	NB	95.09	96.34	89.82
	SVM	97.40	98.37	95.36
	RF	92.48	92.32	90.15
	k-NN	97.22	97.74	95.06
	DT	95.49	95.33	91.17
PPSO	NB	95.57	96.46	91.72
	SVM	97.65	98.48	94.38
	RF	94.86	93.49	92.73
	k-NN	97.57	98.32	94.55
	DT	96.39	95.89	92.99
SDE	NB	94.81	95.68	90.21
	SVM	98.24	98.19	94.39
	RF	96.10	95.66	94.48
	k-NN	95.85	96.55	92.64
	DT	97.72	97.71	93.56

Table 7 Comparison of performance of feature selection algorithm at 60th cycle

Feat. Selection Algo	Classificat.Algo	Accuracy	Sensitivity	Specificity
PSO	NB	95.26	96.4	90.45
	SVM	97.59	98.48	94.79
	RF	92.52	92.64	90.28
	k-NN	96.87	97.83	93
	DT	95.78	95.43	91.32
PPSO	NB	95.43	96.26	91.86
	SVM	97.69	98.06	95.97
	RF	95.47	94.39	93.36
	k-NN	97.48	98.53	93.39
	DT	96.86	95.92	93.24
SDE	NB	95.19	96.33	90.35
	SVM	98.29	98.27	94.57
	RF	96.14	95.71	94.57
	k-NN	96.64	97.25	93.23
	DT	97.81	97.80	93.71

Table 8 Comparison of performance of feature selection algorithm at 80th cycle

Feat. Selection Algo	Algorithms	Accuracy	Sensitivity	Specificity
PSO	NB	94.83	95.98	89.86
	SVM	97.67	98.53	94.81
	RF	93.12	93.66	90.12
	k-NN	97.48	98.37	93.95
	DT	96.12	96.65	92.01
PPSO	NB	95.04	96.14	90.34
	SVM	97.7	98.48	94.59
	RF	95.86	94.68	93.53
	k-NN	97.43	97.74	96.14
	DT	96.91	95.92	93.66
SDE	NB	95.28	96.41	90.43
	SVM	98.32	98.31	94.79
	RF	96.30	95.84	94.61
	k-NN	96.77	97.32	93.38
	DT	97.85	97.83	93.74

Table 9 Comparison of performance of feature selection algorithm at 100th cycle

Feat. Selection Algo	Algo	Accuracy	Sensitivity	Specificity
PSO	NB	95.30	96.45	90.48
	SVM	97.53	98.16	94.93
	RF	93.54	93.69	92.25
	k-NN	97.30	98.00	94.48
	DT	95.45	95.40	90.48
PPSO	NB	95.22	96.35	90.43
	SVM	97.74	98.48	94.79
	RF	94.55	94.62	93.55
	k-NN	97.56	98.33	94.54
	DT	96.32	96.73	92.65
SDE	NB	95.21	96.32	90.33
	SVM	98.34	98.34	94.83
	RF	96.32	95.89	94.65
	k-NN	96.83	97.44	93.42
	DT	97.87	97.88	93.79

- iii. Among all the classifiers, SVM outperformed giving an accuracy of 97.53% with PSO, 97.74% with PPSO and 98.34% with SDE at the 100th cycle. The results of PSO and its improved version PPSO are compared and it is found that PPSO supersedes the PSO in terms of accuracy for almost all the classifiers. PPSO is the improved version of PSO and results have demonstrated the improved performance.
- iv. The performance of PPSO and SDE are compared. It is found that all the classifier results in the increased accuracy of classification for all the instances as seizure and nonseizure except Naïve Bayes and KNN. The performance of Naïve Bayes is increased with PPSO due to multiple probability calculations. PPSO uses a hybrid fitness function which is formed by the combination of probabilistic function and objective function of PSO. Therefore there is an improvement from PSO to PPSO but, with SDE the performance is degraded. KNN calculates nearest neighbors whereas these calculations show significant improvement if there is diversity in the feature set.

10.1 Comparisons among different deep neural network methods

This subsection compares the proposed method to various deep neural network methods (including the extreme learning machine (ELM) with Hard limit activation function, the cascade-forward neural network (CFNN), the recurrent neural network (RNN), the feed-forward neural network (FFNN) with 10 hidden layers and 10 maximum epochs, and the probabilistic neural network (PNN)) to find the most effective solution or non-linear EEG seizure detection problem. Table 10 displays the comparison's results using various measurements. The proposed method provided the most trustworthy findings, as shown in Table 10, followed by the ELM's results, which had an accuracy rate of about 94%, a sensitivity rate of about 39%, and a specificity rate of about 97%. These findings demonstrate that the proposed method accuracy was enhanced by the sequential differential evolution (SDE) and support vector machine (SVM) integration. Additionally, as demonstrated in Fig. 5, the proposed method outperforms deep neural network methods like (FFNN, CFNN, and RNN) and produces high accuracy with minimal error and computationally quick results. These findings demonstrate that the proposed method has superior performance and successful results in the challenging, non-linear dataset for epileptic seizures. See Table 10 and Fig. 5 for further details about various deep neural network methods results.

