

Metaheuristic Optimization Algorithm for Signals Classification of Electroencephalography Channels

Marwa M. Eid^{1,*}, Fawaz Alassery², Abdelhameed Ibrahim³ and Mohamed Saber⁴

¹Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura, 35111, Egypt

²Department of Computer Engineering, College of Computers and Information Technology, Taif University, Taif, 21944, Saudi Arabia

³Computer Engineering and Control Systems Department, Faculty of Engineering, Mansoura University, Mansoura, 35516, Egypt

⁴Department of Electronics and Communications Engineering, Faculty of Engineering, Delta University for Science and Technology, Mansoura, 11152, Egypt

*Corresponding Author: Marwa M. Eid. Email: marwa.3eed@gmail.com

Received: 01 October 2021; Accepted: 05 November 2021

Abstract: Digital signal processing of electroencephalography (EEG) data is now widely utilized in various applications, including motor imagery classification, seizure detection and prediction, emotion classification, mental task classification, drug impact identification and sleep state classification. With the increasing number of recorded EEG channels, it has become clear that effective channel selection algorithms are required for various applications. Guided Whale Optimization Method (Guided WOA), a suggested feature selection algorithm based on Stochastic Fractal Search (SFS) technique, evaluates the chosen subset of channels. This may be used to select the optimum EEG channels for use in Brain-Computer Interfaces (BCIs), the method for identifying essential and irrelevant characteristics in a dataset, and the complexity to be eliminated. This enables (SFS-Guided WOA) algorithm to choose the most appropriate EEG channels while assisting machine learning classification in its tasks and training the classifier with the dataset. The (SFS-Guided WOA) algorithm is superior in performance metrics, and statistical tests such as ANOVA and Wilcoxon rank-sum are used to demonstrate this.

Keywords: Signals; metaheuristics optimization; feature selection; multilayer perceptron; support vector machines

1 Introduction

Digital signal processing is critical for several applications, including seizure detection/prediction, sleep state classification, and categorization of motor imagery. As shown in Fig. 1, digital EEG signal processing consists of three components: a signal collection unit, a feature extraction unit, and a decision algorithm. The EEG signal collected from the scalp, brain surface, or brain interior is used



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

as the system's input. Electrodes, whether invasive or non-invasive, are used to represent the signal acquisition unit. The feature extraction unit is a signal processing device responsible for extracting distinguishing characteristics from a channel (s). For example, in a brain-computer interface (BCI), the decision unit is a hybrid unit that performs categorization, decision-making, and decision-passing to external devices that output the subject's intention [1].

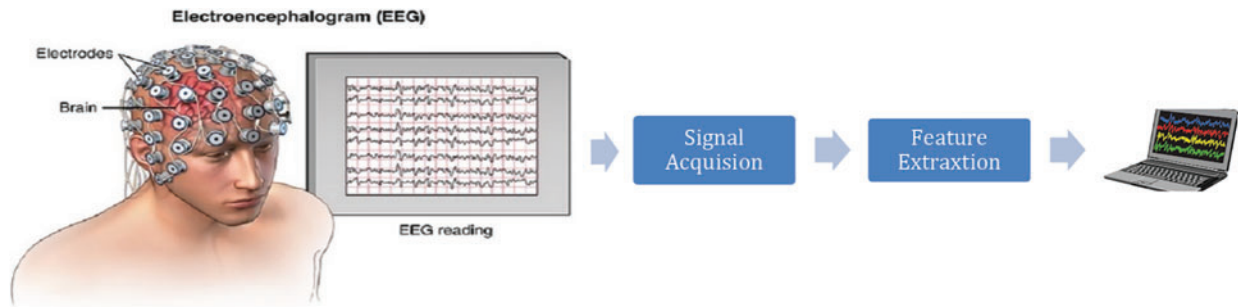


Figure 1: Processing of EEG signals

As previously stated, the interface between the brain and the computer (or another device) may be intrusive or non-invasive. While invasive technologies have recently demonstrated some promise in a variety of applications due to their high accuracy and low noise [2], noninvasive technologies are still widely used for safety applications with some additional signal processing tasks to compensate for noise and resolution limitations. Scalp EEG acquisition devices are usually chosen because they are inexpensive, simple to use, portable, and provide excellent temporal resolution. The scalp EEG waves may be recorded in a variety of modalities, including unipolar and bipolar. The former mode records the voltage differential between all electrodes and a reference electrode, where each electrode-reference pair forms a channel. In contrast, the bipolar mode records the voltage differences between two designated electrodes, each pair forming a channel. The International Federation of Societies for Electroencephalography and Clinical Neurophysiology (IFSECN) proposed an electrode placement method on the scalp dubbed the International 10–20 system [3]. Fig. 2 illustrates the 10–20 EEG electrode locations for electrode insertion on the left and top of the head. These electrodes (channels) depict the activity of several brain regions.

