Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2023) 31: 771 – 791

doi:10.55730/1300-0632.4017

Cognitive load detection using Ci-SSA for EEG signal decomposition and nature-inspired feature selection

Jammisetty YEDUKONDALU, Lakhan Dev SHARMA*
School of Electronics Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India

Received: 30.01.2023 • **Accepted/Published Online:** 03.05.2023 • **Final Version:** 29.09.2023

Abstract: Cognitive load detection is eminent during the mental assignment of neural activity because it indicates how the brain reacts to stimuli. The level of cognitive load experienced during mental arithmetic tasks can be determined using an electroencephalogram (EEG). The EEG data were collected from publicly available datasets, namely, mental arithmetic task (MAT) and simultaneous task workload (STEW). The first phase comprises decomposing the electroencephalogram (EEG) signal into intrinsic mode functions (IMFs) using circulant singular spectrum analysis (Ci-SSA). In the second phase, entropy-based features were evaluated using IMFs. After that, the extracted features were fed to nature-inspired feature selection algorithms: genetic algorithm (GA), binary particle swarm optimization (BPSO), particle swarm optimization (PSO), binary bat algorithm (BBA), and binary dragonfly algorithm (BDA) for optimal selection of features by using machine learning (ML) techniques: K-nearest neighbor (KNN), support vector machine (SVM) to analyse the classification accuracy (Ac), sensitivity (Se), specificity (Sp), precision (Pr), and F-score with 10-fold cross-validation in the third phase. The highest classification Ac, Se, Sp, Pr, and F-score of the MAT dataset were 97.30%, 0.98, 0.97, and 97.40% from multileads, and 96.20%, 0.96, 0.94, and 96.70% from a single lead (F4) of EEG, respectively. However, we achieved 97.98%, 0.98, 0.98, 0.97, and 98.1% values from multi-leads and 96.67%, 0.96, 0.97, 0.95, and 96.90% from a single-lead STEW dataset. When compared to previous state-of-the-art methods, the proposed method (Ci-SSA+BDA+KNN) has proven to be more successful.

Key words: Ci-SSA, cognitive load, EEG, machine learning, nature-inspired algorithms

1. Introduction

1.1. Motivation

Different tasks necessitate various mental resources, including task-related knowledge, attention, working memory, and decision-making. Our brains, on the other hand, have limited resources for processing and combining data. The term cognitive load refers to the amount of work that these limited resources have to do. In neuroscience, continual mental workload revelation has been demonstrated to impact brain dynamics [1]. This cognitive load estimation has the main problem for the researchers to explore different physiological parameters, like heart rate [2], facial expression, pupil dilation [3], and ocular parameters. These parameters would be affected by the visceral motor system or autonomic nervous system. Due to the revelation of stress events, the balance between the sympathetic and parasympathetic nervous systems may be disrupted. The sympathetic nervous system would be activated excessively, while the parasympathetic nervous system would be inhibited.

*Correspondence: devsharmalakhan@gmail.com

Respiration, heart rate variability (HRV), skin conductance, and blood pressure are some of the physiological changes that might occur due to this [4, 5]. The aim of cognitive load detection using EEG signals is to analyse the brain activity of a person to measure the level of cognitive load they are experiencing while performing a particular task. By detecting changes in EEG signals, it is possible to determine the cognitive load of a person in real-time. Particularly, in healthcare, it can be used to monitor cognitive load in patients with neurological disorders such as dementia, traumatic brain injury, etc.

1.2. Related work

Electroencephalography (EEG), which has been effectively utilized to study and quantify many mental functions and workloads [6] in cognitive science and psychology, can be used to monitor and assess cognition. The use of noninvasive EEG readings to study functional changes in the brain is critical in this procedure. Feature extraction, feature selection, and classification [7] are the main steps required for extracting the aspired information from preprocessing of EEG data. Several EEG-based categorization techniques have been presented in the literature [8, 9], and all of them have had a significant influence on the field of medicine. A mental workload detection technique based on the fractal dimension spectrum was developed by Wang et al. [10] to monitor the brain's response to a mental arithmetic task. A mental load characterization technique was developed by Zarjam et al. [11] to assess brain activity across a range of mental arithmetic loads, from low to extremely high. However, the accuracy of current noninvasive stress detection methods needs to be enhanced. To detect stress, a GA hybrid was utilized to select stress features, determine the classifier type, and optimize the classifier's parameters [12]. This resulted in an increase in the stress recognition rate to 89%. Sharma et al. [13] proposed a Bayesian optimized-KNN classifier to identify cognitive load during mental arithmetic using EEG signals. In preprocessing phase, a Savitzky-Golay filter and SWT were employed. Sample entropy features were extracted resulting in an accuracy of 96%. Hou et al. [14] proposed a wavelet packet transform (WPT)-Random Forest (RF) for four class human emotion using differential entropy (DE) features. They achieved an average classification accuracy of 87.3%. Dovile et al. [15] proposed convolutional neural network (CNN) models: EEGNet and DeepConvNet for emotion types and facial inversion stimuli.

Murali et al. [16] introduced a general mixture model (GMM) for emotion recognition and extracted the psychological status of mentally impaired persons using EEG signals. Dursun et al. [17] proposed an artificial neural network (ANN) for the automatic elimination of EOG artifacts from sleep EEG signals. It is an important clinical tool for automatic sleep staging, cognitive load, and other EEG applications. Yalcin et al. [18] proposed a particle swarm optimization (PSO)-based ANN (PSOANN) that was used to classify the epileptic seizures and the diagnosis of epilepsy and achieved an accuracy of 99.45%. Maja et al. [19] proposed empirical mode decomposition (EMD), principal component analysis (PCA), and SVM for feature extraction, selection, and classification, respectively. They achieved an accuracy of 91.45%. In their research, Shon et al. [20] proposed a classifier for detecting emotional stress states using a combination of genetic algorithm (GA)-based KNN and principal component analysis (PCA)-KNN. The accuracy achieved was 71.76% and 65.3% for GA-KNN and PCA-KNN, respectively. Acharya et al. [21] introduced a technique for the automatic classification of epileptic EEG activities in the wavelet framework using principal component analysis and a Gaussian mixture model (GMM) classifier. The proposed approach achieved an accuracy of 99% and can classify EEG segments with a clinically acceptable accuracy using fewer features, which results in less computational cost. Huseyin et al. [22] presented a classification method for obstructive sleep apnea (OSA) based on common features from the

time domain, frequency domain, and nonlinear calculations of heart rate variability analysis, along with feature selection using correlation matrices (CMs). For epileptic seizure detection in EEG, Hassan et al. [23] proposed a method that combines empirical mode decomposition, a mutual information-based best individual feature (MIBIF) selection algorithm, and a multilayer perceptron neural network. Finally, Kai et al. [24] suggested a statistical technique for subject-specific time segment and temporal frequency band selection based on the mutual information between the spatial-temporal patterns from the EEG signals and the corresponding neuronal activities.

Feature selection using nature-inspired algorithms can be considered an optimization problem because it intends to reduce the number of selected features while improving classification accuracy or lowering prediction error [25]. Due to the enormous search space and complex interactions among features, many researchers have considered feature selection to be an optimizer. Based on the assessment criteria of the chosen subset, feature selection techniques are divided into two categories: filters and wrappers [26]. While wrapper approaches analyse the feature subset using a learning algorithm such as classification, filters rely on the data itself to evaluate the feature subset using defined methods. Many GA techniques have been presented to address feature selection [27]. Numerous feature selection techniques have used metaheuristic algorithms to find the best subset in the literature. As a wrapper FS technique, it uses an ant lion optimizer (ALO) [28] that imitates how antlions pursue their prey. A novel technique called grey wolf optimizer (GWO) has been effectively used to resolve feature selection issues in [29]. In [30], a wrapper-based moth-flame optimization method has been proposed for feature selection problems with validated results from 29 benchmark functions and 7 real engineering problems. Based on the literature review, researchers used traditional decomposition methods and feature selection methods. They attained low accuracy for different EEG applications. Thus, in this article, we have employed Ci-SSA for EEG signal decomposition, and entropy-based feature extraction. For feature selection, nature-inspired algorithms were used. Finally, KNN and SVM were employed for classification with different distance/kernel functions.

In this work, Ci-SSA was used to decompose the EEG signals and extract the features from IMFs. It is an automatic process for extracting the signal corresponding with a specific frequency that has been previously specified. This variant of singular spectrum analysis (SSA) differs from earlier versions in that it is applied to any time-series signal and is based on circulant matrices that, once the user has determined the frequency from interest, automatically match this frequency of specific principal components. Ci-SSA has several advantages that make it a useful tool for EEG signal analysis. Some of these advantages include:

- It is a nonparametric approach that makes no data distribution assumptions. EEG signals vary widely among individuals and experimental conditions.
- It is a data-driven method that can detect the EEG signal's underlying structure.
- It is quick and efficient for large datasets. This is important for real-time EEG signal analysis.
- Finally, it provides interpretable results that can be easily visualized and understood. This is particularly useful for identifying EEG signal features.

