



# Chaotic marine predator optimization algorithm for feature selection in schizophrenia classification using EEG signals

Zeynep Garip<sup>1</sup> · Ekin Ekinci<sup>1</sup> · Kasım Serbest<sup>2</sup> · Süleyman Eken<sup>3</sup>

Received: 13 January 2024 / Revised: 27 March 2024 / Accepted: 5 April 2024  
© The Author(s) 2024

## Abstract

Schizophrenia is a chronic mental illness that can negatively affect emotions, thoughts, social interaction, motor behavior, attention, and perception. Early diagnosis is still challenging and is based on the disease's symptoms. However, electroencephalography (EEG) signals yield incredibly detailed information about the activities and functions of the brain. In this study, a hybrid algorithm approach is proposed to improve the search performance of the marine predator algorithm (MPA) based on chaotic maps. For evaluating the performance of the proposed chaotic-based marine predator algorithm (CMPA), benchmark datasets are used. The results of the suggested variation method on the benchmarks show that the Sine Chaotic-based MPA (SCMPA) significantly outperforms the other MPA variants. The algorithm was verified using a public dataset consisting of 14 subjects. Moreover, the proposed SCMPA is essential for EEG electrode selection because it minimizes model complexity and selects the best representative features for providing optimal solutions. The extracted features for each subject were used in the decision tree (DT), random forest (RF), and extra tree (ET) methods. Performance measures showed that the proposed model was successful at differentiating schizophrenia patients (SZ) from healthy controls (HC). In the end, it was demonstrated that the feature selection technique SCMPA, which is the subject of this research, performs significantly better in regard to classification using EEG signals.

**Keywords** Feature selection · Metaheuristic algorithms · Chaotic maps · EEG signals · Classification

## 1 Introduction

Electroencephalography (EEG) is a method used to measure brain electrical activity [1]. A series of electrodes are usually placed on the scalp to measure electrical activity in different areas of the brain. EEG data represent brain waves. These waves have different frequencies and are associated with different brain activities. For example, beta waves are usually associated with alertness and attention, while alpha waves usually occur at rest and with eyes

closed. Delta and theta waves have slower frequencies and are associated with states such as deeper sleep or meditation [2, 3].

EEG is a technique used in many fields, such as neurology, neurophysiology, psychiatry, and sleep medicine. Some common clinical applications of EEG include diagnosing epilepsy [4], detecting sleep disorders [5], evaluating brain function [6], evaluating developmental disabilities in children [7], and identifying psychiatric disorders [8]. One of these psychiatric disorders is schizophrenia. Schizophrenia is associated with abnormalities in brain function, and EEG is used to detect these abnormalities. When EEG recordings of schizophrenia patients are analyzed, it has been observed that beta activity decreases, gamma activity changes, and P300 responses change [9–11].

Recently, artificial intelligence techniques have been used in analyzing EEG data in most fields, as shown in Table 1. Especially by analyzing EEG data using machine learning (ML) algorithms, schizophrenia can be detected in

✉ Ekin Ekinci  
ekinekinci@subu.edu.tr

<sup>1</sup> Faculty of Technology, Computer Engineering Department, Sakarya University of Applied Sciences, Sakarya, Turkey

<sup>2</sup> Faculty of Technology, Mechatronics Engineering Department, Sakarya University of Applied Sciences, Sakarya, Turkey

<sup>3</sup> Faculty of Technology, Information Systems Engineering Department, Kocaeli University, Kocaeli, Turkey

**Table 1** Summary of related work on EEG signals

Ref	Algorithms	Description	Features	Performance criteria (Acc)
[24]	CNN + logistic regression (LR)	They used a fusion of CNN and different ML classifiers	3	Nonsubject based: 98.05% Subject based:90%
[25]	SLBP + HLV + AdaBoostM1 (RF) + CBFS 13	They proposed a local descriptors approach for automated SZ detection using multichannel EEG signals	30 fet 13 fet	99.15% 99.36%
[26]	Decision Trees Random Forest Logistic Regression	They studied the three different types of time frequency feature extraction methods, and ML models for classifying SZ from HC	19	97.98%
[27]	Vision Transformer (LeViT)	They proposed a novel EEG data mapping method with Vision Transformer as feature extractor and classifier for early identification of SZ	19	98.99%
[28]	Boosted trees classifier	They proposed an accurate and easy-to-implement system to detect SZ using EEG signals	19	98.62%
[29]	VGG16	They used a CAD method to automatically distinguish between SZ and HC	19	99.5%
[13]	ResNet-18-SVM	They proposed a transfer learning with CNNs for the diagnosis of SZ from HC	19	98.60%
[30]	Collatz pattern	They proposed a novel Collatz conjecture-based automated SZ detection model using EEG signals	19	99.47%
[31]	VM, K-NN, LDA, DT	They proposed a method for the classification of SZ and HC EEG signals	19	98.9%
[32]	EfficientNetB0-LSTM	They developed CNN-LSTM models were exploited for recognition of SZ from normal subjects	19	99.90%
[33]	kNN	They developed an intelligent algorithm for diagnosing schizophrenia based on brain signal dynamics	19	93.75%

the early stages of the disease. In one of the previous studies, Oh et al. (2019) detected schizophrenia from EEG signals using an eleven-layered convolutional neural network (CNN), a deep learning method. The classification accuracy of the model they developed ranges between 81 and 98% [12]. In another study, Shalhaf et al. (2020) performed schizophrenia detection by converting EEG data into images. Here, they applied images of EEG signals to pretrained AlexNet, ResNet-18, VGG-19, and Inception-v3 algorithms. They evaluated the schizophrenia status of 28 participants using a support vector machine (SVM) classifier. As a result, they achieved 98% classification success with the ResNet-18-SVM algorithm [13]. Supakar et al. (2022) proposed a deep learning-based recurrent neural network (RNN)-long short-term memory (LSTM) model for schizophrenia detection. They classified EEG data obtained from 84 people, including schizophrenia patients and healthy people, with 93–98% accuracy [14]. Sun et al. (2021) developed a hybrid deep learning model by converting time-series EEG data into red–green–blue (RGB) images. As a result, fuzzy entropy features were found to

be more successful than fast Fourier transform features in schizophrenia classification [15].

One of the challenges in analyzing EEG data is working with high-dimensional datasets. Processing data recorded from a large number of electrodes at high sampling frequencies requires high computational power. In this context, it is necessary to extract important features from EEG signals and reduce the sample size [16].

Feature selection is a preprocessing phase that reduces the amount of data and computational complexity by removing unnecessary and redundant features, thereby improving the performance of ML algorithms. Nonetheless, identifying the ideal subset of features in high-dimensional datasets is considered an NP-hard problem. The search space will grow exponentially with an increase in features because a dataset with  $N$  features comprises  $2^{N-1}$  feature subsets. Because the exact methods cannot yield the necessary result in a fair amount of time, metaheuristic algorithms are employed to select the subset of features. Different metaheuristic algorithms are used for feature extraction in ML algorithms [17, 18]. Thus, the data size

can be reduced, and a simple structure can be obtained. In addition, owing to metaheuristic algorithms, overfitting can be prevented, and models that generalize better can be created. The most meaningful features in classifying schizophrenia can be determined and useful in developing practical clinical applications. For example, in a study on classifying schizophrenia, Prabhakar et al. (2020) optimized features in EEG data via the Flower Pollination and Eagle strategies using different evolution algorithms, a backtracking search optimization algorithm (BSA) and a group search optimization algorithm (GSO) [19]. The schizophrenia classification accuracy of the optimization algorithms they proposed varied between 82 and 90%. Khare and Bajaj (2022) used a whale optimization algorithm (WOA) for feature extraction from EEG data. They achieved 92% accuracy in classifying schizophrenia with the six features they identified [20].

In a population-based metaheuristic, the search process consists of two main phases: exploration and exploitation. However, while some metaheuristic algorithms require improvement during the exploitation stage, others must be enhanced during the discovery stage. It is also necessary to enhance both steps in a restricted set of algorithms. During the exploration phase, it is beneficial to behave randomly to cover as much ground as possible in the search space. In contrast, the primary goal of the latter phase is to quickly utilize the locations that show promise. Finding the right balance between these two phases is extremely difficult since population-based metaheuristic algorithms are stochastic. Discrete-time dynamical systems are also referred to as chaotic maps. To effectively find an optimal solution, chaos is incorporated into metaheuristic algorithms to create a balance between exploration and exploitation [21]. Consequently, optimization approaches obtain ergodic and nonrepeating properties of chaos. As a result, it can search more quickly than random search and avoid entering local optima. The performance of optimization algorithms can be greatly improved by all of these advantages [22].

In this study, a new hybrid approach is developed by updating the parameters of the MPA to increase the performance of the algorithm in finding the optimum global solution with a random number sequence obtained from five chaotic maps. Chaotic maps are logistic, tent, henon, sine, and tinkerbell maps. These approaches are known as Chaotic-based Marine Predators Algorithms (Henon Chaotic-based MPA [HCMMPA], Tinkerbell Chaotic-based MPA [TICMPA], Logistic Chaotic-based MPA [LCMPA], Tent Chaotic-based MPA [TECMPA] and Sine Chaotic-based MPA [SCMPA]).

The proposed hybrid approach tries to maximize the accuracy rate while minimizing the number of selected features. The MPA is used to determine significant features

by determining the best foraging strategy for predators and prey in marine environments. The major goal of using a chaotic method is to overcome the drawbacks of an MPA, such as local optimal traps and premature convergence, and ultimately increase the capacity of the search for exploration and exploitation. To enhance the FS performance of the MPA, the proposed technique employs one-dimensional chaotic maps as random number generators. In addition, the decision tree algorithm was chosen to determine the effect of the selected feature on classification. Decision tree construction can handle high-dimensional data and does not require subject expertise, making it ideal for exploratory knowledge mining [23]. For this reason, decision trees (DTs) and an ensemble of DTs are used in the study.

The contributions of the proposed methods can be listed as follows:

- We propose hybrid metaheuristic algorithms by combining MPA and five chaotic maps for feature selection in schizophrenia decision tree-based classification using EEG signals.
- To demonstrate the effectiveness of the SCMPA, the SCMPA is statistically compared with chaotic-based MPA variants on the well-known UCI (Breast, Hepatitis, Liver, Raisin and Heart) (<https://archive.ics.uci.edu/ml>) datasets.
- The ability of the SCMPA to perform feature selection in SCZ classification was verified using EEG signals.

The paper is organized as follows. In Sect. 2, the basic and proposed algorithms are defined and mathematicalized. Section 3 presents the experimental setup, evaluation metrics, details about the EEG signal data and preprocessing and the experimental study. Finally, in Sect. 4, the conclusions are given.

## 2 Methods

### 2.1 Marine predator algorithm (MPA)

The marine predator algorithm (MPA) was developed by Faramarzi et al., inspired by the prey-predator social relationship between marine predators and their prey [34]. Based on the MPA, the transition between phases in the structure of the algorithm is achieved according to the speed ratio between the prey and the predator [35]. These phases include (1) a high-velocity ratio or when prey is moving faster than a predator, (2) a unit-velocity ratio or when both predator and prey are moving at almost the same pace, and (3) a low-velocity ratio when the predator is moving faster than prey.

In the MPA, the initial solution is determined by using a randomly and uniformly distributed search space. The number of predators is  $n$ , the number of iterations is  $m$ , the size of the optimization parameter is  $d$ , and  $Prey$  represents the initial position of the prey.  $X_{max}$  and  $X_{min}$  in Eq. 1 are the maximum and minimum values, respectively, and  $rand$  is a random vector in the range  $[0, 1]$ . In this section, the  $Prey$  matrix, which holds the positions of the initial population, forms the  $Elit$  matrix with the best fitness function.

$$X_0 = X_{min} + rand(X_{max} - X_{min}) \quad (1)$$

Phase 1: In this behavior optimization at the first stage, during the first third of the iterations ( $iter < \frac{1}{3}maxiter$ ) at a high speed rate ( $v \geq 10$ ), the best strategy for the predators is not to move at all.  $P = 0.5$  in Eq. 2 and 3,  $R$  is defined as a vector containing uniformly distributed random numbers between  $[0, 1]$ , and  $R_B$  is defined as a vector containing random numbers based on the normal distribution of Brownian motion. With Eq. 3, the matrices used by  $Prey$ , which move according to Brownian motion, are updated.

$$\overrightarrow{stepsize}_i = \overrightarrow{R}_B \otimes \left( \overrightarrow{Elit}_i - \left( \overrightarrow{R}_B \overrightarrow{Prey}_i \right) \right) i = 1, 2, 3, \dots, n \quad (2)$$

$$\overrightarrow{Prey}_i = \overrightarrow{Prey}_i + \left( P \cdot \overrightarrow{R} \otimes \overrightarrow{stepsize}_i \right) \quad (3)$$

Phase 2: At this stage, the prey and the predators use different movement methods and move at the same speed during the second third of the iterations of the optimization ( $\frac{1}{3}maxiter < iter < \frac{2}{3}maxiter$ ).