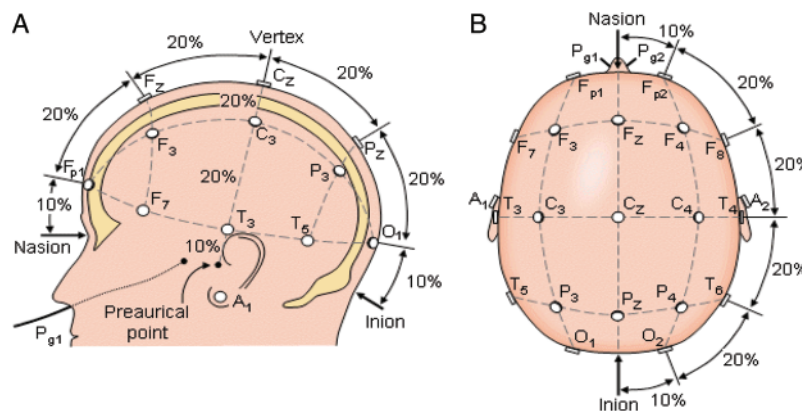


Figure 2: EEG 10–20 electrode placement [3]

The brain regions are shown in Fig. 3. Most of the relevant information about the functioning condition of the human brain is contained in five main brain waves, each with its own distinct frequency band. Delta band within (0–4 Hz), theta band within (3.5–7.5 Hz), alpha band within (7.5–13 Hz), beta band within (13–26 Hz), and gamma band within (26–70 Hz) are these frequency bands [4]. Delta waves are associated with profound slumber. Theta waves are associated with the most meditative state (body asleep/mind awake). Alpha waves are associated with dreams and relaxation. Beta waves are the most prevalent during the waking state of intense concentration. Gamma waves are intimately connected with the brain’s decision-making process. When dealing with mental disease situations, unanticipated changes in brain waves occur, necessitating a significant amount of signal processing to diagnose aberrant conditions [4]. The frequency range, speed, mental state, and waveforms of the EEG are shown in Tab. 1.

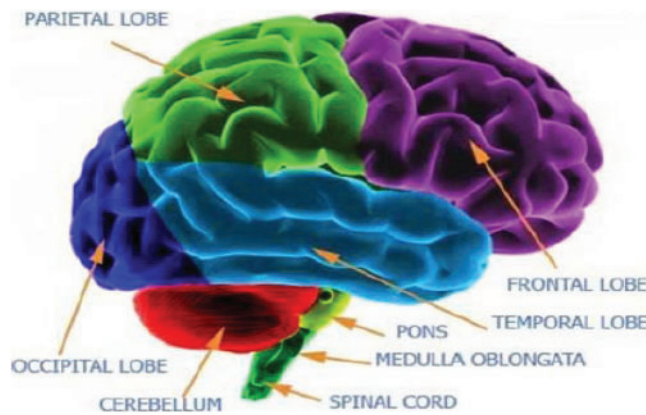


Figure 3: Human brain and various lobes [4]

Table 1: EEG frequency band, speed, state, and waveforms [4]

Frequency band	Speed (Hz)	Mental state	EEG recording (1 Sec.)
Delta	1–4 (Slow)	Deep sleep	
Theta	4–8	Drowsy	
Alpha	8–12	Relaxed	
Beta	12–30 (Fast)	Focused	

The EEG signals collected are often multi-channel in nature. For example, we have two options while classifying these signals: work on a subset of channels chosen based on specific criteria or work

on all channels [5]. The method of EEG data categorization based on channel selection is shown in Fig. 4. Reduce the number of channels in this signal processing environment since the setup procedure with many channels is time-consuming and inconvenient for the subject. Additionally, it increases the computational complexity of the system, which some applications need to be minimal.



Figure 4: General process of EEG signal classification

Seizure prediction and detection is another area where channel reduction may be helpful. The scientific and industry community are particularly interested in medical support systems' portable development. This can detect the onset of epileptic seizures early or even hours in advance by incorporating algorithms, thereby avoiding injury [6,7]. Developing such portable systems based on computationally efficient prediction algorithms that use the fewest possible channels to reduce system power consumption is critical to system longevity. In the processing of EEG data, numerous channel selection techniques have been investigated.

The current work contribution can be summarized as follow.

1. A continuous version of the Guided Whale Optimization based on Stochastic Fractal Search algorithm (Continuous SFS-Guided WOA) is presented.
2. A binary version of the SFS-Guided WOA algorithm (Binary SFS-Guided WOA) is also presented.
3. Two publicly accessible datasets for electroencephalogram (EEG) signal processing, named BCI Competition IV-dataset 2a and BCI Competition IV-data set III, are utilized to test the suggested method.
4. The SFS-Guided WOA algorithm is employed to evaluate the chosen subset of EEG channels of the two datasets.
5. This is used to select the optimum EEG channels for use in Brain-Computer Interfaces (BCIs).
6. Statistical tests such as ANOVA and Wilcoxon rank-sum are used to demonstrate the presented method's performance.

2 Literature Review

Feature selection methods may be categorized as filter-based, wrapper-based, or hybrid-based [8–10]. The advantage of filter-based completely characteristic selection methods over traditional characteristic selection strategies is their speed and capacity to expand to large datasets.

2.1 Related Work

The optimization technique is widely used in various fields of study, including computer science, engineering [11], health [12], agriculture, and feature selection [13]. The primary aim of optimization is to choose the optimal solution to a given problem among the available solutions that match the problem description. Additionally, optimization algorithms have a goal that must be reduced or maximized under the addressed problem [14–16].

Recently, numerous studies have used optimization to resolve given problems, such as the Whale Optimization Algorithm (WOA). WOA was used to locate the optimal weights for training the neural community and developed a multi-objective model of WOA, which was then applied to the problem of

forecasting wind speed. Additionally, WOA was widely employed to determine the final location and length of capacitors used inside the radial system [17]. Additionally, they used WOA to circumvent the difficulty of determining the final length used by a distributed generator [18], and they benefited from the use of WOA for image segmentation [19,20].

2.2 EEG Signal

The nature of EEG alerts may be very complicated since they are no longer linked, however random. The EEG dimension is determined by various factors, most notably the individual's age, gender, psychological state, and intellectual state of the issue [21]. Thus, comprehending the behavior and movement of brain cells involves various linear and nonlinear signal-processing methods that result in the physiological state and circumstances of the issue. Numerous ways are advocated for capturing the dynamic capabilities and sudden changes that may occur. The first step is preprocessing, which includes recording warnings, removing artifacts, signal averaging, output thresholding, and signal enhancement. The second stage is the function extraction technique within the procedure, which determines a feature vector from an ordinary vector [22].