Two nature-inspired algorithms—binary bat algorithm (BBA) and binary dragonfly algorithm (BDA)—techniques were used in this paper for feature selection, and the study was designed for a continuous optimization problem. BDA has several advantages when compared to the aforementioned techniques. BDA has a fast convergence rate, it can quickly converge to a good solution. It is effective for feature selection because of its

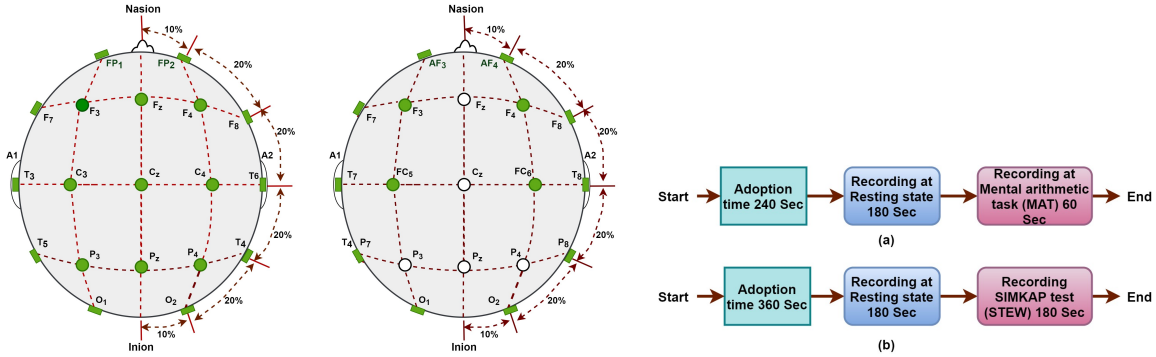
ability to balance exploration and exploitation during the search process. Exploration refers to the ability to search the entire search space to discover new solutions, while exploitation refers to the ability to focus on the most promising regions of the search space to refine existing solutions, and BDA incorporates multiple search strategies including random search, levy flight, and spiral search. Therefore, BDA was employed for feature selection throughout this work. Furthermore, we utilized KNN and SVM classifiers to evaluate the performance of the individual classifier with the reduced feature set by BDA and examine the impact of feature selection on the accuracy of the model. The following is a summary of the paper: Section 2 describes the datasets utilized in this study, Section 3 the methods for this approach, Section 4 the results, Section 5 discussion, and Section 6 conclusion.

2. Datasets

In this article, two publicly available EEG datasets, MAT and STEW, were employed [31, 32].

MAT: MAT data are from 36 healthy people between the ages of 18 and 26. The subjects' EEG signals are recorded using the 10–20 system during and before the performing of mental arithmetic tasks. Power line notch (50 Hz), low-pass (0.5 Hz), high-pass (45 Hz) filters, and 45 Hz cut-off frequencies were utilized to reduce artifacts, and the EEG signals were sampled at 500 Hz, and 20 channels were used as shown in Figure 1.

STEW: The raw EEG data from 48 subjects who took part in a multitasking workload experiment using the SIMKAP multitasking test is included in the STEW dataset. Before the test, the subjects' resting brain activities were recorded and also included. The data was collected using the Emotiv EPOC device, which has a sampling frequency of 128 Hz and 14 channels and records EEG for 2.5 min in each case as depicted in Figure 1.



(a) MAT (b) STEW (c) Experimental protocol

Figure 1. 10-20 EEG electrodes' placement (used leads are colored green) and their experiment protocol.

3. Methods

In this manuscript, Ci-SSA was used to decompose the filtered EEG signal into IMFs. Next, five entropy-based features were extracted from IMFs. Feature selection was done using nature-inspired algorithms, BBA and BDA, to select the significant features from the data. Furthermore, machine learning models, KNN and SVM, were employed to classify the performance metrics. Figure 2 shows the proposed diagram of our methodology. The following sections go through each step in detail.

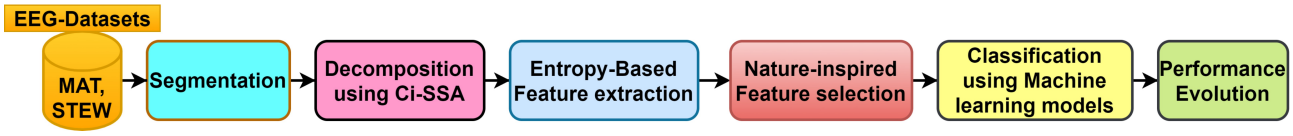


Figure 2. Proposed method diagram.

3.1. Circulant singular spectrum analysis

SSA can be applied to any time series, including nonstationary time series because it is a model-free method. When employing SSA, it is common practice to extract the principal components from the trajectory matrix and then identify the frequencies that correspond to them by, among other factors, looking at their anticipated periodogram and frequency response [33]. The usage of these circulant matrices has a substantial advantage as their eigenvalues and eigenvectors have a closed shape, even though there are fast computation procedures employing Toeplitz matrices' eigenvalues and eigenvectors. Bogalo et al. [7] introduced the CiSSA approach,

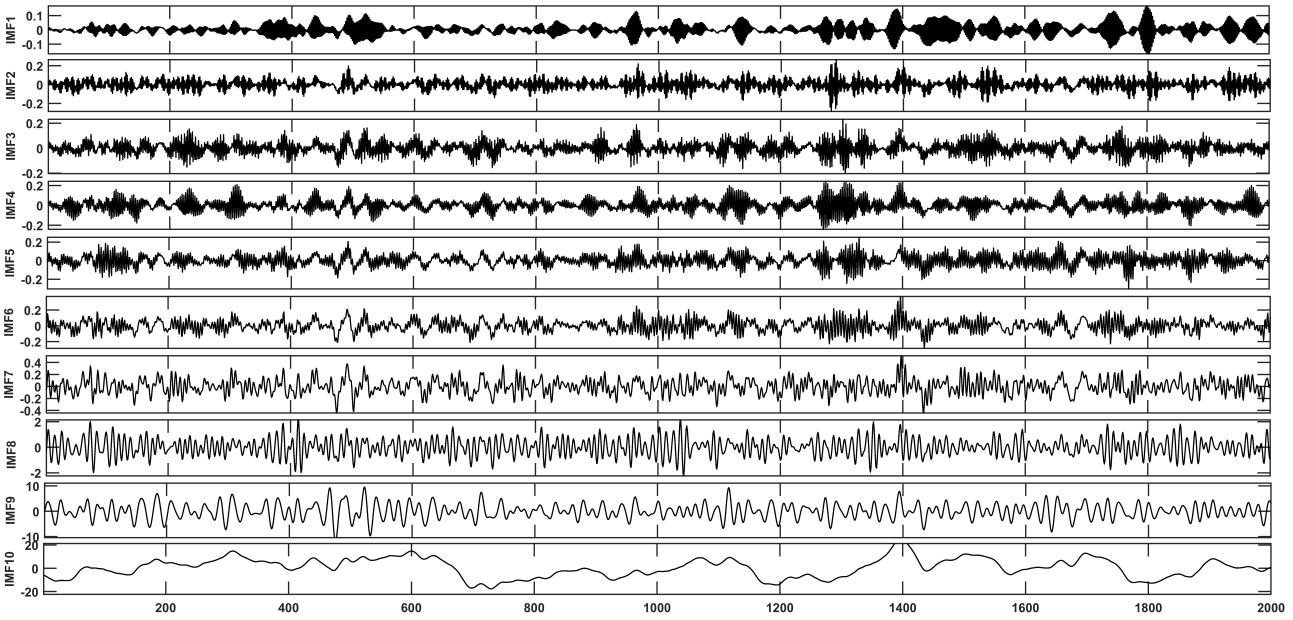


Figure 3. Decomposition of EEG signal using Ci-SSA.

which is a nonparametric signal extraction method. The four steps of CiSSA are grouping, diagonal averaging, decomposition, and embedding. CiSSA is feasibly executed in the four steps below:

step-1: Embedding In this step, to construct an $R \times S$ trajectory matrix A, S = t-R+1, from the original time series as described in the following:

$$A = (a_1|...|a_N) = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_S \\ a_2 & a_3 & a_4 & \cdots & a_{S+1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_R & a_{R+1} & a_{R+2} & \cdots & a_t \end{pmatrix}, \quad (1)$$

where $a_i = (a_i, ...a_{i+R+1})'$ specifies the vector $R \times 1$ with origin at time.

step-2: Decomposition To construct the circulant matrix C_M, which has the following components:

$$\widehat{M_n} = \frac{R-n}{R}\widehat{n}_n + \frac{n}{R}\widehat{n}_{R-n}, n = 0, ...R-1, \tag{2}$$

the eigenvalues of $\widehat{\mu}_b$ of C_M are calculated as follows:

$$\mu_{R,b} = f\left(\frac{b-1}{R}\right). \tag{3}$$

The k^{th} eigenvalue and the corresponding eigenvector are collaborated with frequency $V_b = \dfrac{b-1}{R}$, b=1,..R.

step-3: Grouping We obtain $\widehat{\mu}_b = \mu_R + 2 - b$ when we specify the analogy of power spectral density (PSD). Then, because their associated eigenvectors are complex, these are complex conjugate pairs, $j_b = j^*_{R+2-b}$, where j* denotes the complex conjugate of j. In order to calculate related elements, we turn them into pairs of real eigenvectors. To make the fundamental matrices, we will need to make two-element groups $A_B = b, R+2-b$ for b = 2,... N with A1=1 and $A\dfrac{R}{2}+1 = \dfrac{R}{2}+1$ if R is even. Then, using the frequency D_{A_b}, the elementary matrix is computed as the total of two elementary matrices D_b and D_{R+2b}, those are connected to the eigenvalues $\widehat{\mu}_b$ and $\widehat{\mu}_{R+2-b}$ and the frequency $V_b = \dfrac{b-1}{R}$,

$$\begin{aligned} D_{A_b} &= D_b + D_{R+2-b}, \\ D_{A_b} &= (j_b j_b^H D + j_{R+2-b} j_{R+2-b}^H)D, \\ D_{A_b} &= 2(Y_{j_b} Y_{j_b}^{'} + Z_{j_b} Z_{j_b}^{'})D, \\ D_{A_b} &= (j_b j_b^H + j_b^* j_b^{'})D, \end{aligned} \tag{4}$$

where Y_{j_b} is real part of j_b, imaginary part is Z_{j_b}, and j^H denotes the conjugate transpose of j.