Table 10 Performance results of the proposed method and different deep neural network methods

Methods	Accuracy	Error	Sensitivity	Specificity
FFNN	0.740	0.260	0.844	0.639
CFNN	0.729	0.271	0.792	0.648
RNN	0.746	0.254	0.834	0.651
PNN	0.738	0.262	0.710	0.853
ELM	0.938	0.062	0.389	0.971
SDE-SVM	98.34	0.017	98.34	94.83

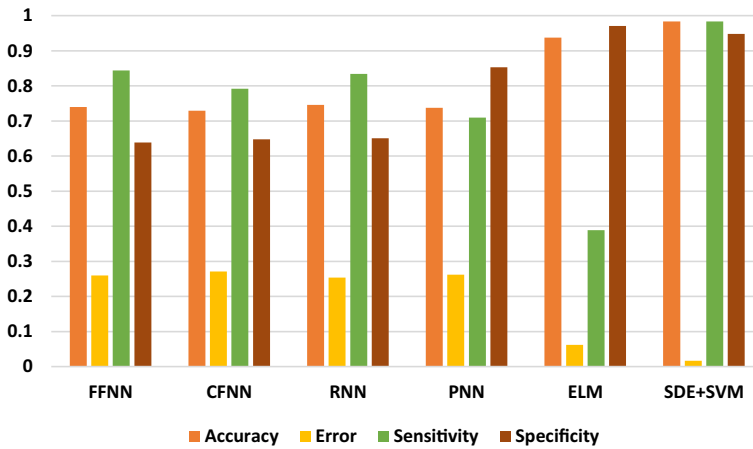


Fig. 5 Visual results for the performance of the proposed method and different deep neural network methods

11 Conclusion and future work

This article presents a broad study of the significance of bio-inspired algorithms for feature optimization for EEG signals. The EEG data of seizure patients have been used to evaluate the method.

The proposed PPSO is developed by the authors. There are various swarm intelligence techniques and many researchers are using them for optimization purposes. PSO is one of the best algorithms among them. Proposed PPSO is a hybrid version of PSO and pseudocode is presented in the paper for research community.

A new hybrid method is developed to select the optimal features. DWT extracts the features from EEG data with four levels of decomposition. The performance is examined with several classifiers. The newly developed Probabilistic Particle Swarm Optimization (PPSO) algorithm and Sequential Differential Evolution (SDE) Algorithm selects the optimal features from the feature space of EEG. The optimal features selected using PSO, PPSO and SDE are further evaluated for their performance with different classifiers. The classification results of PSO, PPSO and SDE are compared. The experimentation results demonstrated that among all the classifiers, SVM outperformed giving accuracy of 97.53% with PSO, 97.74% with PPSO and 98.34% with SDE at 100th cycle, converged state. It is observed that feature extraction and a further selection of optimal features result in improved classification accuracy of EEG data as seizure and non-seizure. In the future work, we will address the problem of parameter optimization to improve the convergence and performance of the proposed algorithm using iterative learning model (Tao et al., 2021). The proposed framework with ML accelerators may be used for large-scale data. It can be applied to other datasets with higher diversity to get the optimal classification results. The bigger size data sets may require specialized hardware to improve performance. Empirical Wavelet Transform (EWT) and Variational Mode Decomposition (VMD) methods build adaptive wavelets for signal decomposition and apply the filtering to remove the noise. The DWT technique is used for said research since it is a dataset independent method, provides better signal compression and data is well transferred into frequency domain. In future work, EWT and VMD will be used for feature extraction.

Declarations

Conflict of interest There is no conflict of interest in this research. This work is not funded by any agency.