2.3 Traditional Optimizer

Optimizer of Genetic Algorithm (GA) is inspired by biology (survival of the fittest). Initialization is a critical GA process. Alternatively, other genetic operators, such as elitism, may be used [23]. The advantages of this optimizer are its simplicity and ability to deal with noisy fitness functions. Due to delayed convergence, premature convergence, and parameter change, complexity is not scalable. This technique is utilized in the construction of image processing filters as well as antennas.

Particle Swarm Optimization (PSO) is another optimization technique that simulates the motions and interactions of individuals in a flock of birds or a school of fish [24]. Every particle is guided by its best-known location and the swarm's optimal position. It is stable, simple to implement, and a suitable model of collaboration, but starting settings are elusive. It has a long convergence time and a high computational cost. Gene clustering, antenna design, vehicle routing, control design, and dimension reduction are only a few examples of uses.

The Grey Wolf Optimizer (GWO) is an algorithm that mimics grey wolf leadership, social structure, and hunting behavior. Encircling and assaulting the victim are the first two phases. This optimizer possesses Exploration and exploitation must be conducted in a balanced manner. While high search accuracy is simple to implement, it results in premature convergence due to the fluctuating positions of the three leaders. The greater the number of variables, the lower performance is achieved. It is utilized in feature selection, parameter adjustment of PID controllers, clustering, robotics, and route finding [25].

The foraging behaviors of humpback whales inspired the Whale Optimization Algorithm (WOA). They catch fish with bubbles as they swirl around a school of fish. It is a simple method for exploring a vast search space that is sluggish to convergence, prone to local optima stagnation, and computationally costly. WOA is applied in route planning, voltage offset reduction, and precision control of laser sensor systems [26].

3 Suggested (SFS-Guided WOA) Algorithm

The Guided WOA is a variant of the standard WOA. In the Guided WOA technique, to address the main disadvantage of this method, the search strategy for a single random whale may be substituted

with an advanced design capable of quickly moving the whales toward the optimal solution or prey. The original WOA compels whales to travel randomly around one another, comparable to the global search. A whale may follow three random whales rather than one to improve exploration performance in the modified WOA (Guided WOA) [27]. This may encourage whales to do more exploration while remaining unaffected by the leading position.

The Stochastic Fractal Search (SFS) technique's diffusion process may generate a sequence of random walks around the optimum solution. This enhances the Guided WOA's exploration capacity by using this diffusion process to find the optimal solution. Gaussian random walks are used as a component of the diffusion process that occurs around the updated optimum position. Algorithm 1 shows the continuous version of the SFS-Guided WOA algorithm. The binary conversion of the algorithm is shown in Algorithm 2, which explains step by step how to convert the continuous algorithm to a binary one to be applied for the tested EEG problem.

Algorithm 1: Continuous SFS-Guided WOA Algorithm

1: **Initialize** Guided WOA population $G_i (i = 1, 2, \dots, n)$ with size n , maximum iterations Max_{iter} , fitness function F_n .

2: **Initialize** Guided WOA parameters $a, A, C, l, r_1, r_2, r_3$

3: **Initialize** Guided WOA parameters w_1, w_2, w_3

4: **Set** $t = 1$

5: **Calculate** fitness function F_n for each G_i

6: **Find** best individual G^*

7: **while** $t \leq Max_{iter}$ (Termination condition) **do**

8: **for** $(i = 1; i < n + 1)$ **do**

9: **if** $(r_3 < 0.5)$ **then**

10: **if** $(|A| < 1)$ **then**

11: **Update** position of current search agent as $G(t + 1) = G^*(t) - A \cdot D$

12: **else**

13: **Select** three random search agents G_{rand1}, G_{rand2} , and G_{rand3}

14: **Update** (z) by the exponential form of

$$z = 1 - \left(\frac{t}{Max_{iter}} \right)^2$$

15: **Update** position of current search agent as

$$G(t + 1) = w_1 * G_{rand1} + z * w_2 * (G_{rand2} - G_{rand3}) + (1 - z) * w_3 * (G - G_{rand1})$$

16: **end if**

17: **else**

18: **Update** position of current search agent as

$$G(t + 1) = D \cdot e^{bl} \cdot \cos(2\pi l) + G^*(t)$$

19: **end if**

20: **end for**

21: **for** $(i = 1; i < n + 1)$ **do**

22: **Calculate** $G_i^* = Gaussian(\mu G^*, \sigma) + (\eta \times G^* - \eta' \times P_i)$

23: **end for**

24: **Update** a, A, C, l, r_3

25: **Calculate** fitness function F_n for each G_i

26: **Find** best individual G^*

(Continued)

Algorithm 1: Continued

27: **Set** $t = t + 1$ 28: **end while**29: **return** G^*

Algorithm 2: Binary SFS-Guided WOA Algorithm

1: **Initialize** SFS-Guided WOA algorithm configuration, including population and parameters2: **Change** current solutions to binary solution (0 or 1)3: **Evaluate** fitness function and determine the best solution4: **Train** k-NN based model and then calculate error5: **while** $t \leq iters_{max}$ **do**6: **Apply** SF5-Guided WOA algorithm7: **Change** updated solution to binary solution (0 or 1) based on the following equation

$$G_d^{(t+1)} = \begin{cases} 1 & \text{if } \frac{1}{1 + e^{-10(G^* - 0.5)}} \geq 0.5 \\ 0 & \text{otherwise} \end{cases}$$

8: **Evaluate** fitness function for each agent9: **Update** parameters10: **Update** best solution11: **end while**12: **Return** optimal solution

4 Results and Discussion

This section discusses the experimental results. The data preprocessing process is explained, including the description and the correlation matrix of tested EEG datasets. Configuration of the suggested algorithm is also discussed. Performance metrics and results discussion are described in detail in this part.