$$j_b = R^{-1/2}(j_{b,1}...j_{b,R}). \tag{5}$$

Step-4: Reconstruction Each matrix D_{Oi} is averaged diagonally to produce the newest time series of length t equal to the preceding one. It is equivalent to calculating the mean of the nondiagonal elements of D_{Oi} or hankelizing this matrix, which is denoted by the operator H $(\cdot)$. The Vautard-Ghil replacement, also known as the Toeplitz SSA, conducts orthogonal diagonalization from a different matrix $G_t = (G_{oi})$ under the premise that it is stationary and has a mean value of a nonzero average.

$$G_{oi} = \frac{1}{t-|o-i|}\sum_{n=1}^{t-|o-i|} D_n D_{n+|o-i|}. \tag{6}$$

As the original series sample lagged the variance-covariance matrix, the G_t is a symmetric Toeplitz matrix. The b^{th} eigentriple is named after the set (t_b, V_b, j_b). The EEG data from MAT and STEW datasets were segmented using a 2000 and 512 window length (where 500 Hz and 128 Hz are the sampling frequencies) and 10-level decomposition was done using Ci-SSA respectively as shown in Figure 3.

3.2. Feature extraction

For cognitive load detection, this work employs entropy-based features [34]. Entropy is a metric of complexity that could be used to describe the dynamics of the brain [35].

3.2.1. Shannon's entropy (SE)

SE illustrates average unpredictability in a random variable and is equivalent to its information content. For a given discrete distribution $P = p_1, \cdots, p_N$, the associated SE is calculated [28].

$$SE[P] = -\sum_{i=1}^{N} p_i \ln(p_i). \tag{7}$$

3.2.2. Approximate entropy (AE)

The term AE is taken from the term Kolmogorov entropy and it is used to define the randomness or unpredictable nature of a finite-length signal. When the phase space embedding dimension (M) is increased from M to M + 1, its arithmetic involves embedding the signal in phase space and computing the rate at which the number of phase space patterns inside a given value (r) rises. The AE is obtained by:

$$AE(M, r, N) = \Phi(r) - \Phi^{M+1}(r). \tag{8}$$

AE can be used as the evaluation criteria of cognitive workload, which could be applied in the ergonomics estimation of the human-computer interaction field [36].

3.2.3. Differential entropy (DE)

Differential entropy is a constant random variable that measures the uncertainty in EEG time series analysis. It is also related to the minimum description length. A mathematical expression is as follows:

$$H(T) = -\int g(t)[\log[g(t)]\mathrm{d}t]. \tag{9}$$

In this case, T is a random variable, and its probability density is given by g(t). Consequently, for every time series T that has a Gaussian distribution, $N(\mu, \sigma^2)$, the DE is written as [37]:

$$H(T) = \frac{1}{2}\log(2\pi e\sigma^2). \tag{10}$$

3.2.4. Information potential (IP)

Renyi's quadratic entropy is estimated nonparametrically by IP. For x_n, a random variable, it was measured using this equation [38]:

$$IP(X) = \frac{1}{N^2}\sum_{a=1}^{N}\sum_{b=1}^{N} k_\sigma(x_b - x_a). \tag{11}$$

In the above equation, x_a, x_b are a^{th}, b^{th} data set samples. N is all samples and $k_\sigma(x_b - x_a)$ expresses a Gaussian kernel function. The sample's fundamental density mean is invariant over time. When using IP for supervised learning, the average of every error signal does not make zero, as most applications require.

3.2.5. Fuzzy entropy (FE)

The degree of similarity among two vectors of length r from an N sample time series is measured by fuzzy entropy [39]. It is calculated based on:

$$FE(r,s,t,N) = \ln[\varphi^r(s,t)] - \ln[\varphi^{s+1}(s,t)]. \tag{12}$$

The function $\varphi^r(s,t)$ is expressed as:

$$\varphi^r(s,t) = \frac{1}{N-r}\sum_{k=1}^{N-r}\frac{1}{N-r-1}\sum_{j=1,j\neq k}^{N-r}(l_{kj}^r). \tag{13}$$

In this case, l_{kj} represents the degree of similarity between any two vectors, while s (gradient), t (width), and r (embedding dimension) control fuzzy boundaries.

3.3. Feature selection using nature-inspired algorithms

If dealing with large datasets containing more features, dimensionality compression can be used as an important step to decrease the number of features. In making a classifier, more work is focused on selecting suitable features and resolving to encode them. In this selective problem of feature selection, we use nature-inspired algorithms, to find an optimal feature subset (SFeat).

3.3.1. Binary bat algorithm (BBA)

A bat uses sound to navigate obstacles, find prey, and find its nest in the dark by following an echo that is reflected off the environment's objects. The bat algorithm was devised in response to this behaviour of bats. This approach employs artificial bats to determine the best solution to an objective function. The bat algorithm, introduced in [40], is discretized in the BBA. Any bat may navigate the nodes and corners of the lattice in the n-dimensional Boolean hypercube which is the search space for the BBA. Each feature is characterised by a binary bat position vector, which is then used to choose features. As a result, a feature is absent when the value is 0 and present when the value is 1. The bat population begins with the position z_i, velocity v_i, and pulse frequency f_p for a specific function f(z).

3.3.2. Binary dragonfly algorithm (BDA)

The dragonfly algorithm (DA) was modelled after dragonflies, as indicated by its name [41]. This approach may be thought of as a swarm intelligence technique to estimate the global optimum of a certain optimal solution. The following lists the mathematical models that simulate dragonfly swarming behaviour:

- To avoid colliding with nearby neighbours, individuals use a strategy known as separation. This behaviour is represented mathematically as Eq. (14).

$$Y_n = -\sum_{r=1}^{S} Z - Z_r, \tag{14}$$

  where Z is the location of the current individual, Zr is that of the Z position's r^{th} neighbour, and S is the size of the neighbourhood.

- Alignment shows how closely an individual's velocity matches that of nearby individuals. The mathematical model for this behaviour is Eq. (15).

$$A_n = \frac{\sum_{r=1}^{S} V_r}{S}, \tag{15}$$

where S is the size of the neighbourhood and V_r is the velocity of the r^{th} neighbour.

- Cohesion is the propensity of individuals to gravitate toward the centre of their group. The mathematical model for this behaviour is Eq. (16).

$$C_n = \frac{\sum_{r=1}^{S} Z_r}{S} - Z. \tag{16}$$

- Other important survival behaviours include attraction to food sources and escape from enemies. Eq. (17) models the attraction to food.

$$F_n = Z^+ - Z, \tag{17}$$

where Z is the current individual position and Z^+ is the food source position.

$$E_n = Z^- + Z, \tag{18}$$

where Z^- is the enemy's position.

DA uses the step vector and position vector to tackle optimization issues. They are defined as two vectors. The following defines a step vector:

$$\Delta Z_{t+1} = (yY_n + aA_n + cC_n + fF_n + eE_n) + w\Delta Z_t, \tag{19}$$

where y stands for separation weight, a for alignment, c for cohesion, f for food factor, e for enemy factor, and w for inertia weight. The step vector is added to the previous position to update the position of dragonflies in a continuous search space. However, the following equations must be used in a binary search space:

$$Z_{t+1} = \begin{cases} \neg Z_t, z < T(\Delta z_{t+1}) \\ Z_t, z \geq T(\Delta z_{t+1}) \end{cases}. \tag{20}$$

$T(\Delta z_{t+1})$ is determined using the formula in Eq. (21).

$$T(\Delta z) = \left| \frac{\Delta z}{\sqrt{\Delta z^2 + 1}} \right|. \tag{21}$$

The binary 0 and 1 values are the only possible solutions to the issue of finite state or FS. Therefore, it is appropriate to apply the binary form of the DA method to resolve this issue. In this work, the solution to the issue is represented by a vector of zeros and ones, where a zero means that the relevant feature is not picked and a one means that this feature is selected. The total number of features in the original dataset is represented by the length of the solution vector. Algorithm 1 is the whole automated procedure and pseudocode of BDA. The objective function is shown in Eq. (22).

$$fitness = \alpha\gamma_E(H) + \left| \frac{E}{T} \right|, \tag{22}$$

Algorithm 1 BDA pseudocode

Create the population Z_n with the values n = 1, 2,..., m.
Set up ΔZ_n (n = 1, 2,..., m).
While t < Maximum iteration.
Calculate every dragonfly.
Update F and E.
Update the w, s, a, c, f, and e coefficients.
Evaluate Y, A, C, F, and E (using Eq. (14) to (18)).
Using Eq. (19) to update step vectors.
Using Eq. (21) to Compute $(T\Delta z)$.
Using Eq. (20) to update Z_{t+1}.
end while

where $\gamma_E(H)$ stands for a certain classifier's classification error rate. Additionally, H is the cardinality of the chosen subset, T is the total number of features in the dataset, α and β are two factors relating to the significance of subset length and classification quality, and $\alpha \in [0,1]$ and $\beta = (1-\alpha)$.

3.3.3. Classification using KNN and SVM

The procedure of classifying a given set of data into classes is known as classification, and both structured and unstructured data can be used with it. In this work, two supervised ML techniques, such as KNN and SVM, were employed to classify the parameter metrics.