While a predator uses Brownian motion,  $Prey$  uses Lévy motion. According to Eq. 4–5, the movements of the first half of the population are updated.  $R_L$  Levy's motion is a vector containing random numbers based on a normal distribution.

$$\overrightarrow{stepsize}_i = \overrightarrow{R}_L \otimes \left( \overrightarrow{Elit}_i - \left( \overrightarrow{R}_L \overrightarrow{Prey}_i \right) \right) \quad (4)$$

$$\overrightarrow{Prey}_i = \overrightarrow{Prey}_i + \left( P \cdot \overrightarrow{R} \otimes \overrightarrow{stepsize}_i \right) \quad (5)$$

According to Eqs. 6–7, the other half of the population is updated. Here,  $CF$  is an adaptive parameter for controlling the step size for predator movement.

$$\overrightarrow{stepsize}_i = \overrightarrow{R}_B \otimes \left( \overrightarrow{R}_B \otimes \overrightarrow{Elit}_i - \overrightarrow{Prey}_i \right) \quad (6)$$

$$\overrightarrow{Prey}_i = \overrightarrow{Elit}_i + \left( P \cdot CF \otimes \overrightarrow{stepsize}_i \right) \quad (7)$$

$$CF = [1 - (Iter./Max.Iter)]^{(2 \cdot Iter./Max.Iter.)}$$

Phase 3: The prey is assumed to move slower than the hunter during the remaining part of the optimization iteration number ( $iter > \frac{2}{3}maxiter$ ). For a low velocity ratio

( $v = 0.1$ ), the best strategy for predators is Lévy. The final phase is modeled according to Eqs. 8–9.

$$\overrightarrow{stepsize}_i = \overrightarrow{R}_L \otimes \left( \overrightarrow{R}_L \otimes \overrightarrow{Elit}_i - \overrightarrow{Prey}_i \right) \quad (8)$$

$$\overrightarrow{Prey}_i = \overrightarrow{Elit}_i + \left( P \cdot CF \otimes \overrightarrow{stepsize}_i \right) \quad (9)$$

Additionally, both eddy formation and fish aggregating devices (FADs) have direct impacts on the algorithm. According to the study of Houssein et al. (2021), sharks spend more than 80% of their time near FADs, and for the remaining 20%, they will make a larger jump in various dimensions, likely in search of a setting with a different distribution of prey [36]. The FADs are considered to be local optima; therefore, being aware of the lengthy leaps prevents them from becoming stranded in local optima. and their influence is assumed to be trapped in particular locations in the search space. It is modeled analytically according to Eq. 10.

$$\overrightarrow{Prey}_i = \begin{cases} \overrightarrow{Prey}_i + CF \left[ \overrightarrow{X}_{min} + \overrightarrow{R} \otimes \left( \overrightarrow{X}_{max} - \overrightarrow{X}_{min} \right) \right] \otimes \overrightarrow{U} & \text{if } (r \leq FADs) \\ \overrightarrow{Prey}_i + [FADs(1-r) + r] \left( \overrightarrow{Prey}_{r1} - \overrightarrow{Prey}_{r2} \right) & \text{if } (r > FADs) \end{cases} \quad (10)$$

## 2.2 Proposed chaotic-based MPA

In this study, a hybrid approach was developed by combining different chaos maps with the MPA algorithm. In this proposed algorithm, random number sequences of chaotic maps, which are the main parameters that affect the performance of the MPA, are used. Periodicity, stochastic natural features, remarkable execution, and sensitivity to the initial conditions are what make chaotic maps unique [37, 38].

The eddy or the effects of fish aggregating device (FAD) behavior on marine predators are highly important in the MPA. FADs and long skips, which are among the basic components of the MPA, prevent the algorithm from reaching a local optimum and increase its performance in finding the optimum global solution.

For these reasons, to improve the performance of the MPA, the value of the uniform random coefficient  $r$  given in Eq. 11, which provides the selection of eddies or the effects of FADs, is updated with random number sequences obtained from one- or two-dimensional chaotic maps. The updated mathematical model is given in Eq. 12.

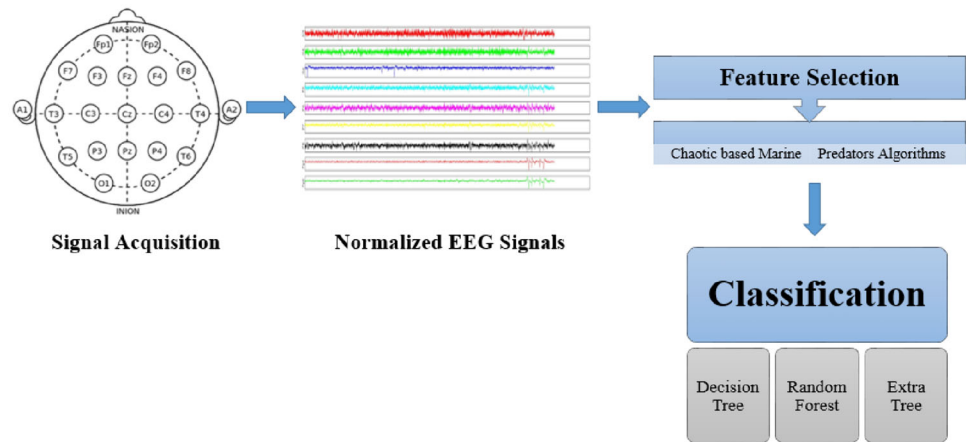
$$r = Chaotic_{(index)} \quad (11)$$

$$\overrightarrow{Prey}_i = \begin{cases} \overrightarrow{Prey}_i + CF \left[ \overrightarrow{X}_{min} + \overrightarrow{R} \otimes \left( \overrightarrow{X}_{max} - \overrightarrow{X}_{min} \right) \right] \otimes \overrightarrow{U} & \text{if } (r \leq FADs) \\ \overrightarrow{Prey}_i + [FADs(1-r) + r] \left( \overrightarrow{Prey}_{r1} - \overrightarrow{Prey}_{r2} \right) & \text{if } (r > FADs) \end{cases} \quad (12)$$

**Table 2** Behavior of chaotic maps

Name	Mathematical Equations	Map
Logistic	$X_{n+1} = \alpha X_n(1 - X_n)$ $\alpha = 3.9$	
Tent	$x_{i+1} = f(x_i, \mu)$ $f(x_i, \mu) = \begin{cases} f_L(x_i, \mu) = \mu x_i, & x_i < 0.5 \\ f_R(x_i, \mu) = \mu(1 - x_i), & x_i \geq 0.5 \end{cases}$ $\mu = 1.41$	
Henon	$x_{n+1} = 1 - \alpha x_n^2 + b y_n$ $y_{n+1} = x_n$ $\alpha = 1.4, b = 0.3$	
Sine	$x_{n+1} = r \sin(\pi x_n)$ $r = 0.867$	
Tinkerbell	$x_{n+1} = x_n^2 - y_n^2 + a x_n + b y_n$ $y_{n+1} = 2x_n y_n + c x_n + d y_n$ $a = 0.9, b = -0.6013,$ $c = 2.0, d = 0.50$	

**Fig. 1** The proposed system architecture



Five chaotic maps are employed in our experiments: logistic, tent, henon, sine, and tinkerbell maps. The mathematical equations and visualized maps are provided in Table 2.

## 2.3 Machine learning

### 2.3.1 Decision tree (DT)

For problems involving classification and regression (also known as supervised learning), DTs are a very helpful tool. The tree is made up of a root node that branches out into multiple decision nodes based on the characteristics of the dataset, signifying the various “questions” that the tree will try to answer and leaf nodes that signify the decision’s result. Since they explicitly match parameters to guide information flow, they train quickly, produce deterministic output (as opposed to the probabilistic output of a neural network), and perform well even with small datasets [39]. To improve performance as a group, they might be coupled in parallel (bagging) or in sequence (boosting). The inner trees of a random forest are fitted to randomly selected subsets of the data using bagging, and the majority decision is used as the output.

### 2.3.2 Random forest (RF)

RF is a well-liked and adaptable ensemble technique that involves a wide range of ML tasks and data types, and it is known for its effectiveness in many real-world applications [40]. Building an ensemble, or forest, of decision trees derived from a randomized version of the tree induction procedure is the basis of the RF family of techniques. Given that decision trees typically have high variation and low bias, which increases their likelihood of benefiting from the averaging process, they are excellent candidates for ensemble approaches. A bootstrapped sample of the

training data was used to train each tree. Therefore, this approach creates diversity among the trees [41].

### 2.3.3 Extremely randomized trees (extra trees-ET)

ET is an ensemble ML technique that shares many similarities with random forests. Its main applications include classification and regression in data mining, image and text analysis, and various ML tasks, and it has become popular because it is highly effective and efficient for a variety of data types [42]. ET creates a collection of decision trees, similar to random forests. They do, however, go beyond the idea of randomness. In ET, a random subset of features is chosen for each node, and the best split is selected from those random features, as opposed to choosing the best split for each node individually. This additional unpredictability can result in more robust models by preventing overfitting. Bootstrapping, which creates several subsamples of the training data with replacement, is a technique commonly used by ETs [43]. One of these bootstrapped datasets serves as the training set for each decision tree in the ensemble. Because the feature selection and bootstrapping steps are randomized, ETs are capable of handling noisy data with great effectiveness [44].

## 3 Experimental study

In this study, we conducted two different experiments. In the first experiment, we evaluate which of five different chaotic map-based MPA algorithms performs well in feature selection on benchmark datasets. We then choose sine, the chaotic map that performs rather well. In the second experiment, classifications are made with DT-based algorithms on EEG signal datasets comprising 14 subjects and datasets obtained as a result of the features selected with the SCMPA. The aim is to determine whether the proposed SCMPA can extract the features that best represent the



dataset. The experimental setup, evaluation criteria, benchmark datasets, feature selection and classification experiments are described below. The proposed system architecture is shown in Fig. 1.