4.1 Data Preprocessing

Two publicly accessible datasets for electroencephalogram (EEG) signal processing are utilized in this work to test our suggested method. [Tab. 2](#) shows the description of the dataset. The Statistics of the EEG Dataset is discussed in [Tab. 3](#) and [Fig. 5](#) shows the correlation matrix of EEG dataset. The BCI Competition IV dataset is the first. The fourth BCI competition was held in 2008 at Austria's Graz University of Technology. For the sake of this research, we will analyze dataset 2a from the competition mentioned above. This dataset is freely accessible through [\[28\]](#). The dataset contains the EEG data of nine healthy individuals. The subjects were healthy and ordinary people. They were instructed to complete the motor imagery activities while seated in a comfy armchair in front of an LCD display. To finish all charges, subjects used four distinct kinds of motor imagery. These activities required the use of one's imagination to move the left or right hand, foot, or tongue. To initiate the experimental paradigm, a brief auditory beep was played. Then, after two seconds, a fixation cross appeared on the LCD and was replaced by an arrow pointing up, down, right, or left. The participants completed one of the imaging tasks involving the mouth, feet, and left or right-hand motions, depending on the

orientation of the needle. The performance subject retained the chosen item's imagination for about three seconds until the fixation cross vanished and the LCD became completely dark. Then, after a brief pause of about two seconds, the next job was resumed. This procedure was repeated 72 times for each of the four activities, totaling 288 instances of motor imagery per participant.

Table 2: Datasets description

Dataset	Subscription	# Rows	# Features/Columns
BCI competition IV- dataset 2a	D1	45,000	22
BCI competition IV- data set III	D2	64,000	19

Table 3: Description statistics of EEG dataset

Feature	s1	s2	s3	s4	s5	s6	s7
Number of values	14304	14304	14304	14304	14304	14304	14304
Minimum	4200	3910	4200	4070	4310	4570	4030
25% percentile	4280	3990	4250	4110	4330	4610	4060
Median	4290	4000	4260	4120	4340	4620	4070
75% percentile	4310	4020	4270	4130	4350	4630	4080
Maximum	4470	4150	4350	4190	4400	4670	4140
Range	270.0	240.0	150.0	120.0	90.00	100.0	110.0
Mean	4298	4007	4262	4120	4340	4618	4071
Std. deviation	32.52	27.33	16.44	17.08	12.33	13.26	17.97
Std. error of mean	0.2719	0.2285	0.1375	0.1428	0.1031	0.1109	0.1503
Sum	61480820	57319530	60958540	58935910	62074800	66060020	58231630
s8	s9	s10	s11	s12	s13	s14	cat
14304	14304	14304	14304	14304	14304	14304	14304
4570	4150	4170	4130	4220	4490	4240	0.000
4600	4190	4220	4190	4270	4590	4340	0.000
4610	4200	4230	4200	4280	4600	4350	0.000
4620	4210	4240	4210	4290	4620	4370	1.000
4670	4260	4290	4270	4330	4720	4490	1.000
100.0	110.0	120.0	140.0	110.0	230.0	250.0	1.000
4614	4200	4229	4200	4277	4603	4358	0.4509
14.67	14.12	15.30	18.62	15.35	25.73	32.01	0.4976
0.1226	0.1181	0.1279	0.1557	0.1284	0.2151	0.2676	0.004161
65999470	60071740	60495390	60073290	61176090	65842840	62338180	6449