KNN: It is a supervised learning algorithm that can be used for both classification and regression tasks. In this algorithm, a data point is classified by finding its k nearest neighbors in the training data, where k is a user-defined hyperparameter. The predicted class of the data point is then determined by a majority vote among the k nearest neighbors. The choice of k is important in KNN, as it can affect the performance of the classifier. A smaller value of k can make the classifier more sensitive to noise in the data, while a larger value of k can make the classifier more biased towards the majority class. Thus, in this work, The KNN was implemented by the distance functions namely, Euclidean, City Block, and Minkowski, and varies the $k \in [3,5,7,9]$.

SVM: It is also a supervised machine learning algorithm that can be used for classification and regression analysis. It finds the best decision boundary that separates data points into different classes by maximizing the margin between the classes. SVM can handle both linear and nonlinear classification tasks by transforming the input data into a higher-dimensional space using a kernel function. Therefore, in this work, the features were mapped using kernel functions namely, radial basis function (RBF), linear, and polynomial. The kernel scale (σ) parameter value was changed from 0.5 to 2.5 at the interval of 0.5.

3.3.4. Performance metrics

In this work, five performance metrics were used to assess the efficacy of classifiers: precision (Pr), F-score (%), and accuracy (Ac%) are three of the standard metrics, while specificity (Sp) and sensitivity (Se) are two of the most commonly used clinical measures. These are formulated as:

- Ac (%) $= \dfrac{TP+TN}{TP+FP+TN+FN} \times 100,$

- $\text{Pr} = \dfrac{TP}{TP+FP}$,
- $\text{F-score}\ (\%) = 2 \times \dfrac{PRE \times recall}{PRE + recall} \times 100$,
- $\text{Se} = \dfrac{TP}{TP+FN}$,
- $\text{Sp} = \dfrac{TN}{TN+FP}$.

4. Results

In our work, we have tested two ML techniques using nature-inspired algorithms for the classification of cognitive load detection. Finally, the performance is analysed in terms of classification parameters including the best optimal solution curve, confusion matrix, and ROC curve. Two models have been used to examine the planned work: (i) multilead approach (MLA), in which features extracted from twenty EEG leads from five brain lobes were ranked and employed and (ii) single-lead approach (SLA), which employed only features extracted from the lead F4 EEG signal. The following subsections describe the MLA and SLA experimental results of two datasets.

4.1. MLA

In the MLA, we analysed the classified parameters on the MAT dataset from PhysioBank using nature-inspired algorithms and classifiers KNN and SVM. Tables 1 and 2 show the results of the BBA-KNN and BDA-KNN classifiers using distance functions with different K values. The highest accuracy of the MAT dataset, 97.30%, was obtained in the BDA-KNN method at a Euclidean function with a K value of 5. In addition, the highest accuracy of the STEW dataset, 97.98%, was obtained in BDA-KNN at a Euclidean function. Tables 3 and 4 show the results of the BBA-SVM and BDA-SVM classifiers using Kernel functions on kernel scale (σ). The highest accuracy of the MAT dataset, 96.65%, was obtained in the BDA-SVM method at RBF kernel function with σ of 2. The highest accuracy of the STEW dataset, 95.55%, was obtained in BDA-SVM. The BDA-KNN method has the greatest accuracy at 97.30%, 97.98%, Se, Sp, Pr, and F-score at a Euclidean function with k=5. KNN has a better potential for cognitive load detection compared with SVM in our proposed algorithms. When compared to MAT dataset, the STEW dataset achieved better results. Thus, the optimal solution curve, confusion matrix, and ROC of MLA: BDA-KNN of the STEW dataset is shown in Figure 4.

Table 1. MLA: MAT dataset results of nature-inspired feature selection methods using KNN classifier.

Feature selection method	Parameters/ distance function	K=3			K=5			K=7			K=9		
		Euclidean	City block	Minkowski	Euclidean	City block	Minkowski	Euclidean	City block	Minkowski	Euclidean	City block	Minkowski
BBA	Ac %	95.00	95.74	95.18	95.65	96.20	94.44	92.77	93.06	92.31	92.13	93.24	91.20
	Se	0.94	0.95	0.95	0.94	0.96	0.92	0.91	0.91	0.91	0.93	0.93	0.93
	Sp	0.95	0.96	0.95	0.96	0.96	0.94	0.94	0.95	0.94	0.93	0.94	0.92
	Pr	0.95	0.96	0.95	0.96	0.94	0.94	0.94	0.93	0.92	0.93	0.93	0.91
	F-score%	94.97	95.69	95.18	95.60	96.70	94.37	92.88	93.05	92.26	0.92	93.26	91.20
BDA	Ac %	95.38	93.16	93.26	**97.30**	96.12	93.53	94.25	92.03	91.76	93.50	92.00	90.87
	Se	0.95	0.93	0.93	**0.98**	0.96	0.92	0.93	0.91	0.91	0.94	0.91	0.89
	Sp	0.96	0.94	0.94	**0.98**	0.96	0.95	0.95	0.94	0.93	0.94	0.93	0.92
	Pr	0.95	0.93	0.93	**0.97**	0.95	0.92	0.93	0.91	0.91	0.94	0.91	0.89
	F-score%	95.31	93.21	93.40	**97.40**	96.30	93.65	94.28	92.28	91.80	93.64	92.35	91.26

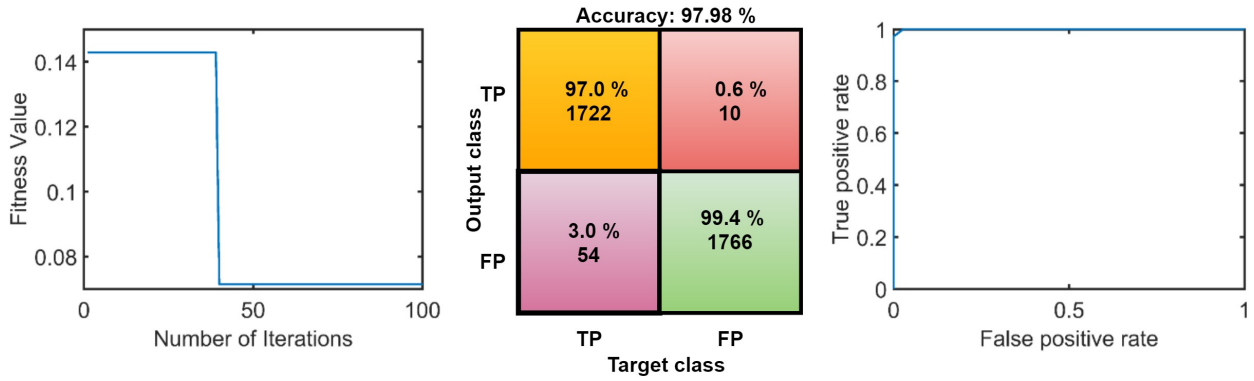


Figure 4. MLA: STEW dataset optimal solution curve, confusion matrix, and ROC.

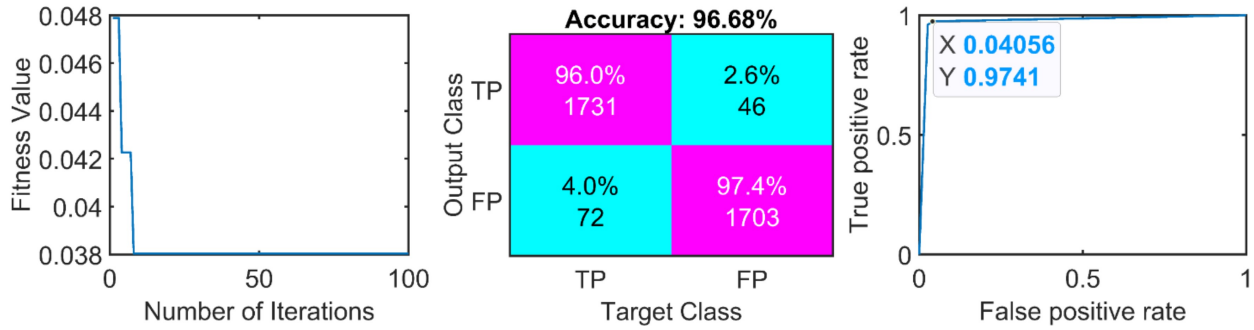


Figure 5. SLA: STEW dataset optimal solution curve, confusion matrix, and ROC.

Table 2. MLA: STEW dataset results of nature-inspired feature selection methods using KNN classifier.

Feature selection method	Parameters/ distance function	K=3			K=5			K=7			K=9		
		Euclidean	City block	Minkowski	Euclidean	City block	Minkowski	Euclidean	City block	Minkowski	Euclidean	City block	Minkowski
BBA	Ac %	94.46	95.40	94.84	96.40	94.25	94.43	93.78	95.55	94.56	93.64	95.94	94.74
	Se	0.94	0.93	0.93	0.96	0.93	0.93	0.92	0.94	0.94	0.92	0.94	0.93
	Sp	0.95	0.95	0.94	0.97	0.94	0.94	0.94	0.95	0.95	0.93	0.96	0.94
	Pr	0.94	0.94	0.93	0.95	0.93	0.94	0.93	0.94	0.94	0.93	0.94	0.93
	F-score%	94.59	95.67	94.90	96.60	94.38	94.76	93.88	95.70	94.65	93.82	96.04	94.93
BDA	Ac %	96.25	96.07	96.10	**97.98**	96.87	96.36	95.73	94.18	93.95	93.44	93.15	93.58
	Se	0.95	0.93	0.92	**0.95**	0.94	0.92	0.94	0.93	0.93	0.93	0.92	0.94
	Sp	0.96	0.96	0.95	**0.98**	0.97	0.96	0.96	0.94	0.94	0.94	0.93	0.94
	Pr	0.93	0.94	0.93	**0.97**	0.95	0.93	0.95	0.93	0.93	0.93	0.93	0.94
	F-score%	96.11	95.82	96.00	**98.00**	96.99	96.45	95.80	94.33	94.00	93.68	93.44	93.67

Table 3. MLA: MAT dataset results of nature-inspired feature selection methods using SVM classifier.