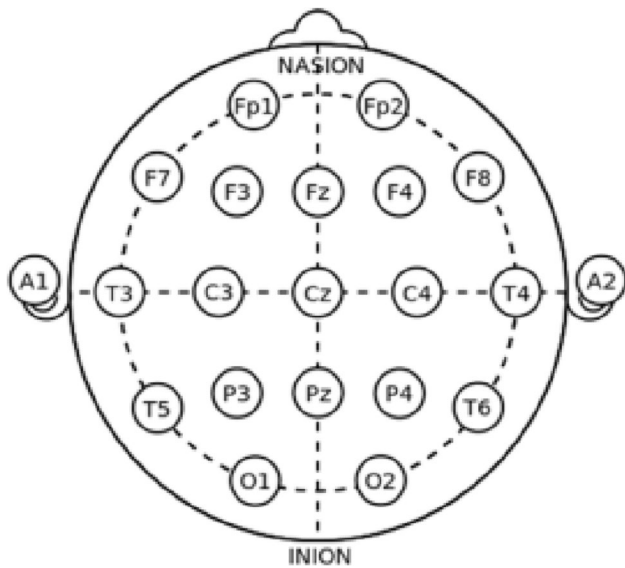


Fig. 2 The standard 10–20 system of EEG electrode maps

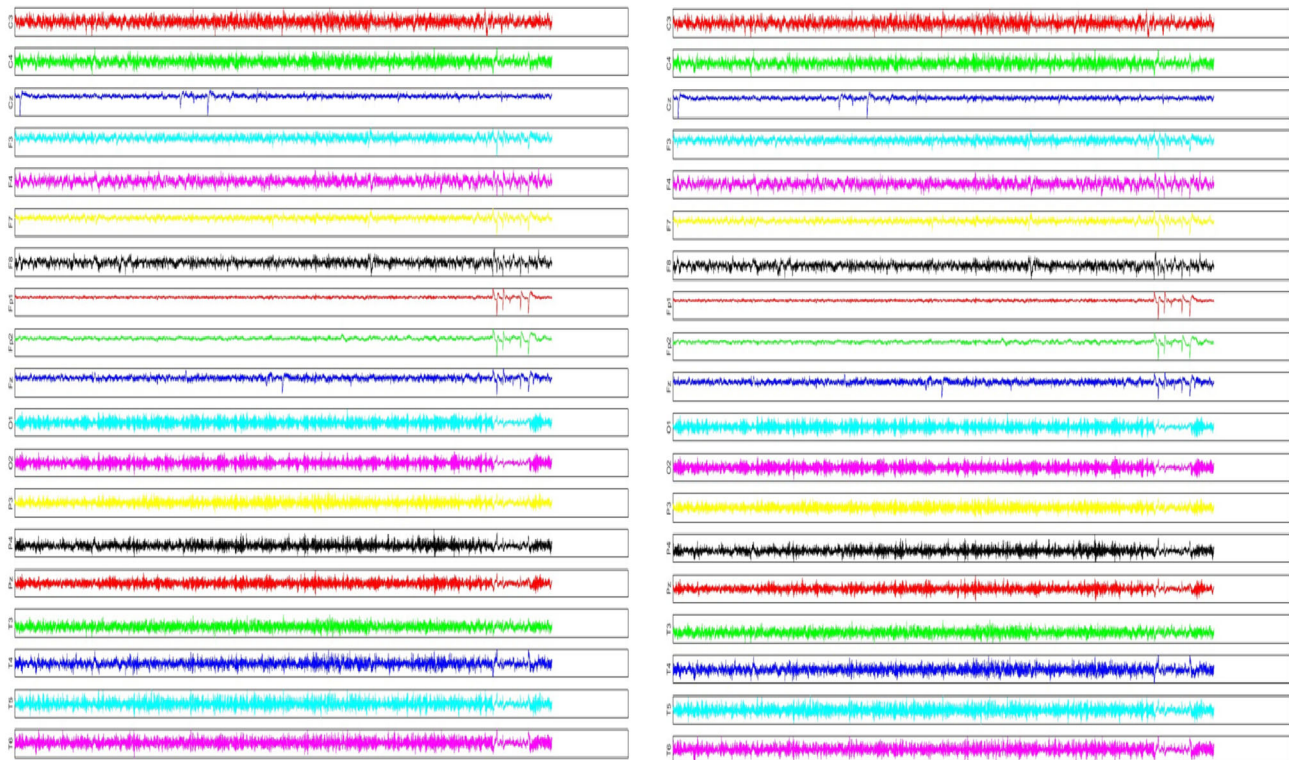


Fig. 3 EEG signal samples of (left column) HCs and (right column) patients with SZ

The proposed system architecture is shown in Fig. 1.

### 3.1 Experimental setup

To verify the feasibility and effectiveness of the proposed SCMPA, experiments employing a dataset of 14 subjects and four different classification algorithms, namely, DT, RF, and ET, are carried out. The classification performances obtained without using any feature selection method (only the original feature set) and obtained through the features selected with SCMPA are also compared. MATLAB 2023b was used to conduct all the numerical experimental studies. SCMPA is run 20 times with a population of 30 and 60 iterations.

To realize ML algorithms, the needed parameters are determined. The main parameters of the algorithms are as follows: the Gini index is utilized as a splitting criterion for DTs to determine which splitting characteristic is optimal for each node. Since the maximum depth of the tree for DT is set to “none”, nodes extend until every leaf is cut. At least two samples are needed for the internal node algorithm to be split. In addition to the parameter settings in the

**Table 3** List of datasets employed in the experiments

Datasets	Number of features	Number of instances	Number of classes	Data category	Missing values	Attribute characteristics
Breast	9	699	2	Biology	Yes	Integer
Hepatitis	19	154	2	Biology	Yes	Categorical, Real, Integer
Liver	5	345	2	Biology	N/A	Categorical, Integer, Real
Raisin	7	900	2	Agriculture	N/A	Real, Integer
Heart	13	270	2	Biology	N/A	Categorical, Real

**Table 4** Statistical results of fitness values evaluated by SCMPA and MPA variants on benchmark datasets

Dataset	Performance Criteria	Algorithms					
		MPA	HCMPA	TICMPA	LCMPA	TECMPA	SCMPA
Breast	Min	0.0056	0.0093	0.0089	<b>0.0033</b>	0.0105	0.0176
	Avg	0.0197	0.0218	<b>0.0173</b>	0.0267	0.0236	0.0245
	Std	0.0101	0.0088	0.0071	0.0129	0.0086	<b>0.0059</b>
Hepatitis	Min	0.1032	0.1341	0.1032	0.1016	<b>0.1011</b>	0.1037
	Avg	<b>1.6758</b>	1.8088	1.7104	1.8759	1.8424	1.7423
	Std	0.0535	<b>0.0321</b>	0.0449	0.0519	0.0450	0.0549
Liver	Min	0.2236	0.2236	<b>0.1915</b>	0.2539	0.2109	0.1949
	Avg	2.8849	2.8975	2.7811	3.0046	<b>2.7133</b>	2.7878
	Std	0.0365	0.0281	0.0390	0.0308	<b>0.0274</b>	0.0406
Raisin	Min	0.0999	<b>0.0964</b>	0.0978	0.1074	0.1169	0.1239
	Avg	1.3086	1.3622	1.3459	<b>1.2936</b>	1.4461	1.4378
	Std	0.0217	0.0176	0.0161	0.0195	0.0182	<b>0.0141</b>
Heart	Min	0.0756	0.0772	0.0588	0.0779	<b>0.0222</b>	0.0282
	Avg	0.1073	0.1189	0.1209	0.1313	0.1063	<b>0.1023</b>
	Std	<b>0.0245</b>	0.0282	0.0373	0.0285	0.0345	0.0520

The best average fitness values are represented as bold

DT, the number of estimators in the ET and RF are set to 100.

### 3.2 Evaluation criteria

The F score with tenfold cross-validation (CV) was used to evaluate the performance of the proposed framework for each subject. The F score is the harmonic mean of the precision and recall. where true positive (TP) is the number of correctly classified healthy signals and false positive (FP) is the number of schizophrenia signals classified as healthy. The number of false negatives (FNs) represents the number of healthy signals classified as schizophrenia, and the number of true negatives (TNs) represents the number of correctly classified schizophrenia signals. In terms of TP, FP, FN, and TN, the formulas for the performance metrics are given below.

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (15)$$

$$F - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (16)$$

In addition to the F score, the performance of each model was assessed using the area under the receiver operating characteristic (ROC) curve (AUC). The overall AUC was used to evaluate the performance of the models across all possible categorization criteria. The range of the AUC value is 0 to 1. The model has a good ability to classify data if the value is near 1. The true positive rate (TPR) was plotted against the false positive rate (FPR) with the TPR on the y-axis and the FPR on the x-axis representing the ROC curve.



**Table 5** Statistical results of fitness values evaluated by SCMPA and MPA variants in the EEG dataset

Subjects	Performance Criteria	MPA	HCMPA	TICMPA	LCMPA	TECMPA	SCMPA
S1	Min	4,81E-02	<b>4,79E-02</b>	<b>4,60E-02</b>	<b>4,73E-02</b>	4,84E-02	<b>4,65E-02</b>
	Avg	5,04E-02	5,05E-02	4,91E-02	5,03E-02	5,09E-02	4,94E-02
	Std	1,86E-03	1,30E-03	2,11E-03	1,43E-03	1,62E-03	1,95E-03
S2	Min	9,39E-02	9,46E-02	<b>9,28E-02</b>	<b>9,38E-02</b>	<b>2,64E-03</b>	<b>2,06E-03</b>
	Avg	9,78E-02	9,74E-02	9,60E-02	9,80E-02	6,92E-02	8,86E-02
	Std	2,39E-03	1,26E-03	2,20E-03	2,07E-03	4,56E-02	3,04E-02
S3	Min	5,30E-02	<b>5,18E-02</b>	<b>5,21E-02</b>	<b>5,11E-02</b>	<b>5,09E-02</b>	<b>5,28E-02</b>
	Avg	5,46E-02	5,43E-02	5,48E-02	5,49E-02	5,51E-02	5,55E-02
	Std	1,10E-03	2,00E-03	1,73E-03	2,41E-03	2,43E-03	1,94E-03
S4	Min	1,56E-03	1,75E-03	8,71E-02	<b>1,09E-03</b>	8,47E-02	<b>1,47E-03</b>
	Avg	7,16E-02	7,33E-02	9,07E-02	6,42E-02	9,06E-02	5,36E-02
	Std	3,67E-02	3,69E-02	1,88E-03	4,32E-02	2,90E-03	4,48E-02
S5	Min	3,54E-03	1,06E-02	<b>2,03E-03</b>	5,30E-02	<b>1,28E-03</b>	<b>1,09E-03</b>
	Avg	4,94E-02	5,01E-02	4,88E-02	5,59E-02	4,41E-02	4,43E-02
	Std	1,62E-02	1,40E-02	1,65E-02	2,10E-03	2,21E-02	2,26E-02
S6	Min	8,49E-02	<b>8,39E-02</b>	8,59E-02	<b>8,44E-02</b>	8,61E-02	<b>8,45E-02</b>
	Avg	9,04E-02	8,92E-02	8,95E-02	9,04E-02	9,07E-02	8,98E-02
	Std	2,23E-03	2,43E-03	1,60E-03	3,46E-03	2,52E-03	2,51E-03
S7	Min	4,37E-03	5,29E-02	8,50E-03	5,41E-02	5,01E-02	<b>3,51E-03</b>
	Avg	5,09E-02	5,53E-02	5,01E-02	5,63E-02	5,60E-02	5,05E-02
	Std	1,65E-02	1,58E-03	1,46E-02	1,73E-03	2,77E-03	1,67E-02
S8	Min	7,94E-02	8,37E-02	7,99E-02	8,27E-02	<b>7,93E-02</b>	7,97E-02
	Avg	8,31E-02	8,52E-02	8,43E-02	8,47E-02	8,42E-02	8,39E-02
	Std	2,35E-03	1,48E-03	2,22E-03	2,05E-03	3,05E-03	2,94E-03
S9	Min	7,98E-02	8,05E-02	9,01E-03	<b>2,98E-03</b>	<b>1,75E-03</b>	<b>3,07E-03</b>
	Avg	8,37E-02	8,50E-02	7,57E-02	7,53E-02	7,54E-02	7,61E-02
	Std	2,34E-03	2,43E-03	2,36E-02	2,55E-02	2,59E-02	2,57E-02
S10	Min	1,66E-03	7,11E-02	6,98E-02	2,13E-03	7,14E-02	<b>1,56E-03</b>
	Avg	5,84E-02	7,34E-02	7,41E-02	6,67E-02	7,44E-02	6,72E-02
	Std	2,98E-02	1,60E-03	2,91E-03	2,27E-02	2,12E-03	2,32E-02
S11	Min	<b>2,60E-03</b>	5,13E-02	3,17E-03	2,69E-03	4,39E-03	4,01E-03
	Avg	3,77E-02	5,32E-02	4,65E-02	4,32E-02	4,74E-02	4,85E-02
	Std	2,34E-02	1,23E-03	1,53E-02	2,12E-02	1,52E-02	1,58E-02
S12	Min	8,11E-04	1,09E-03	9,05E-04	<b>6,21E-04</b>	9,05E-04	9,05E-04
	Avg	1,30E-03	1,40E-03	4,01E-03	1,24E-03	1,30E-03	1,27E-03
	Std	2,75E-04	2,75E-04	8,25E-03	2,94E-04	2,31E-04	3,89E-04
S13	Min	7,88E-02	7,96E-02	8,12E-02	7,91E-02	<b>2,13E-03</b>	<b>2,32E-03</b>
	Avg	8,16E-02	8,29E-02	8,34E-02	8,30E-02	6,69E-02	7,50E-02
	Std	2,08E-03	2,31E-03	1,74E-03	1,96E-03	3,34E-02	2,57E-02
S14	Min	<b>3,35E-03</b>	4,20E-03	5,52E-03	4,01E-03	3,73E-03	4,01E-03
	Avg	2,81E-02	7,52E-02	7,56E-02	2,05E-02	2,02E-02	2,81E-02
	Std	3,80E-02	2,50E-02	2,47E-02	3,31E-02	3,23E-02	3,81E-02

For each subject, those with the best fitness values compared to MPA are in bold

**Table 6** Selected features with SCMPA

Subjects	Fitness Values	Selected Features
S1	4,65E-02	Fp1, Fp2, F7, F4, F8, T3, C3, Cz, T4, T5, P3, Pz, P4, T6, O1
S2	2,06E-03	C3
S3	5,28E-02	Fp1, Fp2, F7, F3, F4, T3, C4, T4, T5, P3, Pz, P4, T6, O1, O2
S4	1,47E-03	T3
S5	1,09E-03	T5
S6	8,45E-02	Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, O1, O2
S7	3,51E-03	O1
S8	7,97E-02	Fp1, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2
S9	3,07E-03	T5
S10	1,56E-03	P4
S11	4,01E-03	T5
S12	9,05E-04	T5
S13	2,32E-03	T5
S14	4,01E-03	C3

When evaluating the performance of a model using data that are not utilized during training, the CV is employed. The fundamental idea of CV is to eliminate some data, use the remaining data to create a model, and then estimate the samples that are omitted. For a total of  $k$  iterations, the data are divided into  $k$  folds, with the excluded samples serving as the test sample in each fold. The ultimate success of the  $k$ -fold CV is determined by averaging the  $k$  performances that are acquired. In this investigation, a  $10 \times CV$  was used.