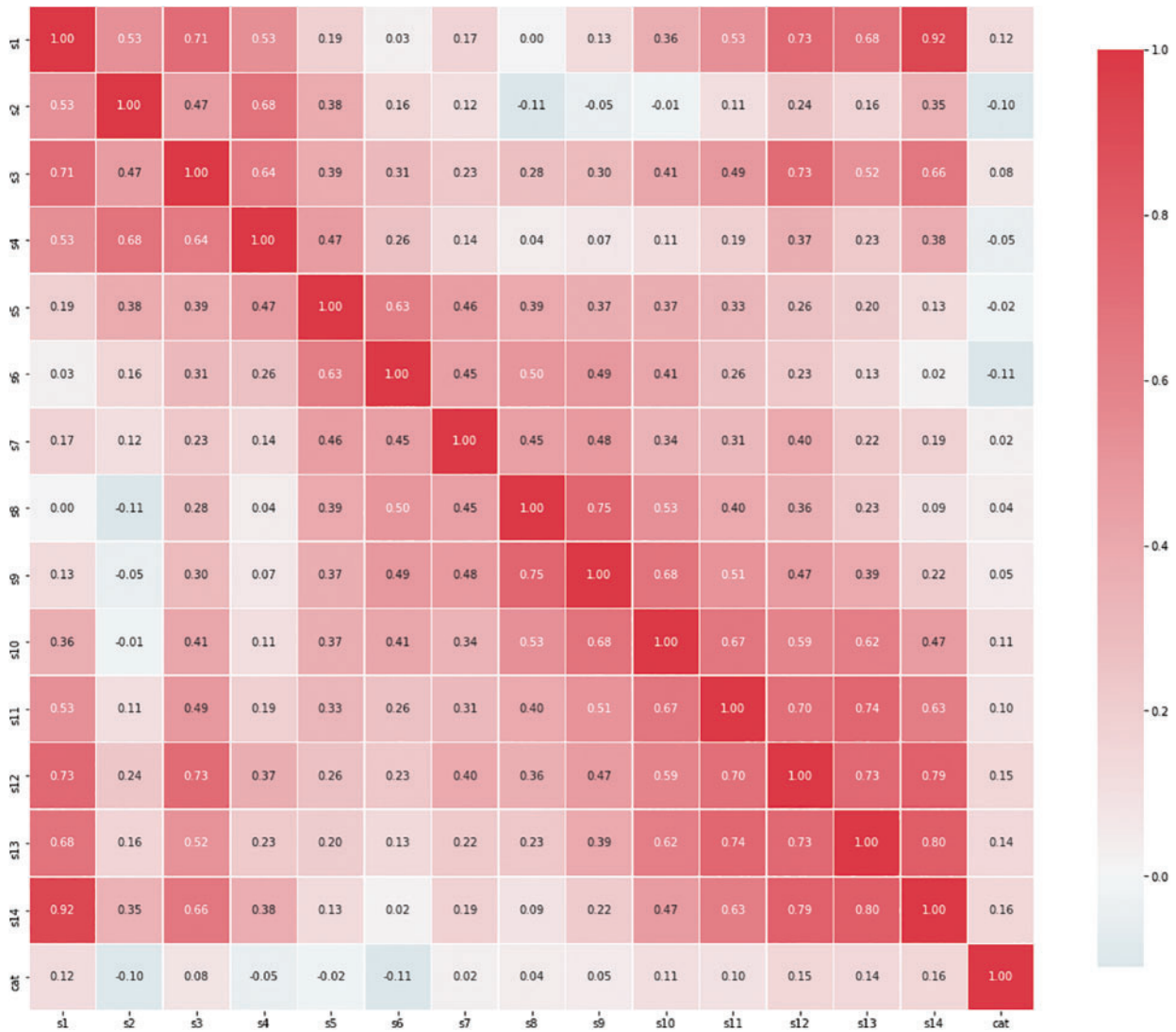


Figure 5: Correlation matrix of EEG dataset

4.2 Configuration

Each dataset is subdivided into three equal-sized segments at random: training, validation, and test. During the learning phase, training is utilized to fine-tune the KNN classifier. Validation is a technique for testing. When determining the fitness function of a particular solution. Normalize data to ensure that all features are contained within the same limits and are handled equally by the machine learning model. One of the simplest methods for scaling data is to use the min-max scaler, which scales and bounds data features between 0 and 1.

4.3 Evaluation Metrics

The evaluation metrics of the suggested method and compared algorithms are shown in Tab. 4. The used variables in Tab. 4 are the number of optimizer’s runs, M , the best solution at the run number

j, g_j^* . In addition, g_j^* vector size, $size(g_j^*)$. A number of tested points, N . Classifier label of the output for a point i , C_i , and label of the class for a point i , L_i . Finally, the total number of features, D , and $Match$, is used to calculate the matching between two inputs.

Table 4: Evaluation metrics

Metric	Value
Average error	$1 - \frac{1}{M} \sum_{j=1}^M \frac{1}{N} \sum_{i=1}^N Match(C_i, L_i)$
Average select size	$\frac{1}{M} \sum_{j=1}^M \frac{M \cdot size(g_j^*)}{D}$
Average fitness	$\frac{1}{M} \sum_{j=1}^M g_j^*$
Best fitness	$Min_{j=1}^M g_j^*$
Worst fitness	$Max_{j=1}^M g_j^*$
Standard deviation	$\sqrt{\frac{1}{M-1} \sum (g_j^* - Mean)^2}$

4.4 Results Evaluation

Results of the experimental for the two tested datasets, D1 and D2, based on the suggested and compared methods are shown in [Tab. 5](#). The results are compared to GWO, GA, WOA, and PSO algorithms. The average error of (0.161956522) for D1 and of (0.027467811) is much better based on the suggested method. The average select error of (0.385714286) and (0.61875) for D1 and D2, respectively, show the performance of the SFS-Guided WOA algorithm. Average, best, worst fitness and standard deviation show the quality of the suggested method compared to other optimization techniques.

Table 5: Experimental results of the suggested and compared methods

Metric	Datasets	bSFS-guided WOA	bGWO	bGA	bWOA	bPSO
Average error	D1	0.161956522	0.165217391	0.163478261	0.165652174	0.162173913
	D2	0.027467811	0.027682403	0.032618026	0.028540773	0.030257511
Average select size	D1	0.385714286	0.5	0.553571429	0.542857143	0.542857143
	D2	0.61875	0.76875	0.7625	0.775	0.71875
Average fitness	D1	0.161956522	0.165217391	0.163478261	0.165652174	0.162173913
	D2	0.027467811	0.027682403	0.032618026	0.028540773	0.030257511
Best fitness	D1	0.006308489	0.009128255	0.006371262	0.012126419	0.010375499
	D2	0.002920846	0.002945634	0.010175433	0.00544218	0.006746602
Worst fitness	D1	0.134782609	0.152173913	0.152173913	0.139130435	0.139130435
	D2	0.021459227	0.021459227	0.021459227	0.025751073	0.025751073
Standard deviation	D1	0.173913043	0.173913043	0.173913043	0.2	0.182608696
	D2	0.034334764	0.034334764	0.055793991	0.0472103	0.055793991

ANOVA and Wilcoxon Signed Rank tests are performed to confirm the suggested method compared to other algorithms. [Tabs. 6](#) and [7](#) show the ANOVA test results of the tested algorithms

based on the first dataset (D1) and the second dataset (D2), respectively. The results indicated that the p-value is less than 0.05. Wilcoxon Signed-Rank test results based on ten runs for the first dataset (D1) and the second dataset (D2) using the suggested and compared algorithms are shown in [Tabs. 8 and 9](#), respectively. The statistical tests results confirm the performance of the SFS-Guided WOA algorithm for the EEG datasets.