References

- Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adeli, H. (2018). Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Computers in Biology and Medicine*, *100*, 270–278.
- Anter, A. M., Abd Elaziz, M., & Zhang, Z. (2021b). Real-time epileptic seizure recognition using Bayesian genetic whale optimizer and adaptive machine learning. *Future Generation Computer Systems*, *127*, 426–434.
- Anter, A. M., & Ali, M. (2020). Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems. *Soft Computing*, *24*(3), 1565–1584.
- Anter, A. M., Bhattacharyya, S., & Zhang, Z. (2020d). Multi-stage fuzzy swarm intelligence for automatic hepatic lesion segmentation from CT scans. *Applied Soft Computing*, *96*, 106677.
- Anter, A. M., Gupta, D., & Castillo, O. (2020b). A novel parameter estimation in dynamic model via fuzzy swarm intelligence and chaos theory for faults in wastewater treatment plant. *Soft Computing*, *24*(1), 111–129.
- Anter, A. M., Huang, G., Li, L., Zhang, L., Liang, Z., & Zhang, Z. (2020c). A new type of fuzzy-rule-based system with chaotic swarm intelligence for multiclassification of pain perception from fMRI. *IEEE Transactions on Fuzzy Systems*, *28*(6), 1096–1109.
- Anter, A. M., Moemen, Y. S., Darwish, A., & Hassanien, A. E. (2020a). Multi-target QSAR modelling of chemo-genomic data analysis based on extreme learning machine. *Knowledge-Based Systems*, *188*, 104977.
- Anter, A. M., Oliva, D., Thakare, A., & Zhang, Z. (2021a). AFCM-LSMA: New intelligent model based on Lévy slime mould algorithm and adaptive fuzzy C-means for identification of COVID-19 infection from chest X-ray images. *Advanced Engineering Informatics*, *49*, 101317.
- Azar, A. T., Anter, A. M., & Fouad, K. M. (2020). Intelligent system for feature selection based on rough set and chaotic binary grey wolf optimisation. *International Journal of Computer Applications in Technology*, *63*(1–2), 4–24.
- Basha, S. H., Anter, A. M., Hassanien, A. E., & Abdalla, A. (2021). Hybrid intelligent model for classifying chest X-ray images of COVID-19 patients using genetic algorithm and neutrosophic logic. *Soft Computing*, 1–16.
- Brest, J., & Maučec, M. S. (2011). Self-adaptive differential evolution algorithm using population size reduction and three strategies. *Soft Computing*, *15*(11), 2157–2174.
- Chebrolu, S., Abraham, A., & Thomas, J. P. (2005). Feature deduction and ensemble design of intrusion detection systems. *Computers & Security*, *24*(4), 295–307.
- Das, S., Abraham, A., Chakraborty, U. K., & Konar, A. (2009). Differential evolution using a neighborhood-based mutation operator. *IEEE Transactions on Evolutionary Computation*, *13*(3), 526–553.
- Das, S., Chang, C. C., Das, A. K., & Ghosh, A. (2017). Feature selection based on bi-objective differential evolution. *Journal of Computing Science and Engineering*, *11*(4), 130–141.
- De Castro, L. N. (2006). *Fundamentals of natural computing: Basic concepts, algorithms, and applications*. CRC Press.
- Deshprabhu, A. A., & Shenvi, N. (2015). Sub-band decomposition of EEG signals and feature extraction for epilepsy classification. *International Journal of Advanced Research in Computer and Communication Engineering*, *4*(3), 108–111.
- ElSoud, M. A., & Anter, A. M. (2016). Computational intelligence optimization algorithm based on meta-heuristic social-spider: Case study on CT liver tumor diagnosis. *Computational Intelligence*, *7*(4), 466–475.
- Ganji, H., Khadivi, S., & Ebadzadeh, M. M. (2019). Support vector-based fuzzy classifier with adaptive kernel. *Neural Computing and Applications*, *31*(7), 2117–2130.
- Garde, P., Thakare, A., Biradar, A., & Pawar, N. (2018). Comparative Study of intelligent classifiers for EEG data. *International Journal of Scientific Research Engineering & Technology (IJSRET), UGC Approved Journal*, *7*(9).
- Hancer, E. (2020). New filter approaches for feature selection using differential evolution and fuzzy rough set theory. *Neural Computing and Applications*, *32*(7), 2929–2944.