Feature selection method	Parameters/ Kernel function	σ=0.5			σ=1			σ=1.5			σ=2			σ=2.5		
		RBF	Linear	Polynomial	RBF	Linear	Polynomial	RBF	Linear	Polynomial	RBF	Linear	Polynomial	RBF	Linear	Polynomial
BBA	Ac %	67.80	86.59	56.91	86.86	89.00	67.77	95.29	90.28	72.94	96.28	85.68	89.41	93.53	85.10	90.66
	Se	0.65	0.85	0.75	0.87	0.87	0.65	0.94	0.88	0.72	0.95	0.85	0.89	0.94	0.85	0.89
	Sp	0.67	0.87	0.38	0.86	0.89	0.67	0.95	0.90	0.73	0.96	0.86	0.90	0.93	0.86	0.90
	Pr	0.65	0.85	0.61	0.86	0.87	0.66	0.95	0.89	0.71	0.95	0.85	0.89	0.94	0.85	0.89
	F-score%	68.00	87.00	63.50	87.00	89.10	67.99	95.50	90.30	73.00	96.25	85.80	89.57	93.60	85.22	90.95
BDA	Ac %	93.19	91.84	67.88	94.17	91.10	72.34	95.47	92.57	72.47	**96.65**	93.88	92.85	93.65	91.89	88.98
	Se	0.92	0.89	0.65	0.93	0.90	0.68	0.95	0.90	0.68	**0.96**	0.92	0.90	0.92	0.89	0.86
	Sp	0.93	0.92	0.67	0.94	0.91	0.71	0.95	0.93	0.73	**0.97**	0.94	0.92	0.94	0.91	0.87
	Pr	0.92	0.89	0.65	0.93	0.89	0.70	0.95	0.89	0.70	**0.95**	0.91	0.91	0.91	0.90	0.85
	F-score%	93.20	92.00	67.97	94.09	91.58	72.80	95.50	92.70	72.99	**96.81**	93.97	93.00	93.90	92.00	89.00

Table 4. MLA: STEW dataset results of nature-inspired feature selection methods using SVM classifier.

Feature selection method	Parameters/ Kernel function	σ=0.5			σ=1			σ=1.5			σ=2			σ=2.5		
		RBF	Linear	Polynomial	RBF	Linear	Polynomial	RBF	Linear	Polynomial	RBF	Linear	Polynomial	RBF	Linear	Polynomial
BBA	Ac %	51.56	91.22	53.83	68.59	91.86	58.96	87.89	91.72	69.73	93.75	92.49	92.42	91.00	89.83	91.11
	Se	0.47	0.89	0.48	0.66	0.89	0.55	0.86	0.88	0.67	0.91	0.90	0.91	0.90	0.86	0.89
	Sp	0.50	0.92	0.53	0.68	0.91	0.57	0.88	0.91	0.69	0.94	0.92	0.92	0.91	0.89	0.91
	Pr	0.49	0.89	0.52	0.66	0.90	0.54	0.86	0.90	0.66	0.92	0.90	0.90	0.90	0.86	0.89
	F-score%	51.97	91.56	54.00	68.90	92.00	59.15	88.22	91.98	69.98	93.95	92.87	92.75	91.25	90.12	91.30
BDA	Ac %	50.83	88.55	50.89	56.32	90.93	52.45	88.63	91.18	57.47	**95.55**	90.73	92.85	91.73	89.97	67.87
	Se	0.49	0.85	0.47	0.52	0.88	0.48	0.85	0.90	0.54	**0.93**	0.88	0.90	0.89	0.87	0.65
	Sp	0.50	0.87	0.48	0.54	0.90	0.52	0.87	0.94	0.57	**0.96**	0.90	0.93	0.92	0.89	0.67
	Pr	0.47	0.85	0.47	0.52	0.89	0.52	0.85	0.89	0.55	**0.94**	0.88	0.91	0.89	0.88	0.65
	F-score%	50.97	88.72	51.00	56.78	90.98	52.70	88.95	91.47	57.80	**95.90**	90.99	93.00	92.00	90.05	68.02

Table 5. SLA: MAT dataset results of nature-inspired feature selection methods using KNN classifier.

Feature selection method	Parameters/ distance function	K=3			K=5			K=7			K=9		
		Euclidean	City block	Minkowski	Euclidean	City block	Minkowski	Euclidean	City block	Minkowski	Euclidean	City block	Minkowski
BBA	Ac %	93.82	93.29	93.84	93.86	94.95	93.96	91.90	92.86	90.77	89.93	92.80	90.88
	Se	0.92	0.92	0.92	0.93	0.93	0.92	0.91	0.91	0.90	0.88	0.91	0.89
	Sp	0.95	0.93	0.94	0.95	0.94	0.93	0.92	0.93	0.91	0.90	0.92	0.91
	Pr	0.92	0.93	0.92	0.94	0.93	0.92	0.90	0.92	0.90	0.89	0.92	0.90
	F-score%	93.99	93.54	93.90	93.97	95.00	94.10	91.92	93.10	90.90	90.23	92.98	91.00
BDA	Ac %	92.14	90.45	89.53	**96.20**	94.96	93.98	93.28	94.27	93.98	90.90	89.76	87.79
	Se	0.91	0.90	0.88	**0.96**	0.92	0.92	0.92	0.93	0.92	0.88	0.87	0.85
	Sp	0.93	0.91	0.89	**0.96**	0.96	0.93	0.93	0.94	0.93	0.90	0.89	0.87
	Pr	0.91	0.90	0.90	**0.94**	0.92	0.92	0.92	0.93	0.94	0.89	0.88	0.86
	F-score%	92.50	90.75	89.53	**96.70**	94.80	94.00	93.40	94.38	93.97	91.00	90.00	87.98

Table 6. SLA: STEW dataset results of nature-inspired feature selection methods using KNN classifier.

Feature selection method	Parameters/ distance function	K=3			K=5			K=7			K=9		
		Euclidean	City block	Minkowski	Euclidean	City block	Minkowski	Euclidean	City block	Minkowski	Euclidean	City block	Minkowski
BBA	**Ac %**	93.52	93.70	93.52	94.26	95.28	94.07	91.57	93.52	92.50	92.78	91.30	90.93
	Se	0.92	0.94	0.92	0.93	0.95	0.93	0.89	0.93	0.92	0.91	0.90	0.90
	Sp	0.95	0.94	0.95	0.96	0.96	0.95	0.94	0.94	0.93	0.94	0.93	0.92
	Pr	0.92	0.94	0.92	0.93	0.95	0.93	0.89	0.93	0.92	0.91	0.90	0.90
	F-score%	93.64	93.72	93.62	94.36	95.29	94.14	91.81	93.55	92.56	92.90	91.42	91.03
BDA	**Ac %**	93.77	91.78	93.59	**96.67**	95.72	95.18	94.89	92.70	93.51	93.89	91.58	93.88
	Se	0.91	0.91	0.90	**0.96**	0.93	0.95	0.92	0.91	0.93	0.93	0.91	0.92
	Sp	0.93	0.92	0.92	**0.97**	0.94	0.95	0.95	0.94	0.94	0.95	0.92	0.94
	Pr	0.93	0.91	0.93	**0.95**	0.95	0.95	0.94	0.93	0.94	0.93	0.92	0.92
	F-score%	93.00	91.40	93.50	**96.90**	95.49	95.12	94.78	92.00	93.45	93.90	91.38	93.99

Table 7. SLA: MAT dataset results of nature-inspired feature selection methods using SVM classifier.

Feature selection method	Parameters/ Kernel function	σ=0.5			σ=1			σ=1.5			σ=2			σ=2.5		
		RBF	Linear	Polynomial	RBF	Linear	Polynomial	RBF	Linear	Polynomial	RBF	Linear	Polynomial	RBF	Linear	Polynomial
BBA	**Ac %**	87.45	80.36	51.83	89.59	78.18	87.76	91.95	78.71	89.65	94.68	79.92	82.94	92.85	79.78	81.18
	Se	0.87	0.80	0.50	0.86	0.77	0.85	0.90	0.77	0.88	0.93	0.78	0.81	0.91	0.77	0.78
	Sp	0.88	0.83	0.50	0.89	0.77	0.8701	0.93	0.79	0.89	0.94	0.80	0.83	0.92	0.79	0.81
	Pr	0.86	0.80	0.50	0.88	0.77	0.864	0.90	0.78	0.85	0.92	0.78	0.81	0.91	0.78	0.79
	F-score%	88.00	80.47	51.90	89.70	78.43	87.81	92.00	78.90	89.80	94.98	79.99	83.00	92.94	79.90	81.40
BDA	**Ac %**	83.24	81.27	51.77	91.56	88.42	68.20	92.90	91.89	80.37	**95.29**	92.80	88.81	92.86	89.87	89.95
	Se	0.82	0.80	0.51	0.90	0.89	0.66	0.90	0.90	0.79	**0.94**	0.91	0.86	0.92	0.88	0.87
	Sp	0.83	0.81	0.52	0.91	0.89	0.68	0.92	0.92	0.80	**0.96**	0.93	0.88	0.93	0.90	0.89
	Pr	0.82	0.80	0.50	0.90	0.88	0.67	0.92	0.90	0.79	**0.94**	0.91	0.8843	0.91	0.89	0.88
	F-score%	83.42	80.41	51.95	91.68	88.55	68.32	92.98	91.90	80.50	**95.52**	92.98	88.95	92.99	89.91	89.98

Table 8. SLA: STEW dataset results of nature-inspired feature selection methods using SVM classifier.