### 3.3 Dataset

The publicly available EEG signal dataset from the Institute of Psychiatry and Neurology in Warsaw, Poland, was used for the experimental study [45]. The dataset included 14 patients (7 males:  $27.9 \pm 3.3$  years, 7 females:  $28.3 \pm 4.1$  years) with paranoid schizophrenia (SZ) and 14 healthy controls (HCs) (7 males:  $26.8 \pm 2.9$  years, 7 females:  $28.7 \pm 3.4$  years). The EEG signals were recorded for fifteen minutes while the patients were in an eyes-closed resting state. Using the conventional 10–20 EEG montage (shown in Fig. 2, where A1 and A2 are reference electrodes) with 19 EEG channels, Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2, data are collected at a sampling frequency of 250 Hz. However, every channel's EEG data contain 225,000 samples; the first 100 s of the signals are taken into account. This dataset is available at <http://dx.doi.org/https://doi.org/10.18150/repod.0107441>. The EEG recordings from HCs and SZ patients are shown in Fig. 3.

We must address significant differences in the raw EEG data since this frequently results in the ML algorithm underperforming when there is a significant difference in the numerical attributes of the dataset. In this study, we

normalize the raw EEG data and transform the data into the  $[0, 1]$  range by using min–max normalization.

$$x_{i_{scaled}} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (17)$$

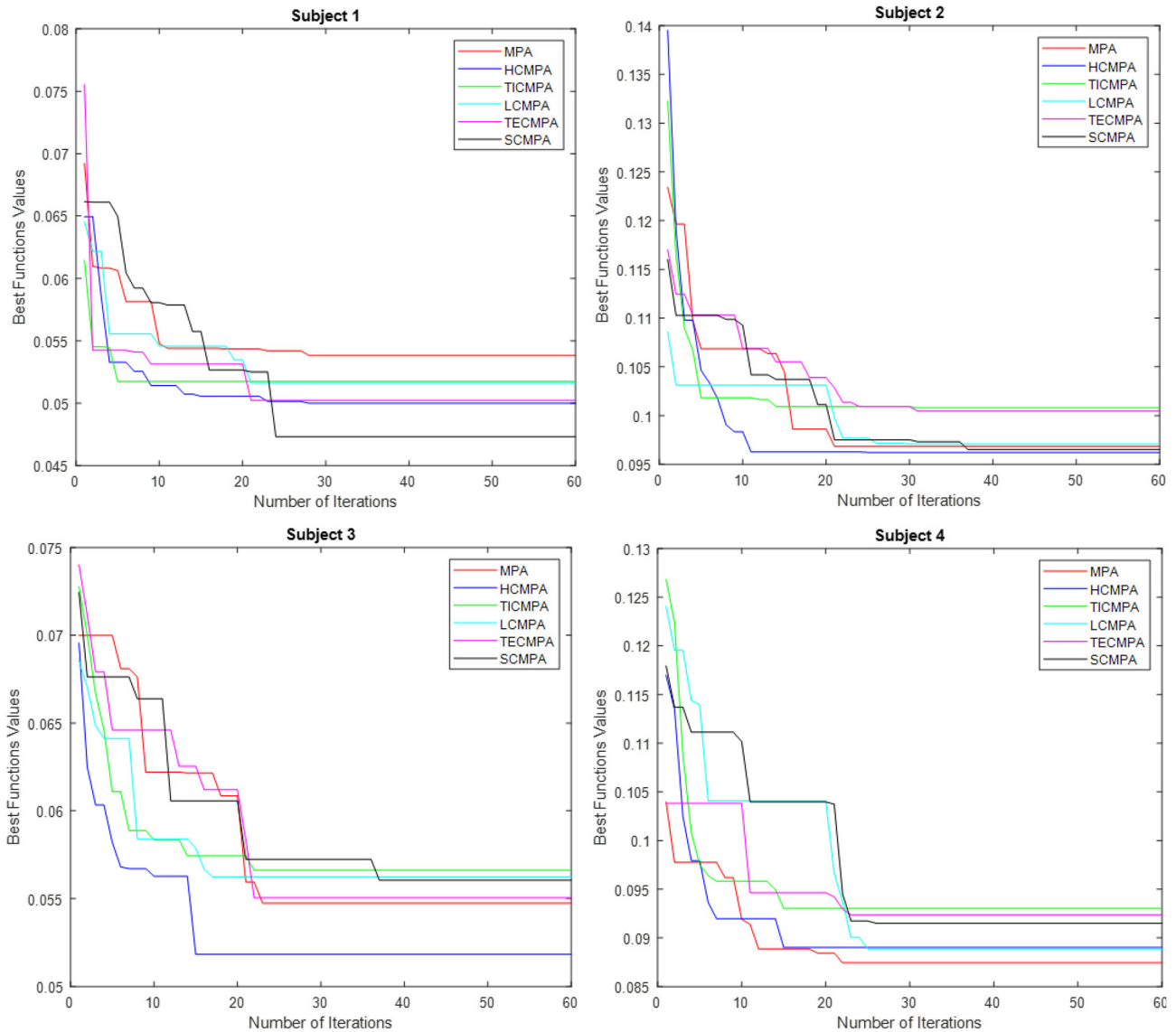
where  $x_i$  represents the  $i_{th}$  data point,  $x_{min}$  represents the minimum-valued data point,  $x_{max}$  represents the maximum-valued data point, and  $x_{i_{scaled}}$  is the normalized form of  $x_i$ .

### 3.4 The performance of the feature selection algorithms

#### 3.4.1 Case study-I: Comparison of the performance of the original MPA and chaotic map-based MPA algorithms on benchmark datasets

In this study, five datasets (<https://archive.ics.uci.edu/ml>) are selected to verify the effectiveness and efficiency of the algorithm developed for feature selection. The number of features, number of instances, number of classes, and category of data definitions for the selected datasets are given in Table 3. The main objective of these experiments is to evaluate the performance of the MPA on different chaotic maps and determine the optimal chaotic map. The number of populations and the maximum number of generations for each method are 30 and 60, respectively. The mean (mean), standard deviation (std), and minimum (best) values of the different evaluation indices after 20 runs for each dataset are shown in Table 4. Additionally, Table 4 compares chaotic-based MPAs with different chaotic maps with the original MPA in terms of feature selection.

Table 4 shows that the chaotic MPA outperforms the original MPA when various chaotic maps are used. Additionally, it should be noted that across all the datasets, the SCMPA and tent map achieved the best fitness values with



**Fig. 4** Convergence of each algorithm for each subject

the fewest selected feature results when compared to the others. Compared with the other maps, the breast has the best performance at 0.0033 using LCMPA. Hepatitis showed its best performance at 0.1011 with TECMPA. While Liver demonstrated superior performance at 0.1915 using the TICMPA when compared to other map options, Raisin revealed a top performance value of 0.0964 through the HCMPA. The heart exhibited the highest performance value at 0.0222 according to the TECMPA compared to the other maps. All these results prove that chaotic-based MPA algorithms reduce the number of selected features while increasing the quality compared to the MPA.

### 3.4.2 Case study II: Comparison of the performance of the original MPA algorithm and the chaotic map-based MPA algorithm on the EEG dataset

Metaheuristic algorithms were run on the EEG dataset, and the results obtained are given in Table 5. According to Table 5, the most successful algorithm is the SCMPA, and feature selection is performed with the SCMPA.

The features selected for each subject with SCMPA are given in Table 6. According to Table 6, approximately 21% of the features for subjects 1 and 3 were eliminated. Approximately 95% of the features are eliminated for Subjects S2, S4, S5, S7, S9, S10, S11, S12, S13 and S14. For Subjects S6 and S8, approximately 5% of the features

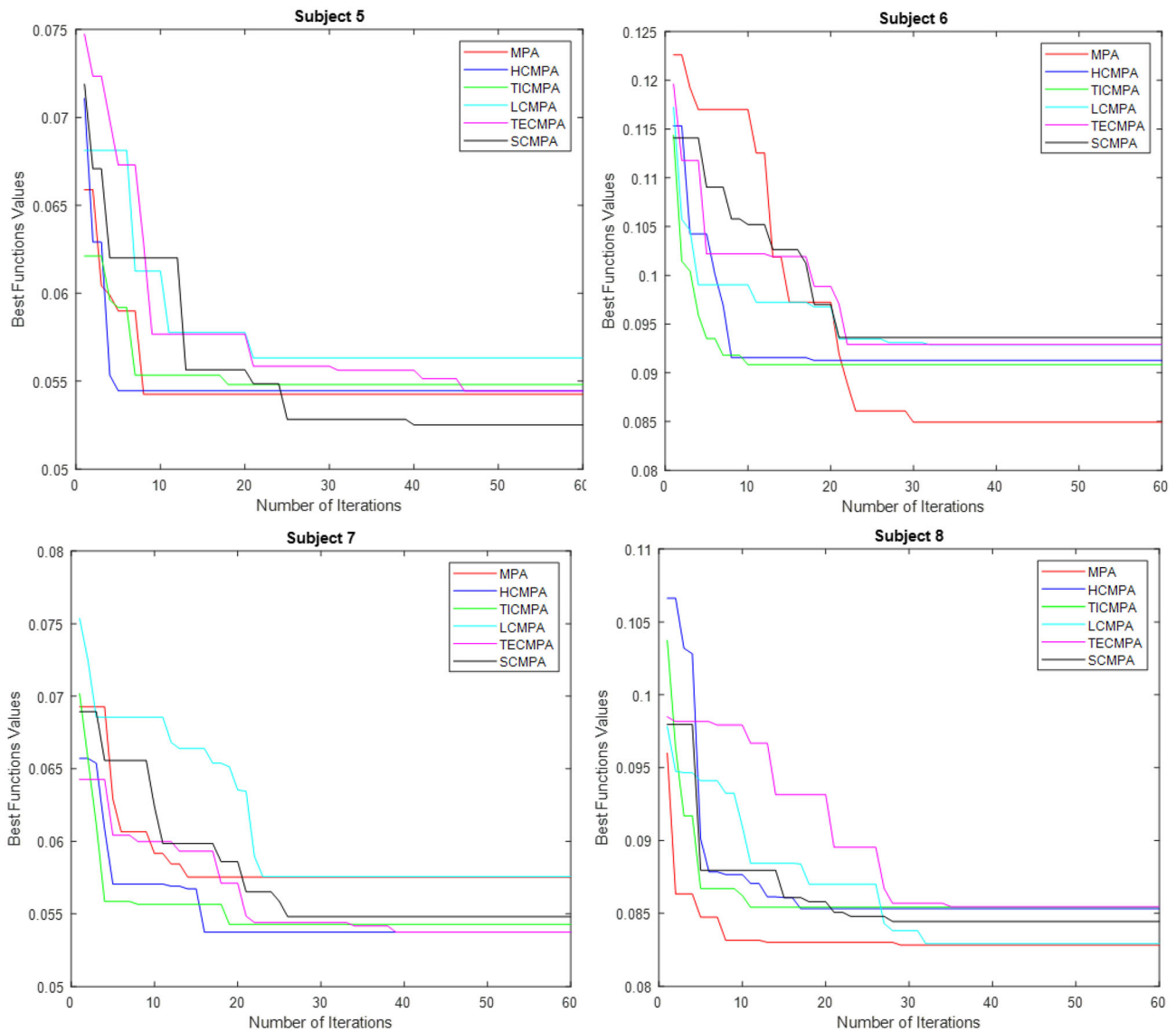


Fig. 4 continued

are eliminated. The convergence graph obtained from each subject of the SCMPA algorithm is also given in Fig. 4.

Additionally, the total rank of MPA and its variants was calculated using Friedman's mean rank test for the EEG dataset as shown in Table 7. Results show that most of the subjects SCMPA ranked first compared to other algorithms.

### 3.5 Classification experiments

In this section, to verify the effectiveness of the SCMPA in SZ recognition, we conduct two types of classification experiments: original feature set-based and selected feature set-based.

#### 3.5.1 Experiment-I: Original feature set-based EEG classification

Classification was performed with tenfold cross-validation via the DT, RF and ET algorithms on the original dataset for each subject, and the obtained performance values are given in Table 8.