- Harender, B., & Sharma, R. K. (2017, May). DWT based epileptic seizure detection from EEG signal using k-NN classifier. In *2017 international conference on trends in electronics and informatics (ICEI)* (pp. 762–765). IEEE.
- Ho-Huu, V., Nguyen-Thoi, T., Truong-Khac, T., Le-Anh, L., & Vo-Duy, T. (2018). An improved differential evolution based on roulette wheel selection for shape and size optimization of truss structures with frequency constraints. *Neural Computing and Applications*, 29(1), 167–185.
- Jin, Z., Zhou, G., Gao, D., & Zhang, Y. (2020). EEG classification using sparse Bayesian extreme learning machine for brain–computer interface. *Neural Computing and Applications*, 32(11), 6601–6609.
- Kalbhori, S. D., & Harpale, V. K. (2016, August). The review of detection and classification of epileptic seizures using wavelet transform. In *2016 international conference on computing communication control and automation (ICCUBEA)* (pp. 1–5). IEEE.
- Li, J., Zhang, S., Zhang, L., Lei, C., & Zhang, J. (2020). Unsupervised nonlinear feature selection algorithm via kernel function. *Neural Computing and Applications*, 32(11), 6443–6454.
- Nedic, N., Prsic, D., Dubonjic, L., Stojanovic, V., & Djordjevic, V. (2014). Optimal cascade hydraulic control for a parallel robot platform by PSO. *The International Journal of Advanced Manufacturing Technology*, 72(5), 1085–1098.
- Nedic, N., Stojanovic, V., & Djordjevic, V. (2015). Optimal control of hydraulically driven parallel robot platform based on firefly algorithm. *Nonlinear Dynamics*, 82(3), 1457–1473.
- Parsopoulos, K. E. (2009, July). Cooperative micro-differential evolution for high-dimensional problems. In *Proceedings of the 11th annual conference on Genetic and evolutionary computation* (pp. 531–538).
- Price, K., Storn, R. M., & Lampinen, J. A. (2006). *Differential evolution: a practical approach to global optimization*. Berlin: Springer.
- Quintero-Rincón, A., Pereyra, M., d’Giano, C., Risk, M., & Batatia, H. (2018). Fast statistical model-based classification of epileptic EEG signals. *Biocybernetics and Biomedical Engineering*, 38(4), 877–889.
- Rajesh, K., Sabarinathan, V., Sarath Kumar, V., & Sugumaran, V. (2015). Eye state prediction using EEG signal and C4.5 decision tree algorithm. *International Journal of Applied Engineering Research*, 10, 167–171.
- Ríos, E. M. A., Sánchez, A. S., Lasheras, F. S., & Crespo, M. D. M. S. (2020). Genetic algorithm based on support vector machines for computer vision syndrome classification in health personnel. *Neural Computing and Applications*, 32(5), 1239–1248.
- Rocca, P., Oliveri, G., & Massa, A. (2011). Differential evolution as applied to electromagnetics. *IEEE Antennas and Propagation Magazine*, 53(1), 38–49.
- Ru, Y., Li, J., Chen, H., & Li, J. (2022). Epilepsy detection based on variational mode decomposition and improved sample entropy. *Computational Intelligence and Neuroscience*, 2022, 1–11.
- Sathish, E., Sivakumar, N., Simon, S. P., & Raghavan, S. (2017). Genetic algorithm based feature selection for classification of focal and non-focal intracranial electroencephalographic signals. *Journal of Scientific & Industrial Research*, 76, 614–619.
- Shen, S., Gong, W., & Cai, Z. (2016, July). Classification guided differential evolution. In *2016 IEEE congress on evolutionary computation (CEC)* (pp. 3276–3283). IEEE.
- Subasi, A., Kevric, J., & Canbaz, M. A. (2019). Epileptic seizure detection using hybrid machine learning methods. *Neural Computing and Applications*, 31(1), 317–325.
- Sufang, W. (2020). An adaptive ensemble classification framework for real-time data streams by distributed control systems. *Neural Computing & Applications*, 32(9), 4139–4149.
- Tao, H., Li, X., Paszke, W., Stojanovic, V., & Yang, H. (2021). Robust PD-type iterative learning control for discrete systems with multiple time-delays subjected to polytopic uncertainty and restricted frequency-domain. *Multidimensional Systems and Signal Processing*, 32(2), 671–692.
- Thakare, A., & Gore, S. (2019, September). A methodology for classification of seizure disorder using EEG Signals. In *2019 5th international conference on computing, communication, control and automation (ICCUBEA)* (pp. 1–7). IEEE.
- Thakare, A., & Kharche, D. (2017). Data clustering using hybrid swarm intelligence method with multi-objective functions. *International Journal of Control Theory and Applications*, 10(8), 177–183.
- Tzamourta, K. D., Tzallas, A. T., Giannakeas, N., Astrakas, L. G., Tsalikakis, D. G., Angelidis, P., & Tsipouras, M. G. (2019). A robust methodology for classification of epileptic seizures in EEG signals. *Health and Technology*, 9(2), 135–142.
- Wang, H., Wu, Z., & Rahnamayan, S. (2011). Enhanced opposition-based differential evolution for solving high-dimensional continuous optimization problems. *Soft Computing*, 15(11), 2127–2140.
- Wang, X., Gong, G., Li, N., & Qiu, S. (2019). Detection analysis of epileptic EEG using a novel random forest model combined with grid search optimization. *Frontiers in Human Neuroscience*, 13, 52.