Feature selection method	Parameters/ Kernel function	σ=0.5			σ=1			σ=1.5			σ=2			σ=2.5		
		RBF	Linear	Polynomial	RBF	Linear	Polynomial	RBF	Linear	Polynomial	RBF	Linear	Polynomial	RBF	Linear	Polynomial
BBA	**Ac %**	76.76	86.94	56.20	93.43	85.00	59.35	93.89	84.17	88.33	92.96	83.52	87.96	92.69	82.59	86.30
	Se	0.68	0.87	0.58	0.90	0.84	0.58	0.907	0.84	0.87	0.907	0.82	0.87	0.91	0.81	0.86
	Sp	0.86	0.87	0.55	0.98	0.86	0.61	0.976	0.84	0.90	0.955	0.86	0.89	0.95	0.84	0.86
	Pr	0.68	0.86	0.55	0.90	0.84	0.58	0.9072	0.84	0.87	0.907	0.82	0.8714	0.91	0.81	0.86
	F-score%	76.89	87.03	59.74	93.73	85.19	61.93	94.12	84.18	88.52	93.15	83.96	88.10	92.85	82.94	86.30
BDA	**Ac %**	87.50	86.02	56.30	93.52	86.57	57.31	**94.07**	84.81	88.15	91.85	83.80	88.61	90.74	83.43	87.31
	Se	0.81	0.86	0.59	0.91	0.87	0.61	**0.92**	0.84	0.88	0.90	0.83	0.88	0.88	0.84	0.87
	Sp	0.92	0.86	0.55	0.97	0.87	0.55	**0.97**	0.86	0.88	0.94	0.85	0.89	0.94	0.83	0.88
	Pr	0.80	0.86	0.59	0.91	0.87	0.55	**0.91**	0.84	0.88	0.90	0.83	0.88	0.88	0.84	0.87
	F-score%	88.80	86.08	56.74	93.73	86.59	61.32	**94.25**	85.04	88.15	91.63	84.11	88.73	91.07	83.47	87.35

4.2. SLA

We utilized features of lead F4 in the SLA. It represents the frontal lobe of the cerebral cortex, which is responsible for decision-making, future projections, preparation, attention, and emotional feelings. Here,

features extracted from the F4 lead have been used separately to analyse the ability of a single-lead EEG signal to detect cognitive load. Tables 5 and 6 show the results of the BBA-KNN and BDA-KNN classifiers using four distance functions with different K values. The highest accuracy of the MAT dataset was 96.20%, which was obtained in the BDA-KNN method at a Euclidean function with K=5. Similarly, STEW dataset accuracy of 96.67% was obtained in BDA-KNN. Tables 7 and 8 show the results of the BBA-SVM and BDA-SVM using Kernel functions on σ. The MAT dataset 95.29% accuracy was obtained in the BDA-SVM method at RBF with σ of 2. The highest accuracy of the STEW dataset, 94.07%, was obtained in the BDA-SVM method at kernel function RBF with σ of 1.5. Finally, we reported that the BDA-KNN is the best method for cognitive load detection compared to BBA. The optimal solution curve, confusion matrix, and ROC curve of SLA of STEW dataset are shown in Figure 5.

5. Discussion

The proposed technique is compared to known methodologies in Table 9. Various systems in the body respond to factors that change an organism's surroundings. As a result, different biosignals, namely EEG, GSR, ECG, and else, may be utilized to assess cognitive load. Apart from EEG, the effectiveness of our enhanced technique has been compared to various biosignals to demonstrate its efficacy.

Table 9. The proposed technique is compared to existing methods.

Year/reference	Method	Signal	Accuracy %
2017/[9]	SVM, HHT	EEG	89.07
2021/[13]	SWT,BO-KNN	EEG	96.07
2020/[43]	Supervise classification	ECG	75.02
2022/[46]	SWT, WOA, and SVM	EEG	97.25
2022/[52]	BSSA-KNN	EEG	95.5
2021/[53]	SWT, ADASYN-SVM	EEG	94
This work	**Ci-SSA, BDA, and KNN**	**EEG**	**97.98 (STEW) 97.30 (MAT)**

Researchers developed several methods for detecting cognitive load using EEG data. The work in [9] proposes a method for stress identification utilizing EEG signals that employ the HHT for time-frequency domain imperative attribute identification. As a classifier, SVM was utilized, and they were able to attain the accuracy of about 89.07%. Four levels of stress may be identified with a mean accuracy of 67.06%, three levels of stress with the accuracy of 75.22%, and two levels of stress with the accuracy of 85.71% using their technique. With the help of an EEG signal, the authors in [43] suggested a method for detecting stress in construction workers. Fixed and sliding windowing methodologies were used to analyse the frequency and time-based features from EEG signals. Finally, the authors used a few supervised learning approaches to assess workers' stress levels while on the job. The fixed windowing technique with the Gaussian SVM gave the highest accuracy of 80.32%, according to the results. Emotional Stress level detection using GA-based feature selection on EEG signals was first proposed by Shon et al. [44]. They used a genetic algorithm-based feature selection and a KNN classifier to analyse EEG signals and quantify stress in humans using the Database for Emotion Analysis using physiological signals from the DEAP dataset. Experimental outcomes show that by utilizing their method, stress could be predicted with a classification accuracy of 71.76% using GA-KNN. Sharma et al. [53] proposed an SWT-SVM for cognitive load performance identification using the MAT dataset and attained an accuracy of 94%. Yedukondalu et al. [54] proposed circulant singular spectrum analysis to decompose the contaminated EEG signal into IMFs. They effectively reduce the eyeblink artifact from EEG signals using evolution metrics.

In comparison to existing feature selection methods, our proposed binary dragonfly algorithm (BDA) gives better and comparable results as shown in Table 10.

Table 10. Comparison of the proposed method (BDA) with existing feature selection methods.

Reference	Signal	Feature selection method	Accuracy (%)
[44]	EEG	GA	73
[45]	EEG	Binary Harris Hawk optimization	96.88
[46]	EEG	Whale optimization algorithm	97.52
[48]	EEG	PSO	84
[49]	EEG	Bayesian optimization	92
[50]	EEG	Extreme gradient boosting Bayesian optimization	94.44
[51]	EEG	Evolutionary computation	96.97
This work	**EEG**	**BDA**	**97.98 (STEW), 97.30 (MAT)**

In comparison to previous methods for cognitive load detection analysis, our technique provides superior and comparable results. We have compared several decompositions and feature extraction methods to our proposed (Ci-SSA) method. Feature extraction and decomposition are done with the help of empirical mode decomposition (EMD), discrete wavelet transforms (DWT), and wavelet packet decomposition (WPD). DWT describes both the time domain and the frequency domain of a waveform. The problem with DWT is the time variance, which is important in statistical signal processing applications like EEG. SWT is used to rectify the issue that the DWT does not have translation invariance. This is a redundant technique because the number of samples coming out of each level of SWT is the same as the number coming in. The Hilbert transform can be used for EMD in the Hilbert–Huang Transform. It takes information about time and frequency from a signal that is not linear and does not stay the same. The idea behind this classifier is that an EEG signal can be decomposed into its intrinsic modes. This method is harder than others because it uses more than one mode at once. It is a way to show signals that are data-driven, adaptive, and have more than one resolution. It is very sensitive to noise, and mixing modes may make things even worse. So, the features taken from EEG signals have a greater effect on how well they can be classified. In this work, circulant singular spectrum analysis (Ci-SSA) technique was used for the decomposition of EEG signals into their constituent components (IMFs), particularly in the analysis of nonstationary and nonlinear EEG signals. However, like any other signal processing technique aforesaid, Ci-SSA has some potential biases that can affect its accuracy and reliability. Some of the potential biases of Ci-SSA for EEG signal decomposition are:

- Choice of window length: The choice of window length for the decomposition of EEG signals can affect accuracy, particularly for nonstationary signals.
- Overfitting: It can be overfitted to identify spurious components that do not represent physiologically meaningful EEG activity.
- A bias for low-frequency components: When the window length is long, CSSA decomposes EEG signals into low-frequency components. This bias can cause higher frequency components to be lost, which may include valuable EEG activity information.

Based on this, we have set the embedding dimension length (L) as 18 (L=18), and the window length as 2000, 512 samples for MAT, and STEW datasets, respectively.

Similarly, we have compared feature selection methods such as the genetic algorithm (GA), particle swarm optimization (PSO), Bayesian optimization (BO), differential evolution (DE), and whale optimization algorithm (WOA) to our proposed method (BDA). GAs can also be limited in scalability when the number of features is large. This can make it difficult to explore the entire search space and find the best subset of features. As the number of features increases, the time required to evaluate fitness and generate new solutions increases. PSO can converge slowly to the optimal feature subset, especially when the search space is large or complex. WOA can suffer from premature convergence, which occurs when the algorithm gets trapped in a local minimum and fails to explore the search space further. Table 11 shows the parameter settings for nature-inspired algorithms. The BBA and BDA models have fewer parameters compared to others, resulting in a lower convergence rate that increases with more parameters. Additionally, low convergence leads to a shorter response time. Efficiency: BDA is a fast and efficient optimization algorithm that can quickly identify the most relevant features from a large number of features. Accuracy: The algorithm is designed to identify the optimal subset of features that can provide the best accuracy for a given task, which can significantly improve the performance of the machine learning model. Robustness: The algorithm is able to handle noisy and incomplete data and can adapt to changing environments, making it a reliable approach for feature selection. According to the studied literature, the proposed decomposition (Ci-SSA) and feature selection (BDA) can overcome the limitations of previous state-of-the-art methods for cognitive load detection.