Table 6: ANOVA test of the first dataset (D1)

	SS	DF	MS	F (DFn, DFd)	P value
Treatment (between columns)	0.00012	4	2.99E-05	F (4, 45) = 33.25	P < 0.0001
Residual (within columns)	4.05E-05	45	9E-07	-	-
Total	0.00016	49	-	-	-

Table 7: ANOVA test of the second dataset (D2)

	SS	DF	MS	F (DFn, DFd)	P value
Treatment (between columns)	0.0002	4	5.01E-05	F (4, 45) = 69.76	P < 0.0001
Residual (within columns)	3.23E-05	45	7.18E-07	-	-
Total	0.000233	49	-	-	-

Table 8: Wilcoxon signed rank test of the first dataset (D1)

	bSFS-Guided WOA	bGWO	bGA	bWOA	bPSO
Theoretical median	0	0	0	0	0
Actual median	0.162	0.1652	0.1635	0.1657	0.1622
Number of values	10	10	10	10	10
Wilcoxon signed rank test					
Sum of signed ranks (W)	55	55	55	55	55
Sum of positive ranks	55	55	55	55	55
Sum of negative ranks	0	0	0	0	0
P value (two tailed)	0.002	0.002	0.002	0.002	0.002
Exact or estimate?	Exact	Exact	Exact	Exact	Exact
P value summary	**	**	**	**	**
Significant (alpha = 0.05)?	Yes	Yes	Yes	Yes	Yes
How big is the discrepancy?					
Discrepancy	0.162	0.1652	0.1635	0.1657	0.1622

Table 9: Wilcoxon signed rank test of the first dataset (D1)

	bSFS-Guided WOA	bGWO	bGA	bWOA	bPSO
Theoretical median	0	0	0	0	0
Actual median	0.02747	0.02768	0.03262	0.02854	0.03026
Number of values	10	10	10	10	10
Wilcoxon signed rank test					
Sum of signed ranks (W)	55	55	55	55	55
Sum of positive ranks	55	55	55	55	55
Sum of negative ranks	0	0	0	0	0
P value (two tailed)	0.002	0.002	0.002	0.002	0.002
Exact or estimate?	Exact	Exact	Exact	Exact	Exact
P value summary	**	**	**	**	**
Significant (alpha = 0.05)?	Yes	Yes	Yes	Yes	Yes
How big is the discrepancy?					
Discrepancy	0.02747	0.02768	0.03262	0.02854	0.03026

The average error of the suggested (bSFS-Guided WOA) and compared algorithms (bPSO, bWOA, bGA and bGWO) over the two tested datasets (D1 and D2) is shown in Fig. 6. The figure indicates the performance of the suggested method over the tested datasets. Residual, Homoscedasticity, QQ plots and heat map of the suggested and compared algorithms over the first tested dataset (D1) and the second dataset (D2) are shown in Figs. 7 and 8, respectively.

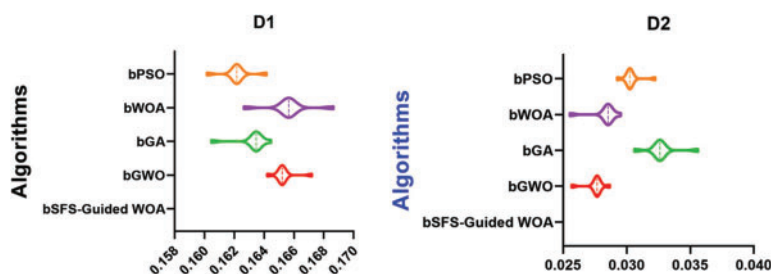


Figure 6: Suggested and compared algorithms average error over the two tested datasets (D1 and D2)

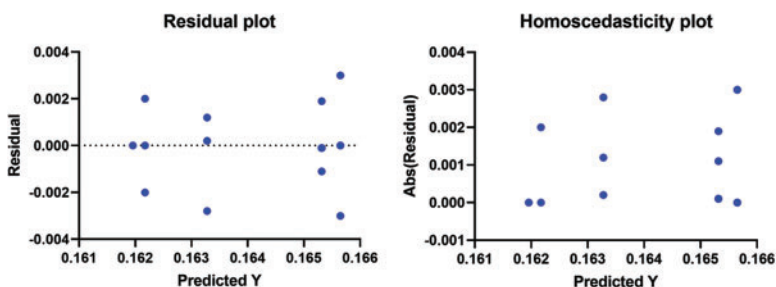


Figure 7: Continued

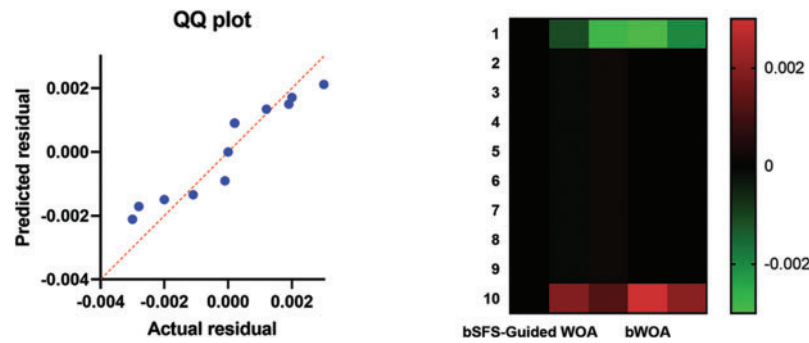


Figure 7: Residual, Homoscedasticity, QQ plots and heat map of the suggested and compared algorithms over the first tested dataset (D1)