An examination of the results reveals that the DT algorithm achieved relatively less successful classification than did the other two algorithms. However, when evaluated in general, it cannot be ignored that all the classification models applied were successful. This study aimed to determine whether the same or higher success rate can be achieved with a smaller feature set. Thus, instead of

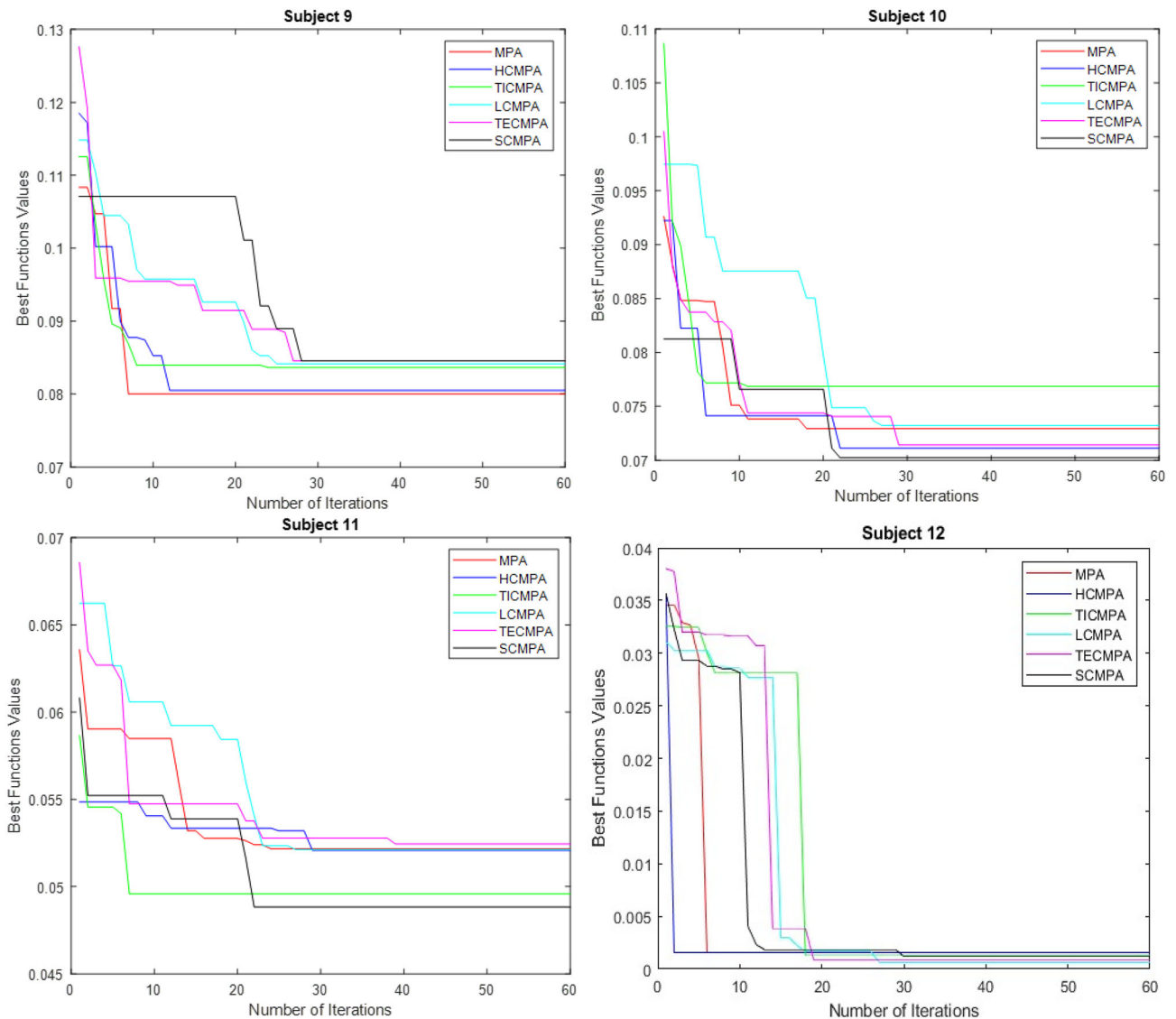


Fig. 4 continued

working with the entire feature set, one can focus on important features and work with a smaller feature set.

### 3.5.2 Experiment II: Selected feature set-based EEG classification

Determining which subset of features performs best in classification is crucial since reducing the number of dimensions in the data can result in reduced calculation time and information processing costs. In this section, classification is performed for each subject based on the features selected with the SCMPA, and the performance values are given in Table 9.

Based on Table 9 above, the results show that selected feature set-based EEG classification with a tenfold CV

gives more accurate results than does the original feature set-based EEG classification.

SZs and HCs were accurately classified according to the results displayed above. The confusion matrix and ROC curves, shown in Fig. 5, suggest that nearly 100% accuracy can be attained in the robust discrimination of HCs and SZ patients by using only one electrode.

The comparison between classification with original features and SCMPA selected features are given in Table 10. Based on the results, it can be clearly seen that the same or even better results can be achieved with fewer features.



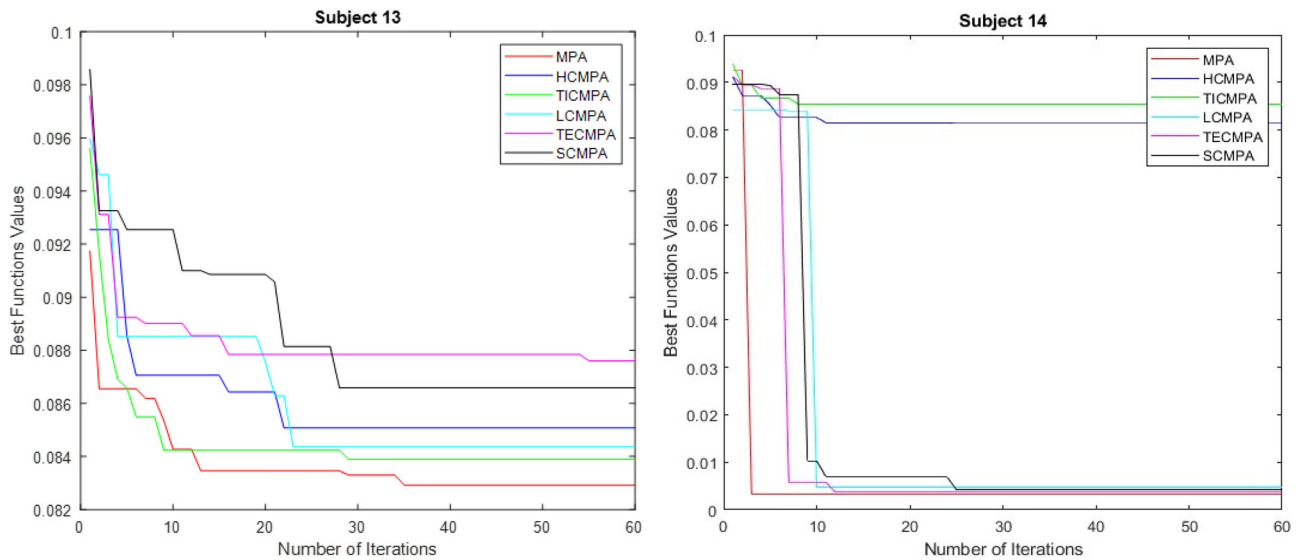


Fig. 4 continued

Table 7 Friedman's mean rank statistics

Subjects	Algorithm	MPA	HCMPA	TICMPA	LCMPA	TECMPA	SCMPA
S1	Friedman Mean Rank	3.3000	3.9000	2.7000	3.7000	4.2000	3.2000
	Rank	3	5	1	4	6	2
S2	Friedman Mean Rank	4.2000	3.6000	2.5000	4.2000	2.8000	3.7000
	Rank	5	3	1	6	2	4
S3	Friedman Mean Rank	3	2.9000	4	3.1000	3.8000	4.2000
	Rank	2	1	5	3	4	6
S4	Friedman Mean Rank	3.1000	3.9000	4.2500	3.2000	4.2500	2.3000
	Rank	2	4	5	3	6	1
S5	Friedman Mean Rank	3.4000	3.7000	3.2000	4.3000	3.2000	3.2000
	Rank	4	5	3	6	2	1
S6	Friedman Mean Rank	4.2000	3.0500	3.0500	3.8000	3.9000	3
	Rank	6	2	3	4	5	2
S7	Friedman Mean Rank	3.6000	3.4500	2.1000	4.3000	4.2500	3.3000
	Rank	4	3	1	6	5	2
S8	Friedman Mean Rank	2.5000	4.6500	3.3000	3.7500	3.6000	3.2000
	Rank	1	6	3	5	4	2
S9	Friedman Mean Rank	3.5000	4.2000	2.8000	4.2000	3.1500	3.1500
	Rank	4	5	1	6	3	2
S10	Friedman Mean Rank	2.1000	3.3500	4.0500	3.9000	3.9000	3.7000
	Rank	1	2	6	4	5	3
S11	Friedman Mean Rank	2.9000	4.5000	2.4000	3.8000	3.5000	3.9000
	Rank	2	6	1	4	3	5
S12	Friedman Mean Rank	3.4000	4.2500	4.1000	3.1500	3.5500	2.5500
	Rank	3	6	5	2	4	1
S13	Friedman Mean Rank	2.7000	3.3000	4.4000	4.1000	3.1000	3.4000
	Rank	1	3	6	5	2	4
S14	Friedman Mean Rank	2.9500	4.8000	4.9000	2.9000	2.6000	2.8500
	Rank	4	5	6	3	1	2

**Table 8** Performance of each classification algorithm for each subject (original features)

Subjects	Model	Accuracy	AUC	Recall	Precision	F1-Score
S1	DT	0.89	0.89	0.89	0.89	0.89
	RF	0.95	0.99	0.93	0.97	0.95
	ET	0.96	0.99	0.93	0.98	0.95
S2	DT	0.83	0.83	0.83	0.83	0.83
	RF	0.91	0.97	0.87	0.94	0.9
	ET	0.9	0.97	0.83	0.96	0.89
S3	DT	0.88	0.88	0.89	0.88	0.88
	RF	0.94	0.98	0.97	0.92	0.94
	ET	0.95	0.99	0.98	0.92	0.95
S4	DT	0.86	0.86	0.86	0.86	0.86
	RF	0.92	0.98	0.95	0.89	0.92
	ET	0.91	0.98	0.96	0.87	0.91
S5	DT	0.91	0.91	0.91	0.91	0.91
	RF	0.96	0.99	0.97	0.94	0.96
	ET	0.96	0.99	0.98	0.94	0.96
S6	DT	0.78	0.78	0.79	0.78	0.78
	RF	0.89	0.96	0.88	0.89	0.89
	ET	0.9	0.97	0.9	0.91	0.9
S7	DT	0.89	0.89	0.9	0.89	0.89
	RF	0.95	0.99	0.97	0.93	0.95
	ET	0.95	0.99	0.98	0.92	0.95
S8	DT	0.86	0.86	0.86	0.86	0.86
	RF	0.92	0.97	0.94	0.9	0.92
	ET	0.93	0.98	0.97	0.9	0.93
S9	DT	0.82	0.82	0.82	0.83	0.82
	RF	0.93	0.98	0.93	0.91	0.93
	ET	0.92	0.98	0.94	0.91	0.93
S10	DT	0.85	0.85	0.85	0.85	0.85
	RF	0.93	0.98	0.93	0.93	0.93
	ET	0.93	0.98	0.93	0.93	0.93
S11	DT	0.92	0.92	0.92	0.92	0.92
	RF	0.96	0.99	0.97	0.96	0.96
	ET	0.96	0.99	0.98	0.95	0.97
S12	DT	0.97	0.97	0.97	0.97	0.97
	RF	0.98	1	0.99	0.97	0.98
	ET	0.98	1	1	0.97	0.98
S13	DT	0.87	0.87	0.88	0.87	0.87
	RF	0.94	0.98	0.95	0.92	0.94
	ET	0.94	0.99	0.97	0.91	0.94
S14	DT	0.92	0.92	0.92	0.93	0.92
	RF	0.93	0.98	0.97	0.90	0.93
	ET	0.93	0.98	0.98	0.89	0.93