Table 11. Parameter settings for nature-inspired algorithms.

GA		**PSO and BPSO**		**BBA**		**BDA**	
No. of chromosomes(N)	10	No. of particles(N)	10	No. of bats (N)	10	No. of dragonflies (N)	10
Maximum iterations	100	Maximum iterations	100	Maximum iterations	100	Maximum iterations	100
Crossover rate(CR)	0.8	Cognitive factor(C1)	2	—	—	—	—
Mutation rate(MR)	0.3	Social factor(C2)	2	—	—	—	—
——-	—	Inertia weight(W)	1	—	—	—	—

Figure 6 represents the boxplots of the top five features of the BDA-KNN approach for the task and rest dimensions. The interclass distribution of the selected features is depicted in these plots. The rest signal feature mean entropy values are low compared to the task signal.

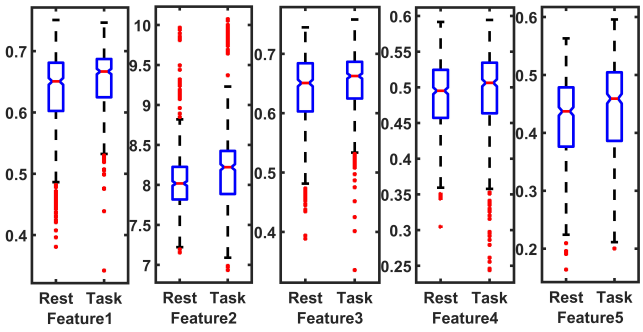


Figure 6. Boxplots describing the separability of the top 5 features as the rest and task states.

For the purpose of analysing the spatial brain activity during both the resting and task states, topographic maps have been created using characteristics computed from EEG signals. Figure 7 displays the topographic maps of a randomly chosen subject from the standard STEW dataset. Through the entropy values of five features, SE, AE, DE, IP, and FE, in distinct brain regions, these graphs show changes in activity levels.

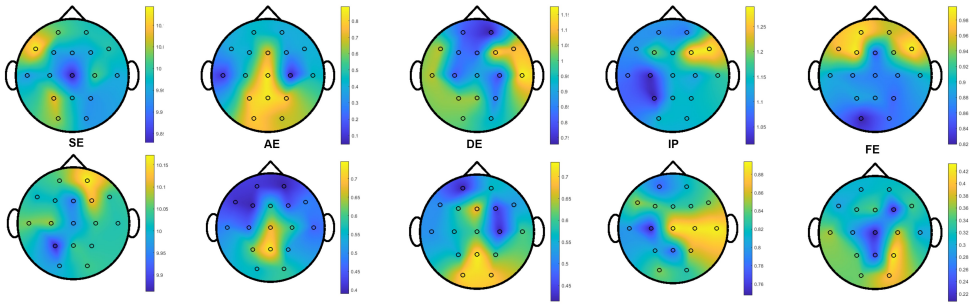
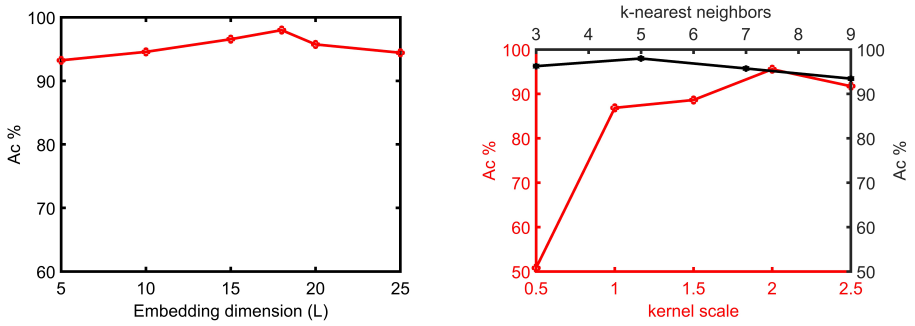


Figure 7. Topographic map for entropy-based features of Rest (topmost) and Task (bottom) states.

5.1. Experiment on parameters tuning and Statistical analysis

One of the most important C-SSA tuning parameters is the embedding dimension (L). Figure 8 shows the results of testing the proposed method for various values of this parameter. L = 18 yielded an effective result, so these values were applied throughout the research. In a similar vein, we tested the suggested approach by utilizing a variety of tuning parameters of BDA and classifiers namely, KNN and SVM with varied K and σ values of different distance and kernel functions. The statistical analysis of the EEG data was performed to see whether



(a) (b)

Figure 8. Plots showing the variation in the tuning parameters for classification Ac (%) of (a) Ci-SSA (b) BDA.

there was a statistically significant difference between the different brain EEG signals at rest and the condition of mental task load as determined by the ANOVA test with $p < 0.05$.

6. Conclusion and future work

The Ci-SSA method is used in this research to decompose the multitasking EEG signals from two datasets namely, MAT and STEW, into IMFs, which are commonly used to extract entropy-based features such as SE, AE, DE, IP, and FE. The total number of recorded channels was utilized to extract features, which were then input into nature-inspired algorithms such as GA, PSO, BPSO, BBA, and BDA. A unique and practical way to feature selection is provided by BDA for optimal feature selection, it is a purely natural selection strategy. The classification was done with the help of KNN and SVM by using the optimal feature subset (SFeat). The proposed framework BDA-KNN detected cognitive load with the highest accuracies for the MAT dataset and the STEW dataset, which were 97.30%, 97.98% and 96.20%, 96.67% in MLA and SLA approaches,

respectively. Similarly, we have achieved better Se, Sp, Pr, and F- score values. When compared to the existing feature selection and previous state-of-the-art methods, the presented method has been demonstrated to be more successful. The results indicate that using EEG for cognitive load detection is feasible. Our future work suggestions are:

- Cognitive load detection can be improved with a larger dataset and a wider age range.
- Noise and artifact reduction techniques can be utilized in the preprocessing stage.
- Classification accuracy may be increased with the use of deep learning models.
- Other data-driven decomposition methods, features, and feature selection techniques may be explored.

Acknowledgement

The authors are thankful to the coordinator of Microwave Lab, VIT-AP University for computational resources.

Conflict of interest

The authors declare no conflicts of interest.

Data access statement

In this work, two publicly available MAT [31] and STEW [32] datasets were used.

Ethics statement

Not applicable.

Funding statement

No funding was received.

References

[1] Stress, Healthy Workplaces Manage. Campaign Guide: Managing stress and psychosocial risks at work. OSHA, European Agency for Safety Health at Work, 2013.

[2] Duman RS. Neurobiology of stress, depression, and rapid acting antidepressants: remodeling synaptic connections. Depression and anxiety. 2014;31 (4):291-296. doi: 10.1002/da.22227

[3] Deschênes A, Forget H, Daudelin-Peltier C, Fiset D, Blais C. Facial expression recognition impairment following acute social stress. Journal of vision. 2015 ;15 (12):1383-1383. doi: 10.1167/15.12.1383

[4] Wielgosz J, Schuyler BS, Lutz A, Davidson RJ. Long-term mindfulness training is associated with reliable differences in resting respiration rate. Scientific reports. 2016 ;6 (1):27533. doi: 10.1038/srep27533

[5] He J, Li K, Liao X, Zhang P, Jiang N. Real-time detection of acute cognitive stress using a convolutional neural network from electrocardiographic signal. IEEE Access. 2019 ;7:42710-42717. doi: 10.1109/ACCESS.2019.2907076

[6] Dedovic K, Renwick R, Mahani NK, Engert V, Lupien SJ et al. The Montreal Imaging Stress Task: using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain. Journal of Psychiatry and Neuroscience. 2005 ;30 (5):319-325.

[7] Bógalo J, Poncela P, Senra E. Circulant singular spectrum analysis: A new automated procedure for signal extraction. Signal Processing. 2021 ;179:107824. doi: 10.1016/j.sigpro.2020.107824

[8] Asif A, Majid M, Anwar SM. Human stress classification using EEG signals in response to music tracks. Computers in biology and medicine. 2019 ;107:182-196. doi: 10.1016/j.compbiomed.2019.02.015

[9] Vanitha V, Krishnan P. Real time stress detection system based on EEG signals. 2017.