- Xue, B., Fu, W., & Zhang, M. (2014, July). Differential evolution (DE) for multi-objective feature selection in classification. In *Proceedings of the companion publication of the 2014 annual conference on genetic and evolutionary computation* (pp. 83–84).
- Zhang, Y., Liu, B., Cai, J., & Zhang, S. (2017). Ensemble weighted extreme learning machine for imbalanced data classification based on differential evolution. *Neural Computing and Applications*, 28(1), 259–267.
- Zhao, S. Z., Suganthan, P. N., & Das, S. (2011). Self-adaptive differential evolution with multi-trajectory search for large-scale optimization. *Soft Computing*, 15(11), 2175–2185.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



Anuradha Thakare is a Professor in Department of Computer Engineering in Pimpri Chinchwad College of Engineering, Pune India. Anuradha received her Ph.D in Computer Science and Engineering from SGB Amravati University and M.E. degree in Computer Engineering from Savitribai Phule Pune University. She is serving for education and research from last 23 years. Her area of research is Evolutionary Computing, Artificial Intelligence, Machine Learning, Biomedical Engineering, Healthcare Analytics, High Performance Computing etc. She published 100+ research papers in reputed Journals and Conferences with indexing in Scopus, SCI, SCIE, Web of Science, ACM, Pubmed etc. She has authored and edited six books published by CRC Taylor & Francis, IGI Global, Wiley etc. She received Research grants from AICTE-AQIS, QIP-SPPU, BCUD-SPPU Pune, Maharashtra State Commission for Women and National Commission for Women. She worked as reviewer for Journal of International Blood Research, IEEE transactions and other Scopus indexed Journals. Anuradha is PhD guide in Computer Engineering in SPPU, Pune. She has been a General

Chair of IEEE International Conference ICCUBEA 2018 and Advisory member for International Conferences. Delivered 25+ expert talks on Machine Learning, Evolutionary Algorithms, Outcome Based Education etc. she worked with industries like DRDO, NCL etc for research projects. She is working as Subject Chairman for various Computer Engineering subjects under Savitribai Phule Pune University (SPPU). She contributed for SPPU syllabus Content designing and revision.



Ahmed M. Anter is an Associate Professor of Computer Science at the Computer Science and Information Technology (CSIT), Egypt-Japan University of Science and Technology (E-JUST), Alexandria, Egypt. Anter is also with the Computers and Artificial Intelligence, Beni-suef University, Egypt. Anter received M.Sc. and Ph.D. degrees in computer science from the Faculty of Computer Science and Information Systems, Mansoura University, in 2010 and 2016, respectively. From 2006 to 2010, he was a team leader of software development at CITC, Mansoura University, and from 2011 to 2014, he was a lecturer at the Faculty of Computer Science and Information Systems, Jazan University, Saudi Arabia. Also, Anter joined Shenzhen University as a post-doctoral fellow with the School of Biomedical Engineering, China, from 2018–2021. He has a good publication record with over 55 scientific research publications and serves as a reviewer for various international journals and conferences. Also, Anter serves as academic editor for prestigious journals. His main research interests include pattern recognition and intelligent systems, machine learning, medical image processing, meta-heuristics, optimization, neuroscience, and fuzzy systems.



Ajith Abraham received Ph.D. degree in Computer Science from Monash University, Melbourne, Australia (2001) and a Master of Science degree from Nanyang Technological University, Singapore (1998). Ajith's research experience includes over 33 years in the Industry and Academia. He has a worldwide academic experience. As an Investigator / Co-Investigator, he is part of research grants worth over 100 Million US\$. Ajith works in a multi-disciplinary environment involving machine intelligence, cyber-physical systems, Internet of things, network security, sensor networks, Web intelligence, data mining and applied to various real world problems. He has given more than 150 plenary lectures and conference tutorials in these areas. He has authored / co-authored over 1400+ publications (with colleagues from nearly 70 countries) and has an h-index of 105+! Some of his research works have also won best paper awards at International conferences. Currently he is the Dean of Faculty of Computing and Data Science at FLAME University, Pune India. Prior to that he was the Founding Director of Machine Intelligence Research Labs (MIR Labs - [http://](http://www.mirlabs.org)

www.mirlabs.org), which has members from 100+ countries. He was the Chair of IEEE Systems Man and Cybernetics Society Technical Committee on Soft Computing (2008-2021) and a Distinguished Lecturer of IEEE Computer Society representing Europe (2011-2013). He serves/served the editorial board of over 50 International journals.