**Table 9** Performance of each classification algorithm for each subject (selected features)

Subjects	Model	Accuracy	AUC	Recall	Precision	F1-Score
S1	DT	0.89	0.89	0.89	0.89	0.89
	RF	0.95	0.99	0.93	0.97	0.95
	ET	0.95	0.99	0.92	0.98	0.95
S2	DT	0.99	0.99	0.99	0.99	0.99
	RF	0.99	0.99	0.99	0.99	0.99
	ET	0.99	0.99	0.99	0.99	0.99
S3	DT	0.87	0.87	0.87	0.88	0.88
	RF	0.93	0.98	0.97	0.91	0.94
	ET	0.94	0.99	0.98	0.92	0.95
S4	DT	0.99	0.99	0.99	1	0.99
	RF	0.99	0.99	0.99	1	0.99
	ET	0.99	0.99	0.99	1	0.99
S5	DT	0.99	0.99	0.99	1	0.99
	RF	0.99	0.99	0.99	1	0.99
	ET	0.99	0.99	0.99	1	0.99
S6	DT	0.78	0.78	0.78	0.78	0.78
	RF	0.89	0.96	0.88	0.89	0.89
	ET	0.9	0.97	0.89	0.91	0.9
S7	DT	0.99	0.99	0.99	0.99	0.99
	RF	0.99	0.99	0.99	0.99	0.99
	ET	0.99	0.99	0.99	0.99	0.99
S8	DT	0.85	0.85	0.85	0.86	0.85
	RF	0.92	0.97	0.94	0.9	0.92
	ET	0.93	0.97	0.96	0.9	0.93
S9	DT	0.99	0.99	0.99	0.99	0.99
	RF	0.99	0.99	0.99	0.99	0.99
	ET	0.99	0.99	0.99	0.99	0.99
S10	DT	0.99	0.99	0.99	0.99	0.99
	RF	0.99	0.99	0.99	0.99	0.99
	ET	0.99	0.99	0.99	0.99	0.99
S11	DT	0.99	0.99	0.99	1	0.99
	RF	0.99	0.99	0.99	1	0.99
	ET	0.99	0.99	0.99	1	0.99
S12	DT	0.99	0.99	0.99	1	0.99
	RF	0.99	0.99	0.99	1	0.99
	ET	0.99	0.99	0.99	1	0.99
S13	DT	0.99	0.99	0.99	1	0.99
	RF	0.99	0.99	0.99	1	0.99
	ET	0.99	0.99	0.99	1	0.99
S14	DT	0.99	0.99	0.99	0.99	0.99
	RF	0.99	0.99	0.99	0.99	0.99
	ET	0.99	0.99	0.99	0.99	0.99

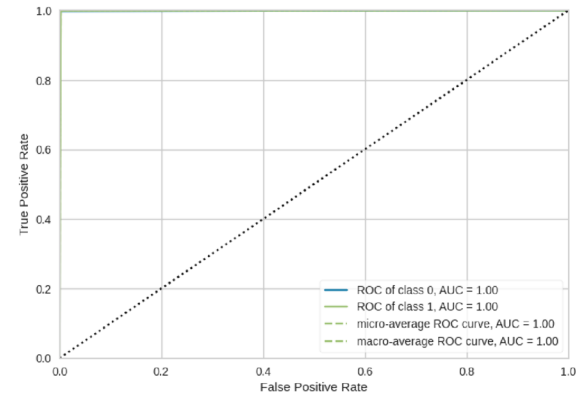
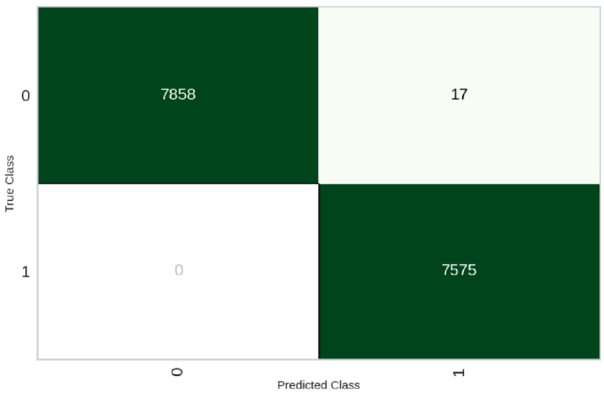
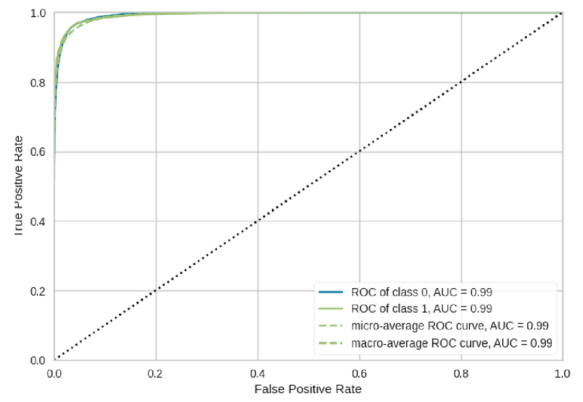
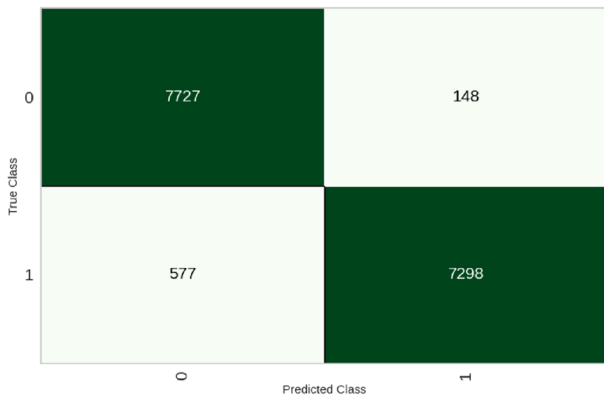


Fig. 5 Confusion matrices and ROC curves for each subject

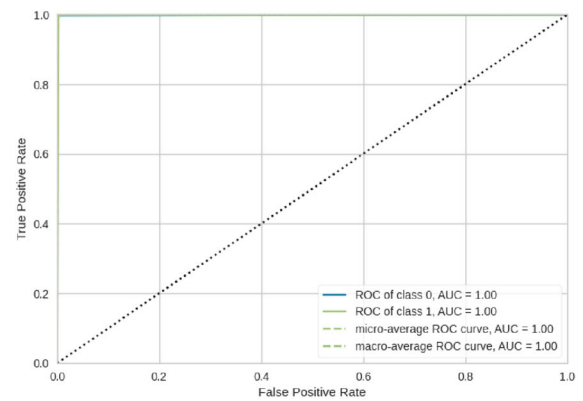
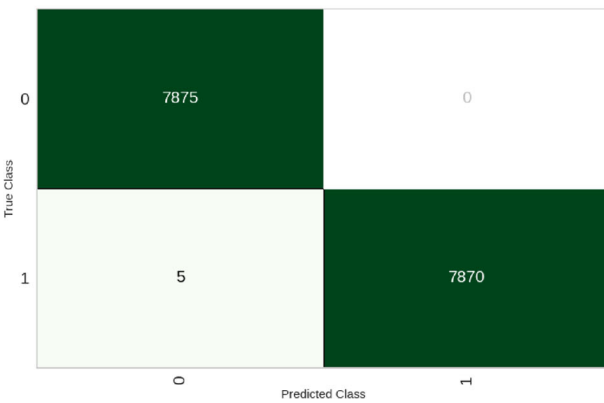
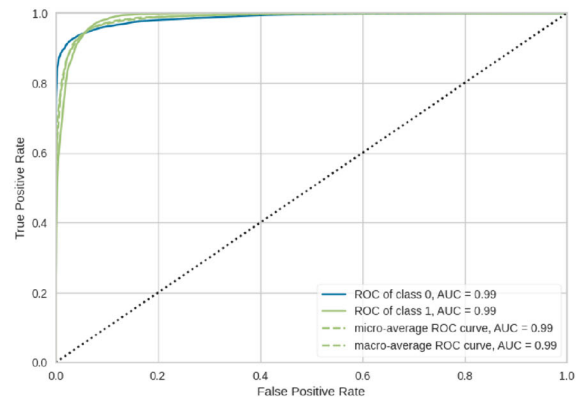
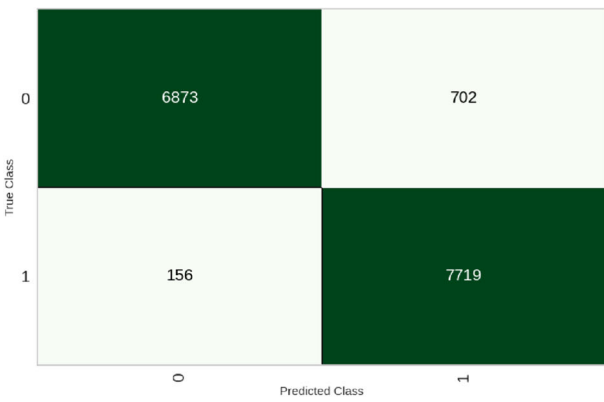


Fig. 5 continued

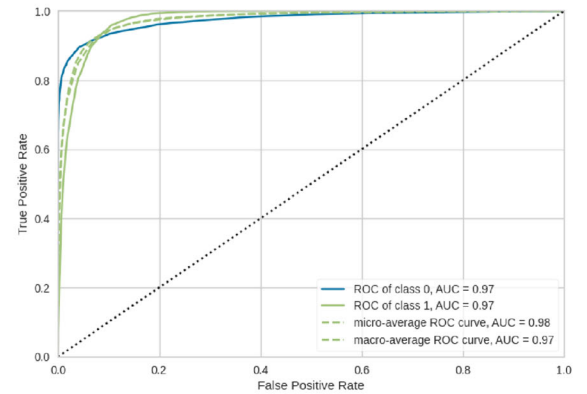
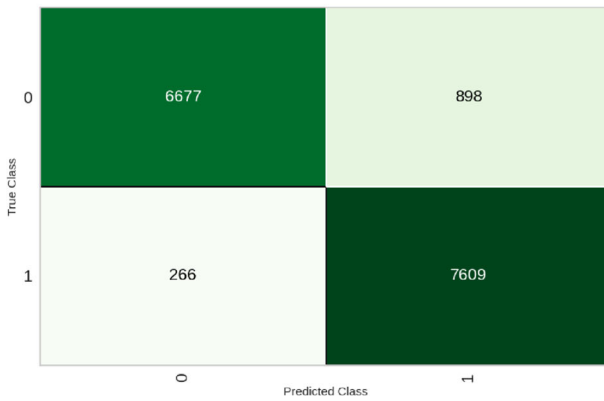
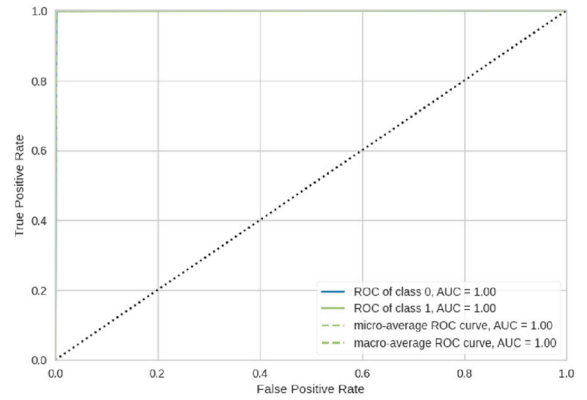
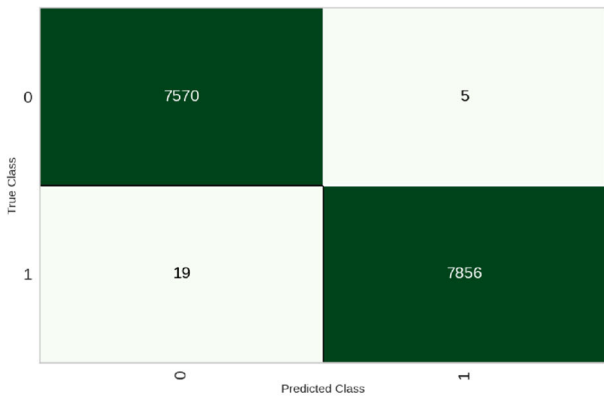
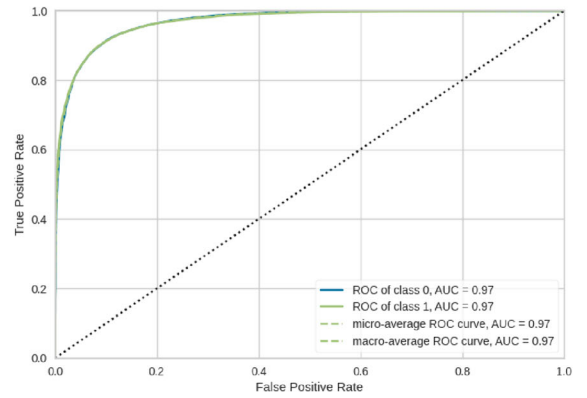
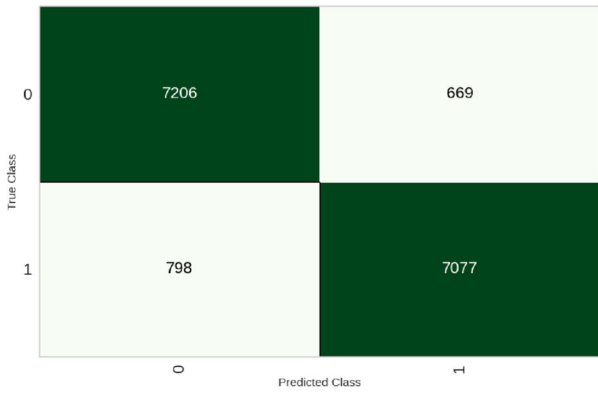
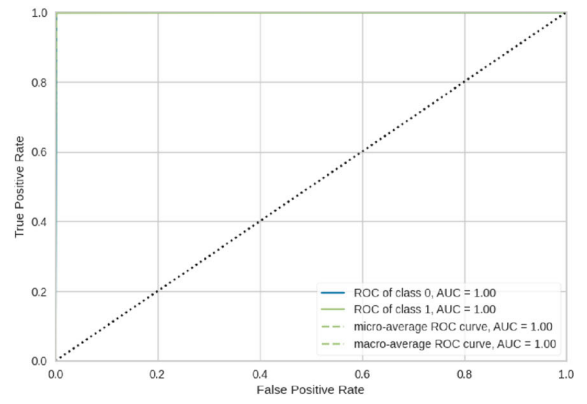
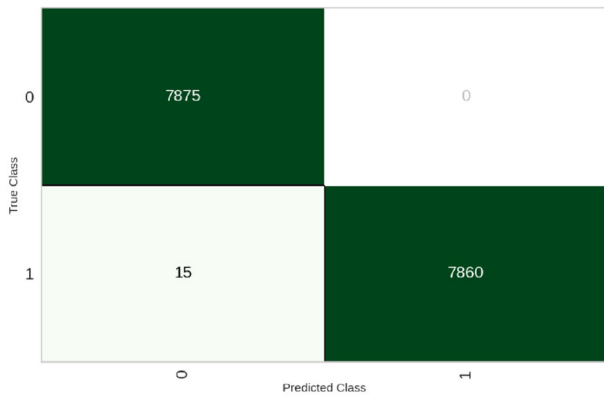