[10] Wang Q, Sourina O. Real-time mental arithmetic task recognition from EEG signals. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2013 ;21 (2):225-232. doi: 10.1109/TNSRE.2012.2236576

[11] Zarjam P, Epps J, Lovell NH. Characterizing mental load in an arithmetic task using entropy-based features. In2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA) 2012 ;(pp. 199-204). IEEE. doi: 10.1109/ISSPA.2012.6310545

[12] Shon D, Im K, Park JH, Lim DS, Jang B et al. Emotional stress state detection using genetic algorithm-based feature selection on EEG signals. International Journal of environmental research and public health. 2018 ;15(11):2461. doi: 10.3390/ijerph15112461

[13] Sharma LD, Chhabra H, Chauhan U, Saraswat RK, Sunkaria RK. Mental arithmetic task load recognition using EEG signal and Bayesian optimized K-nearest neighbor. International Journal of Information Technology. 2021 ;13:2363-9. doi: 10.1007/s41870-021-00807-7

[14] Zhang J, Min Y. Four-classes human emotion recognition via entropy characteristic and random Forest. Information Technology and Control. 2020 ;49 (3):285-298. doi: 10.5755/j01.itc.49.3.23948

[15] Komolovaitė D, Maskeliūnas R, Damaševičius R. Deep convolutional neural Network-based visual stimuli classification using electroencephalography signals of healthy and Alzheimer's disease subjects. Life. 2022;12 (3):374. doi: 10.3390/life12030374

[16] Krishna NM, Sekaran K, Vamsi AV, Ghantasala GP, Chandana P et al. An efficient mixture model approach in brain-machine interface systems for extracting the psychological status of mentally impaired persons using EEG signals. Ieee Access. 2019 Jun 10;7:77905-14. doi: 10.1109/ACCESS.2019.2922047

[17] Dursun M, Özşen S, Güneş S, Akdemir B, Yosunkaya Ş. Automated elimination of EOG artifacts in sleep EEG using regression method. Turkish Journal of Electrical Engineering and Computer Sciences. 2019;27 (2):1094-108. doi:10.3906/elk-1809-180

[18] Yalcin N, Tezel G, Karakuzu C. Epilepsy diagnosis using artificial neural network learned by PSO. Turkish Journal of Electrical Engineering and Computer Sciences. 2015;23 (2):421-432. doi:10.3906/elk-1212-151

[19] Milicevic M, Mazic I. Optimal set of EEG features in infant sleep stage classification. Turkish Journal of Electrical Engineering and Computer Sciences. 2019;27 (1):605-14. doi:10.3906/elk-1710-28

[20] Shon D, Im K, Park JH, Lim DS, Jang B et al. Emotional stress state detection using genetic algorithm-based feature selection on EEG signals. International Journal of environmental research and public health. 2018 ;15 (11):2461. doi:10.3390/ijerph15112461

[21] Acharya UR, Sree SV, Alvin AP, Suri JS. Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework. Expert Systems with Applications. 2012 ;39 (10):9072-8. doi:10.1016/j.eswa.2012.02.040

[22] Gürüler H, Şahin M, FERİKOĞLU A. Feature selection on single-lead ECG for obstructive sleep apnea diagnosis. Turkish Journal of Electrical Engineering and Computer Sciences. 2014;22 (2):465-78. doi:10.3906/elk-1207-132

[23] Hassan KM, Islam MR, Nguyen TT, Molla MK. Epileptic seizure detection in EEG using mutual information-based best individual feature selection. Expert Systems with Applications. 2022 ;193:116414. doi:10.1016/j.eswa.2021.116414

[24] Ang KK, Chin ZY, Zhang H, Guan C. Mutual information-based selection of optimal spatial–temporal patterns for single-trial EEG-based BCIs. Pattern Recognition. 2012 ;45 (6):2137-44. doi:10.1016/j.patcog.2011.04.018

[25] Hichem H, Elkamel M, Rafik M, Mesaaoud MT, Ouahiba C. A new binary grasshopper optimization algorithm for feature selection problem. Journal of King Saud University-Computer and Information Sciences. 2022;34 (2):316-328. doi: 10.1016/j.jksuci.2019.11.007

[26] Liu H, Motoda H. Feature extraction, construction and selection: A data mining perspective. Springer Science & Business Media; 1998 Aug 31.

[27] Oh IS, Lee JS, Moon BR. Hybrid genetic algorithms for feature selection. IEEE Transactions on pattern analysis and machine intelligence. 2004 ;26 (11):1424-37. doi: 10.1109/TPAMI.2004.105

[28] Mirjalili S. The ant lion optimizer. Advances in engineering software. 2015 ;83:80-98. doi: 10.1016/j.advengsoft.2015.01.010

[29] Emary E, Zawbaa HM, Hassanien AE. Binary ant lion approaches for feature selection. Neurocomputing. 2016 ;213:54-65. doi: 10.1016/j.neucom.2016.03.101

[30] Mirjalili S. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. Knowledge-based systems. 2015 ;89:228-49. doi: 10.1016/j.knosys.2015.07.006

[31] Zyma I, Tukaev S, Seleznov I, Kiyono K, Popov A et al. Electroencephalograms during mental arithmetic task performance. Data. 2019 ;4 (1):14. doi: 10.3390/data4010014

[32] Lim WL, Sourina O, Wang LP. STEW: Simultaneous task EEG workload data set. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2018 ;26 (11):2106-14. doi: 10.1109/TNSRE.2018.2872924

[33] de Carvalho M, Rua A. Real-time nowcasting the US output gap: Singular spectrum analysis at work. International Journal of Forecasting. 2017 ;33 (1):185-98. doi: 10.1016/j.ijforecast.2015.09.004

[34] Hekim M. The classification of EEG signals using discretization-based entropy and the adaptive neuro-fuzzy inference system. Turkish Journal of Electrical Engineering and Computer Sciences. 2016;24 (1):285-97. doi: 10.3906/elk-1306-164

[35] Can YS. Stressed or just running? Differentiation of mental stress and physical activity by using machine learning. Turkish Journal of Electrical Engineering and Computer Sciences. 2022;30 (1):312-327. doi: 10.3906/elk-2102-138

[36] Greene BR, Faul S, Marnane WP, Lightbody G, Korotchikova I et al. A comparison of quantitative EEG features for neonatal seizure detection. Clinical Neurophysiology. 2008;119 (6):1248-61. doi: 10.1016/j.clinph.2008.02.001

[37] Li X, Jiang Y, Hong J, Dong Y, Yao L. Estimation of cognitive workload by approximate entropy of EEG. Journal of Mechanics in Medicine and Biology. 2016 ;16 (06):1650077. doi: 10.1142/S0219519416500779

[38] MILICEVIC M, MAZIC I. Optimal set of EEG features in infant sleep stage classification. Turkish Journal of Electrical Engineering and Computer Sciences. 2019;27 (1):605-14. doi: 10.3906/elk-1710-28

[39] Principe JC, Xu D, Erdogmuns D. Renyi's entropy, divergence and their nonparametric estimators. Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives. 2010:47-102. doi: 10.1007/978-1-4419-1570- 2_2

[40] de Magalhães, Inês Alexandra Teixeira Antunes. Feel My Heart: Emotion Recognition Using the Electrocardiogram. 2021.

[41] Mirjalili S, Mirjalili SM, Yang XS. Binary bat algorithm. Neural Computing and Applications. 2014 ;25:663-681. doi: 10.1007/s00521-013-1525-5

[42] Mirjalili S. Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. Neural computing and applications. 2016 ;27:1053-73. doi: 10.1007/s00521-015-1920-1

[43] Jebelli H, Hwang S, Lee S. EEG-based workers' stress recognition at construction sites. Automation in Construction. 2018; 93:315-24. doi: 10.1016/j.autcon.2018.05.027

[44] Shon Dongkoo, et al. Emotional stress state detection using genetic algorithm-based feature selection on EEG signals. International Journal of environmental research and public health 15.11 (2018): 2461. doi: 10.3390/ijerph15112461

[45] Yedukondalu J, Sharma LD. Cognitive load detection using circulant singular spectrum analysis and Binary Harris Hawks Optimization based feature selection. Biomedical Signal Processing and Control. 2023 ;79:104006. doi: 10.1016/j.bspc.2022.104006

[46] Sharma LD, Bohat VK, Habib M, Ala'M AZ, Faris H et al. Evolutionary inspired approach for mental stress detection using EEG signal. Expert Systems with Applications. 2022 ;197:116634. doi: 10.1016/j.eswa.2022.116634

[47] Atkinson J, Campos D. Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers. Expert Systems with Applications. 2016 ;47:35-41. doi: 10.1016/j.eswa.2015.10.049

[48] Li R, Ren C, Zhang X, Hu B. A novel ensemble learning method using multiple objective particle swarm optimization for subject-independent EEG-based emotion recognition. Computers in biology and medicine. 2022 Jan 1;140:105080. doi: 10.1016/j.compbiomed.2021.105080

[49] Kamhi S, Zhang S, Ait Amou M, Mouhafid M, Javaid I et al. Multi-Classification of Motor Imagery EEG Signals Using Bayesian Optimization-Based Average Ensemble Approach. Applied Sciences. 2022 ;12 (12):5807. doi: 10.3390/app12125807

[50] Thenmozhi T, Helen R. Feature selection using extreme gradient boosting Bayesian optimization to upgrade the classification performance of motor imagery signals for BCI. Journal of Neuroscience Methods. 2022 ;366:109425. doi: 10.1016/j.jneumeth.2021.109425

[51] Nakisa B, Rastgoo MN, Tjondronegoro D, Chandran V. Evolutionary computation algorithms for feature selection of EEG-based emotion recognition using mobile sensors. Expert Systems with Applications. 2018 ;93:143-55. doi: 10.1016/j.eswa.2017.09.062

[52] Yedukondalu J, Sharma LD. Cognitive load detection using Binary salp swarm algorithm for feature selection. In2022 IEEE 6th Conference on Information and Communication Technology (CICT) 2022; (pp. 1-5). IEEE. doi: 10.1109/CICT56698.2022.9997949

[53] Sharma LD, Saraswat RK, Sunkaria RK. Cognitive performance detection using entropy-based features and lead-specific approach. Signal, Image and Video Processing. 2021 ;15 (8):1821-8. doi: 10.1007/s11760-021-01927-0

[54] Yedukondalu J, Sharma LD. Circulant Singular Spectrum Analysis and Discrete Wavelet Transform for Automated Removal of EOG Artifacts from EEG Signals. Sensors. 2023 ;23 (3):1235. doi: 10.3390/s23031235