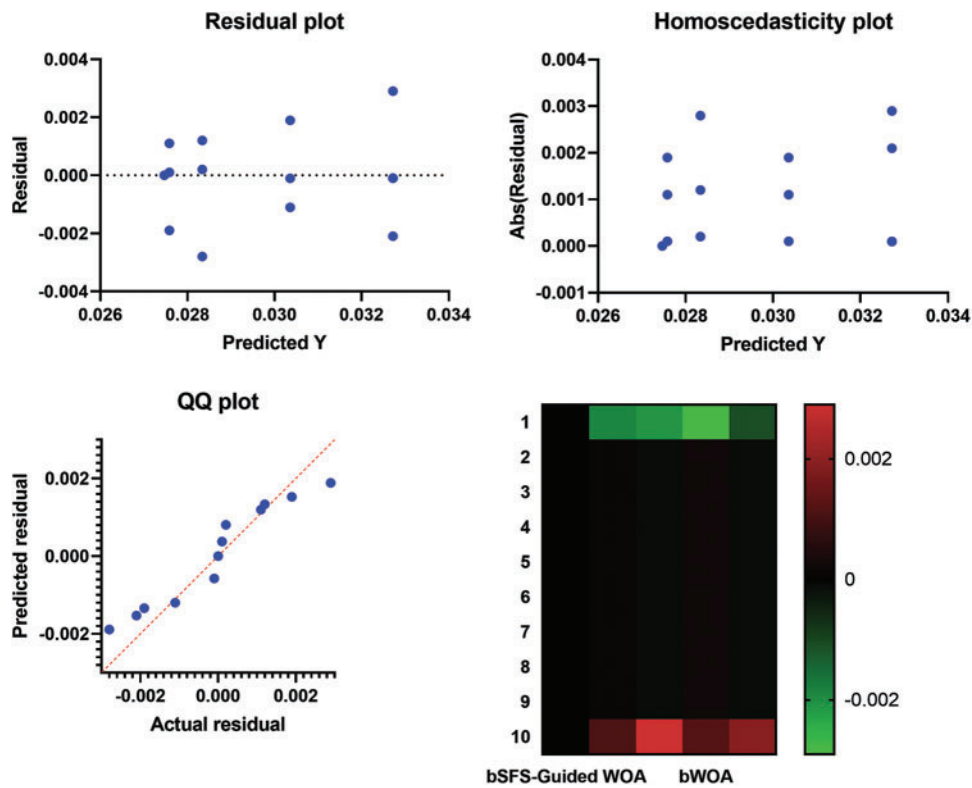


Figure 8: Residual, Homoscedasticity, QQ plots and heat map of the suggested and compared algorithms over the first tested dataset (D2)

5 Conclusion

In this work, the Guided Whale Optimization Method (Guided WOA) algorithm based on Stochastic Fractal Search (SFS) technique is used to evaluate the chosen subset of channels for EEG datasets. This method is used to select the optimum EEG channels for use in Brain-Computer Interfaces (BCIs). The (SFS-Guided WOA) algorithm is superior in terms of performance metrics, and statistical tests such as ANOVA and Wilcoxon rank-sum are used to demonstrate this. The results for the two tested datasets based on the suggested and compared methods (GWO, GA, WOA,

and PSO algorithms) show the quality of the recommended method. The average error and average select error confirm the performance of the SFS-Guided WOA algorithm. Other metrics, such as average, best, worst fitness and standard deviation, also show the quality of the suggested method compared to other optimization techniques. The average error of the presented (bSFS-Guided WOA) algorithm and compared algorithms (bPSO, bWOA, bGA and bGWO) indicates the performance of the recommended method over the tested datasets. Residual, Homoscedasticity, QQ plots and heat map of the suggested and compared algorithms are also tested over the two datasets. The recommended method in this work will be tested for other datasets in the future.

Acknowledgement: The authors thank Taif University Accessibility Center for the study participants. We deeply acknowledge Taif University for supporting this study through Taif University Researchers Supporting Project Number (TURSP-2020/150), Taif University, Taif, Saudi Arabia.

Funding Statement: Funding for this study is received from Taif University Researchers Supporting Project No. (Project No. TURSP-2020/150), Taif University, Taif, Saudi Arabia.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [2] N. Birbaumer, "Breaking the silence: Brain-computer interfaces (BCI) for communication and motor control," *Psychophysiology*, vol. 43, pp. 517–532, 2006.
- [3] J. S. Ebersole and T. A. Pedley, *Current Practice of Clinical Electroencephalography*. 3rd ed., Philadelphia: Lippincott Williams & Wilkins, pp. 72–99, 2003.
- [4] B. J. Baars and N. M. Gage, *Cognition, Brain, and Consciousness: Introduction to Cognitive Neuroscience*. 2nd ed., Elsevier Academic Press, 2010.
- [5] D. Garrett, D. A. Peterson, C. W. Anderson and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 141–144, 2003.
- [6] R. S. Fisher, "Therapeutic devices for epilepsy," *Annals of Neurology*, vol. 71, no. 2, pp. 157–168, 2012.
- [7] D. Chandler, J. Bisasky, J. L. V. M. Stanislaus and T. Mohsenin, "Real-time multichannel seizure detection and analysis hardware," *IEEE Biomedical Circuits and Systems Conf. (BioCAS)*, San Diego, CA, USA, 2011.
- [8] M. M. Fouad, A. I. El-Desouky, R. Al-Hajj and E.-S. M. El-Kenawy, "Dynamic group-based cooperative optimization algorithm," *IEEE Access*, vol. 8, pp. 148378–148403, 2020.
- [9] E. -S. El-Kenawy and M. Eid, "Hybrid gray wolf and particle swarm optimization for feature selection," *international journal of innovative computing*, *Information and Control*, vol. 16, no. 3, pp. 831–844, 2020.
- [10] M. M. Eid, E. -S. M. El-Kenawy and A. Ibrahim, "A binary sine cosine-modified whale optimization algorithm for feature selection," in *4th National Computing Colleges Conf. (NCCC 2021)*, Taif, Saudi Arabia, IEEE, pp. 1–6, 2021.
- [11] Z. Malki, E. S. Atlam, G. Dagnew, A. R. Alzighaibi, E. Ghada *et al.*, "Bidirectional residual LSTM-based human activity recognition," *Computer and Information Science*, vol. 13, no. 3, pp. 40–48, 2020.
- [12] H. Hashim, E. S. Atlam, M. Almaliki, R. El-Agamy, M. M. El-Sharkasy *et al.*, "Integrating data warehouse and machine learning to predict on COVID-19 pandemic empirical data," *Journal of Theoretical and Applied Information Technology*, vol. 99, no. 1, pp. 159–170, 2021.