Fig. 5 continued

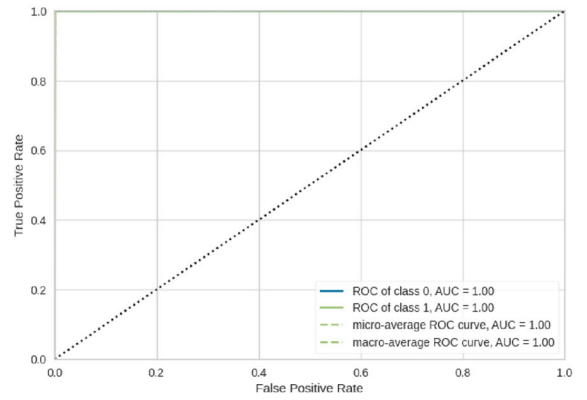
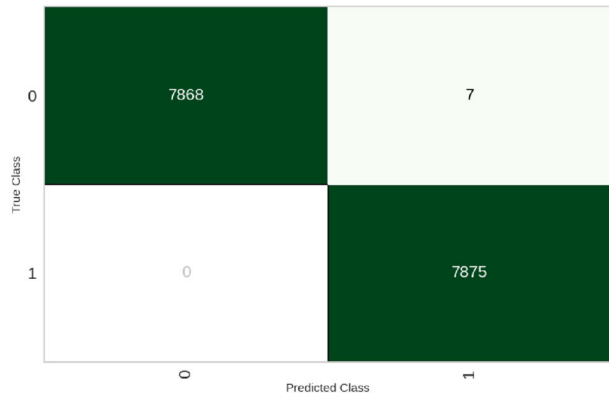
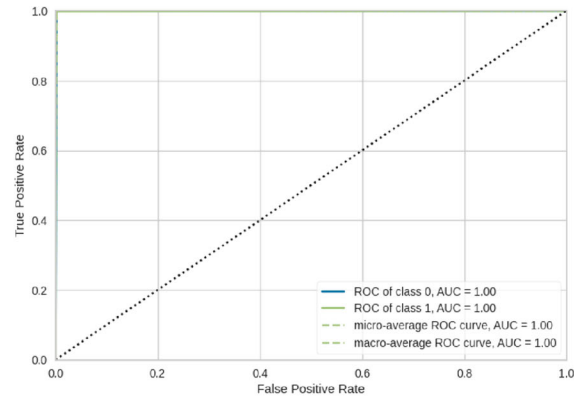
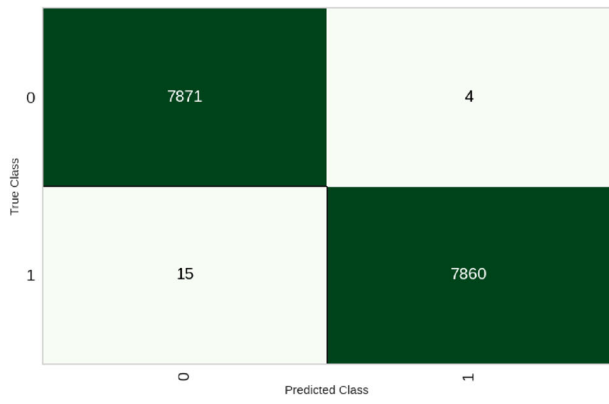


Fig. 5 continued

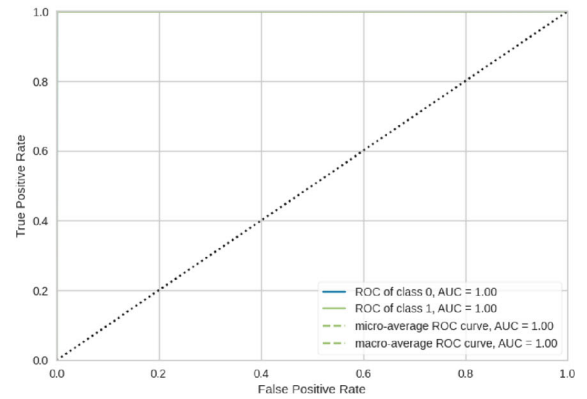
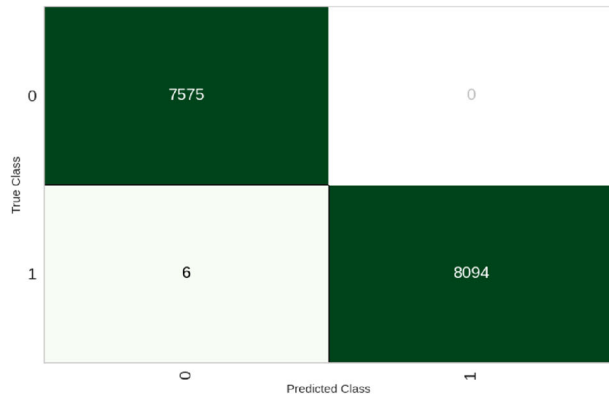
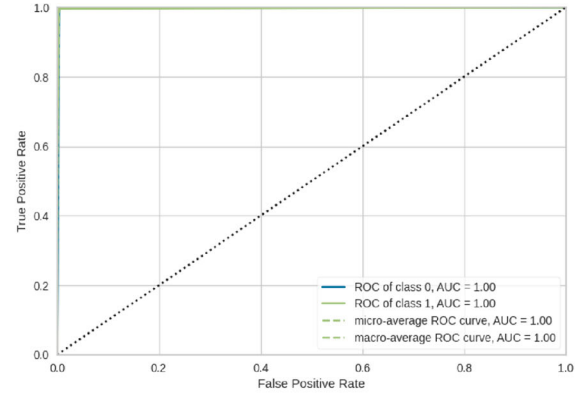
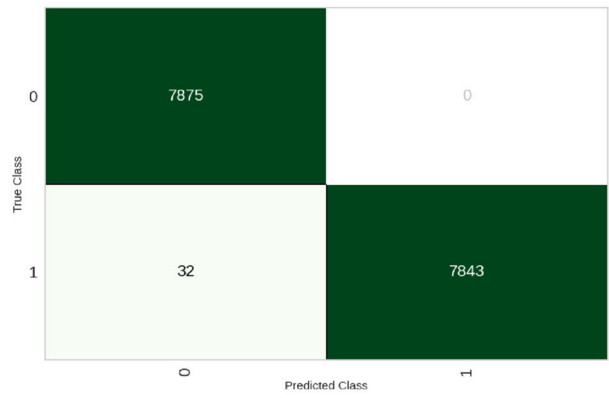


Fig. 5 continued



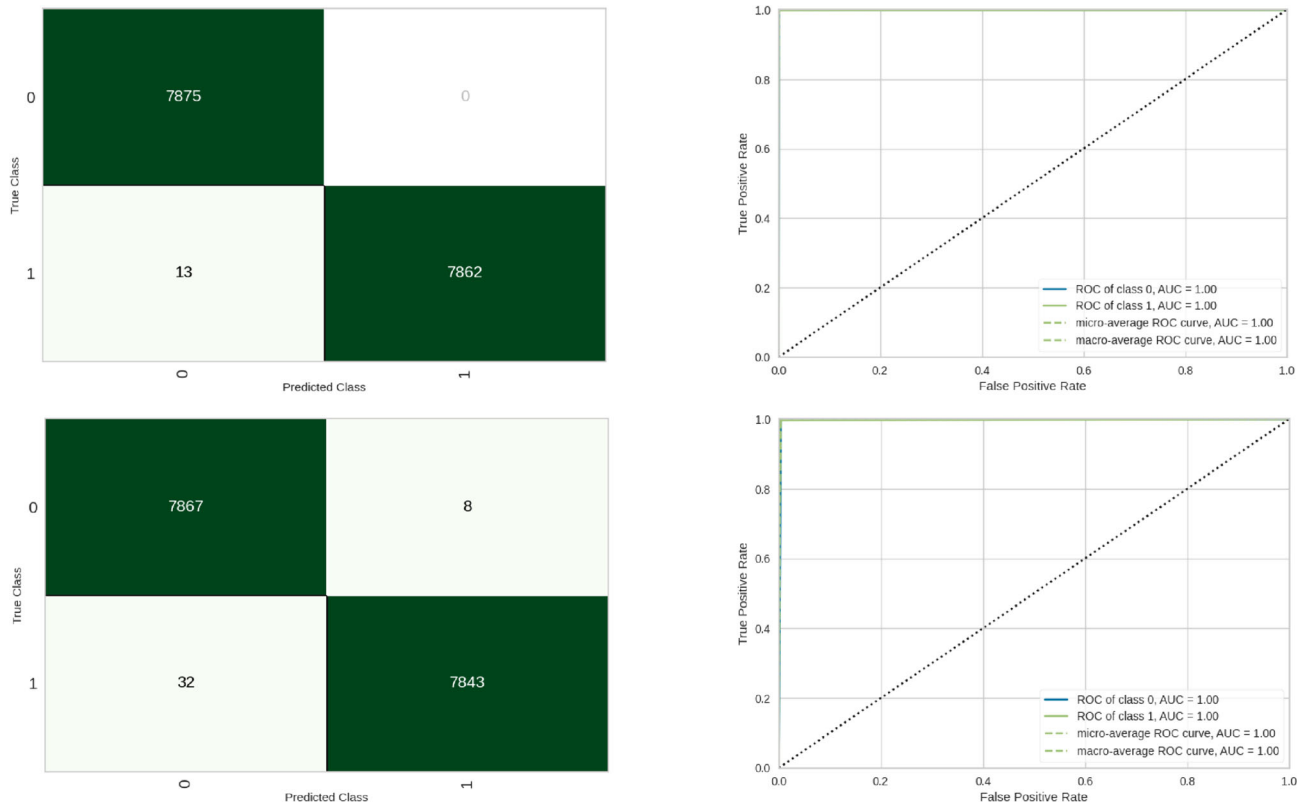


Fig. 5 continued

Table 10 Comparison of the F1-score of DT, RF and ET with original features and SCMPA selected features for EGG dataset

Features	Models	Subjects													
		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14
Original	DT	0.89	0.83	0.88	0.86	0.91	0.78	0.89	0.86	0.82	0.85	0.92	0.97	0.87	0.92
	RF	0.95	0.9	0.94	0.92	0.96	0.89	0.95	0.92	0.93	0.93	0.96	0.98	0.94	0.93
	ET	0.95	0.89	0.95	0.91	0.96	0.9	0.95	0.93	0.93	0.93	0.97	0.98	0.94	0.93
SCMPA	DT	0.89	<b>0.99</b>	0.88	<b>0.99</b>	<b>0.99</b>	0.78	<b>0.99</b>	0.85	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
	RF	0.95	<b>0.99</b>	0.94	<b>0.99</b>	<b>0.99</b>	0.89	<b>0.99</b>	0.92	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
	ET	0.95	<b>0.99</b>	0.95	<b>0.99</b>	<b>0.99</b>	0.9	<b>0.99</b>	0.93	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>

The best classification results are given as bold

### 4 Conclusions

In this study, a new hybrid algorithm is proposed that involves modifying the parameters of the MPA algorithm to improve the algorithm’s performance by using a random number sequence derived from five chaotic maps. The goal of the proposed algorithms is to minimize the number of features chosen during feature selection while maximizing the accuracy rate for the EEG signal classification task. Compared with other chaotic-based MPAs, the HCMPA, TICMPA, LCMPA, and SCMPA can be used to select

more representative features from the experimental dataset. The SCMPA selects a different number of features for each subject. These features are fed into a DT based on three different classifiers. The experimental results show that more stable and accurate results are achieved via classification via SCMPA-selected features. The only shortcoming with SCMPA is that SCMPA handles only single-objective continuous optimization problems. Consequently, the following can be regarded as the main topics of future research: devising SCMPA to address binary, multi-objective, and discrete space optimization problems.