- [13] M. Farsi, D. Hosahalli, B. R. Manjunatha, I. Gadb, E. S. Atlam *et al.*, “Parallel genetic algorithms for optimizing the SARIMA model for better forecasting of the NCDC weather data,” *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 1299–1316, 2021.
- [14] A. Ibrahim, A. Tharwat, T. Gaber and A. E. Hassanien, “Optimized superpixel and adaboost classifier for human thermal face recognition,” *Signal, Image and Video Processing*, vol. 12, pp. 711–719, 2018.
- [15] A. Ibrahim, S. Mohammed, H. A. Ali and S. E. Hussein, “Breast cancer segmentation from thermal images based on chaotic salp swarm algorithm,” *IEEE Access*, vol. 8, no. 1, pp. 122121–122134, 2020.
- [16] A. Ibrahim, H. A. Ali, M. M. Eid and E. -S. M. El-Kenawy, “Chaotic harris hawks optimization for unconstrained function optimization,” in *2020 16th Int. Computer Engineering Conf. (ICENCO)*, Cairo, Egypt, IEEE, pp. 153–158, 2020.
- [17] E. -S. M. El-Kenawy, S. Mirjalili, A. Ibrahim, M. Alrahmawy, M. El-Said *et al.*, “Advanced meta-heuristics, convolutional neural networks, and feature selectors for efficient COVID-19 x-ray chest image classification,” *IEEE Access*, vol. 9, pp. 36019–36037, 2021.
- [18] E. M. Hassib, A. I. El-Desouky, L. M. Labib and E. -S. M. T. El-Kenawy, “Woa + brnn: An imbalanced big data classification framework using whale optimization and deep neural network,” *Soft Computing*, vol. 24, pp. 5573–5592, 2020.
- [19] E. M. Hassib, A. I. El-Desouky, E. M. El-Kenawy and S. M. El-Ghamrawy, “An imbalanced big data mining framework for improving optimization algorithms performance,” *IEEE Access*, vol. 7, pp. 170774–170795, 2019.
- [20] E.-S. M. El-Kenawy, S. Mirjalili, S. S. M. Ghoneim, M. M. Eid, M. El-Said *et al.*, “Advanced ensemble model for solar radiation forecasting using sine cosine algorithm and newton’s laws,” *IEEE Access*, vol. 9, pp. 115750–115765, 2021.
- [21] A. Bohr and K. Memarzadeh, “The rise of artificial intelligence in healthcare applications,” *Artificial Intelligence in Healthcare*, vol. 2020, pp. 25–60, 2020.
- [22] J. B. Glattfelder, “The consciousness of reality,” in *Information—Consciousness—Reality*. The Frontiers Collection, Cham: Springer, pp. 515–595, 2019.
- [23] S. S. M. Ghoneim, T. A. Farrag, A. A. Rashed, E. -S. M. El-Kenawy and A. Ibrahim, “Adaptive dynamic meta-heuristics for feature selection and classification in diagnostic accuracy of transformer faults,” *IEEE Access*, vol. 9, pp. 78324–78340, 2021.
- [24] A. Ibrahim, M. Noshay, H. A. Ali and M. Badawy, “PAPSO: A power-aware VM placement technique based on particle swarm optimization,” *IEEE Access*, vol. 8, pp. 81747–81764, 2020.
- [25] E. -S. M. El-Kenawy, M. M. Eid, M. Saber and A. Ibrahim, “Mbgwo-sFS: Modified binary grey wolf optimizer based on stochastic fractal search for feature selection,” *IEEE Access*, vol. 8, no. 1, pp. 107635–107649, 2020.
- [26] A. Ibrahim, S. Mirjalili, M. El-Said, S. S. M. Ghoneim, M. Al-Harhi *et al.*, “Wind speed ensemble forecasting based on deep learning using adaptive dynamic optimization algorithm,” *IEEE Access*, vol. 9, pp. 125787–125804, 2021.
- [27] E. S. M. El-Kenawy, A. Ibrahim, S. Mirjalili, M. M. Eid and S. E. Hussein, “Novel feature selection and voting classifier algorithms for COVID-19 classification in CT images,” *IEEE Access*, vol. 8, pp. 179317–179335, 2020.
- [28] T. Michael, M. K.-Robert, A. Ad, B. Niels, B. Christoph *et al.*, “Review of the BCI competition IV,” *Frontiers in Neuroscience*, vol. 6, no. 55, pp. 1–31, 2012.