**Author contributions** ZG: Supervision, Conceptualization, Methodology, Software, Writing, Reviewing and Editing. EE: Supervision, Conceptualization, Methodology, Software, Writing, Reviewing and Editing. KS: Supervision, Methodology, Writing. SE: Supervision, Methodology, Writing.

**Funding** Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK). No funding.

**Data availability** The datasets used in the study are publicly available.

## Declarations

**Conflict of interest** The authors declare no competing interests.

**Ethical approval** Not applicable.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Niedermeyer, E., da Silva, F.L. (eds.): *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins (2005)
- Rochais, C., Sébilleau, M., Ménoret, M., Oger, M., Henry, S., Hausberger, M., Cousillas, H.: Attentional state and brain processes: state-dependent lateralization of EEG profiles in horses. *Sci. Rep.* **8**(1), 10153 (2018)
- Lagopoulos, J., Xu, J., Rasmussen, I., Vik, A., Malhi, G.S., Eliassen, C.F., Ellingsen, Ø.: Increased theta and alpha EEG activity during nondirective meditation. *J. Altern. Complement. Med. Altern. Complement. Med.* **15**(11), 1187–1192 (2009)
- Smith, S.J.: EEG in the diagnosis, classification, and management of patients with epilepsy. *J. Neurol. Neurosurg. Psychiatry Neurosurg. Psychiatry* **76**(suppl 2), ii2–ii7 (2005)
- Peter-Derex, L., Berthomier, C., Taillard, J., Berthomier, P., Bouet, R., Mattout, J., Bastuji, H.: Automatic analysis of single-channel sleep EEG in a large spectrum of sleep disorders. *J. Clin. Sleep Med. Clin. Sleep Med.* **17**(3), 393–402 (2021)
- Wendling, F., Ansari-Asl, K., Bartolomei, F., Senhadji, L.: From EEG signals to brain connectivity: a model-based evaluation of interdependence measures. *J. Neurosci. Methods* **183**(1), 9–18 (2009)
- Ünal, Ö., Özcan, Ö., Öner, Ö., Akcakin, M., Aysev, A., Deda, G.: EEG and MRI findings and their relation with intellectual disability in pervasive developmental disorders. *World J. Pediatr.* **5**, 196–200 (2009)
- Choi, K.M., Kim, J.Y., Kim, Y.W., Han, J.W., Im, C.H., Lee, S.H.: Comparative analysis of default mode networks in major psychiatric disorders using resting-state EEG. *Sci. Rep.* **11**(1), 22007 (2021)
- Eroglu, C., Brand, A., Hildebrandt, H., Kedzior, K.K., Mathes, B., Schmiedt, C.: Working memory related gamma oscillations in schizophrenia patients. *Int. J. Psychophysiol.* **64**(1), 39–45 (2007)
- Baradits, M., Kakuszi, B., Bálint, S., Fullajtár, M., Mód, L., Bitter, I., Czobor, P.: Alterations in resting-state gamma activity in patients with schizophrenia: a high-density EEG study. *Eur. Arch. Psychiatry Clin. Neurosci.* **269**, 429–437 (2019)
- Turetsky, B.I., Dress, E.M., Braff, D.L., Calkins, M.E., Green, M.F., Greenwood, T.A., Light, G.: The utility of P300 as a schizophrenia endophenotype and predictive biomarker: clinical and sociodemographic modulators in COGS-2. *Schizophr. Res. Res.* **163**(1–3), 53–62 (2015)
- Oh, S.L., Vicsesh, J., Ciaccio, E.J., Yuvaraj, R., Acharya, U.R.: Deep convolutional neural network model for automated diagnosis of schizophrenia using EEG signals. *Appl. Sci.* **9**(14), 2870 (2019)
- Shalhaf, A., Bagherzadeh, S., Maghsoudi, A.: Transfer learning with deep convolutional neural network for automated detection of schizophrenia from EEG signals. *Phys. Eng. Sci. Med.* **43**, 1229–1239 (2020)
- Supakar, R., Satvaya, P., Chakrabarti, P.: A deep learning based model using RNN-LSTM for the detection of schizophrenia from EEG data. *Comput. Biol. Med. Biol. Med.* **151**, 106225 (2022)
- Sun, J., Cao, R., Zhou, M., Hussain, W., Wang, B., Xue, J., Xiang, J.: A hybrid deep neural network for classification of schizophrenia using EEG data. *Sci. Rep.* **11**(1), 1–16 (2021)
- Wan, Z., Yang, R., Huang, M., Zeng, N., Liu, X.: A review on transfer learning in EEG signal analysis. *Neurocomputing* **421**, 1–14 (2021)
- Atban, F., Ekinci, E., Garip, Z.: Traditional machine learning algorithms for breast cancer image classification with optimized deep features. *Biomed. Signal Process. Control* **81**, 104534 (2023)
- Ay, Ş., Ekinci, E., & Garip, Z.: A comparative analysis of meta-heuristic optimization algorithms for feature selection on ML-based classification of heart-related diseases. *J. Supercomput.* 1–30 (2023)
- Prabhakar, S.K., Rajaguru, H., Lee, S.W.: A framework for schizophrenia EEG signal classification with nature inspired optimization algorithms. *IEEE Access* **8**, 39875–39897 (2020)
- Khare, S.K., Bajaj, V.: A hybrid decision support system for automatic detection of Schizophrenia using EEG signals. *Comput. Biol. Med.* **141**, 105028 (2022)
- Bingol, H., Alatas, B.: Chaos based optics inspired optimization algorithms as global solution search approach. *Chaos Solitons Fract* **141**, 110434 (2020)
- Sayed, G., Tharwat, A., Hassanien, A.: Chaotic dragonfly algorithm: an improved metaheuristic algorithm for feature selection. *Appl. Intell.* **49**, 188–205 (2019)
- Che, Y., Che, K., & Li, Q.: Application of decision tree in PE teaching analysis and management under the background of big data. *Comput. Intell. Neurosci.* 2022. (2022)
- Hassan, F., Hussain, S.F., Qaisar, S.M.: Fusion of multivariate EEG signals for schizophrenia detection using CNN and machine learning techniques. *Inf. Fusion* **92**, 466–478 (2023)
- Kumar, T.S., Rajesh, K.N., Maheswari, S., Kanhangad, V., Acharya, U.R.: Automated schizophrenia detection using local descriptors with EEG signals. *Eng. Appl. Artif. Intell.* **117**, 105602 (2023)
- Gosala, B., Kapgate, P.D., Jain, P., Chaurasia, R.N., Gupta, M.: Wavelet transforms for feature engineering in EEG data processing: An application on Schizophrenia. *Biomed. Signal Process. Control* **85**, 104811 (2023)
- Li, B., Wang, J., Guo, Z., Li, Y.: Automatic detection of schizophrenia based on spatial-temporal feature mapping and LeViT with EEG signals. *Expert Syst. Appl.* **224**, 119969 (2023)

28. Agarwal, M., Singhal, A.: Fusion of pattern-based and statistical features for Schizophrenia detection from EEG signals. *Med. Eng. Phys.* **112**, 103949 (2023)
29. Aslan, Z., Akin, M.: A deep learning approach in automated detection of schizophrenia using scalogram images of EEG signals. *Phys. Eng. Sci. Med.* **45**(1), 83–96 (2022)
30. Baygin, M., Yaman, O., Tuncer, T., Dogan, S., Barua, P.D., Acharya, U.R.: Automated accurate schizophrenia detection system using Collatz pattern technique with EEG signals. *Biomed. Signal Process. Control* **70**, 102936 (2021)
31. Das, K., Pachori, R.B.: Schizophrenia detection technique using multivariate iterative filtering and multichannel EEG signals. *Biomed. Signal Process. Control* **67**, 102525 (2021)
32. Bagherzadeh, S., Shahabi, M.S., Shalbaf, A.: Detection of schizophrenia using hybrid of deep learning and brain effective connectivity image from electroencephalogram signal. *Comput. Biol. Med.* **146**, 105570 (2022)
33. Goshvarpour, A., Goshvarpour, A.: Schizophrenia diagnosis by weighting the entropy measures of the selected EEG channel. *J. Med. Biol. Eng.* **42**(6), 898–908 (2022)
34. Faramarzi, A., Heidarinejad, M., Mirjalili, S., Gandomi, A.H.: Marine predators algorithm: a nature-inspired metaheuristic. *Expert Syst. Appl.* **152**, 113377 (2020)
35. Yousri, D., Fathy, A., Rezk, H.: A new comprehensive learning marine predator algorithm for extracting the optimal parameters of supercapacitor model. *J. Energy Storage* **42**, 103035 (2021)
36. Houssein, E.H., Abdelminaam, D.S., Ibrahim, I.E., Hassaballah, M., Wazery, Y.M.: A hybrid heartbeats classification approach based on marine predators algorithm and convolution neural networks. *IEEE Access* **9**, 86194–86206 (2021)
37. Arora, S., Anand, P.: Chaotic grasshopper optimization algorithm for global optimization. *Neural Comput. Appl.* **31**, 4385–4405 (2019)
38. Zawbaa, H.M., Emary, E., Grosan, C.: Feature selection via chaotic antlion optimization. *PLoS ONE* **11**(3), e0150652 (2016). <https://doi.org/10.1371/journal.pone.0150652>
39. Kingsford, C., Salzberg, S.L.: What are decision trees? *Nat. Biotechnol.* **26**(9), 1011–1013 (2008)
40. Biau, G., Scornet, E.: A random forest guided tour. *TEST* **25**, 197–227 (2016)
41. Cutler, A., Cutler, D. R., & Stevens, J. R.: Random forests. *Ensemble Mach. Learn.: Methods Appl.* 157–175. (2012)
42. Afshar, F., Seyedabrishami, S., Moridpour, S.: Application of extremely randomized trees for exploring influential factors on variant crash severity data. *Sci. Rep.* **12**(1), 11476 (2022)
43. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. *Mach. Learn.* **63**, 3–42 (2006)
44. Geurts, P., & Louppe, G.: Learning to rank with extremely randomized trees. In: *Proceedings of the Learning to Rank Challenge* (pp. 49–61). PMLR (2011)
45. Olejarczyk E, Jernajczyk W.: “EEG in schizophrenia”, (2017). <https://doi.org/10.18150/repod.0107441>, (2017)RepOD, V1

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



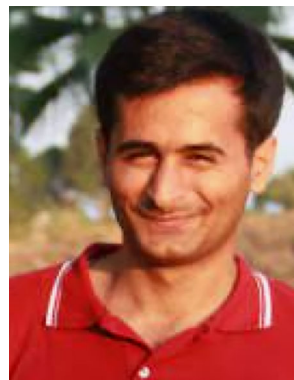
**Zeynep Garip** received Ph.D. degree from the Department of Mechatronics Engineering, Sakarya University, Turkey, in 2018. She is currently an Assoc. Professor at the Sakarya University of Applied Sciences, Turkey. Her current research interests include metaheuristics algorithms and image processing.



**Ekin Ekinci** received Ph.D. degree from the Department of Computer Engineering, Kocaeli University, Turkey, in 2019. She is currently an Assoc. Professor at the Sakarya University of Applied Sciences, Turkey. Her current research interests include artificial intelligence, machine learning and nature language processing.



**Kasim Serbest** is an Assoc. Professor in Mechatronics Engineering. His research involves movement biomechanics, design and development of assistive devices, and machine learning for motion analysis.



**Süleyman Eken** received the M.S. and Ph.D. degrees in computer engineering from Kocaeli University, Izmit, Turkey. He was a Research Assistant at Kocaeli University from 2010 to 2019, where he is currently working as an Assoc. Professor of Information Systems Engineering. His main research work focuses on distributed systems and big data analysis.