Computers in Biology and Medicine 148 (2022) 105931

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

EEG-based BCI: A novel improvement for EEG signals classification based on real-time preprocessing

Said Abenna *, Mohammed Nahid, Hamid Bouyghf, Brahim Ouacha

Faculty of Sciences and Technologies, Hassan II University, Casablanca, Morocco

ARTICLE INFO

Keywords:
Brain–computer interface (BCI)
Electroencephalogram (EEG)
Signal processing
Feature extraction
Feature selection
Machine learning
Optimization

ABSTRACT

This work aims to improve EEG signal binary and multiclass classification for real-time BCI applications. Therefore, our paper discusses the results of a new real-time approach that was integrated into a complete prediction system, where we proposed a new trick to eliminate the effect of EEG's non-stationarity nature. This improvement can increase the accuracy from 50% using raw EEG to the order of 90% after preprocessing step in the binary case and from 28% to 78% in the multiclass case. Then, we chose to filter all signals by the proposed bandpass filter automatically optimized using the sine cosine algorithm (SCA) to find the optimal bandwidth that contains the entire EEG characteristics in beta waves. Moreover, we used a common spatial pattern (CSP) filter to eliminate the correlation between all extracted features. Then, the light gradient boosting machine (LGBM) classifier is also combined with SCA algorithm to build better prediction models. As a result, the outcome system was applied on UCI and PhysioNet datasets to get excellent accuracy values of higher than 99% and 95%, respectively, using the data acquired only from three channels. On the other hand, the related works used all the data acquired from 14 channels to find an accuracy value between 70% and 98.5%, which shows the robustness of our method to improve EEG signal prediction quality.

1. Introduction

EEG signals started in 1924 with the first acquisition by Hans Berger using a brain–computer interface (BCI) system [1–3]. Where the human brain is composed by billions of neurons, these neurons are connected between them via synapses and are distributed to other body organs from the spinal cord [4,5]. These neurons transmit information at the level of tiny electrical signals between each organ and various brain areas. So, BCI is used to diagnosis complex disease or to control external devices based on acquired brain activities when applying any physical movement [6–9]. BCIs are generally composed by an acquisition device and a processing step. Where the brain activities are analysed according several acquisition methods such as intracortical recording of neurons (INR), electrocorticography (ECoG), EEG, magnetoencephalography (MEG), ...etc. [10,11]. Where EEG signals are the results of acquisition at the scalp using a set of electrodes positioned following the international system 10–20 [6,12]. These signals acquired from the scalp mean a sum inhibition of dendrites and excitation potential of a set of neurons that have a vertical and parallel orientation below the crust [6,13]. Then, EEG-related BCI is a qualitative technological breakthrough aimed primarily at helping people with physical needs communicate effectively with their environment, inversely, with built-in devices to compensate for amputated limbs that are often uncomfortable and feel constant stress and pain due to their close association with muscles or nerves [14,15], but BCI systems allow remote control of prostheses, for example, via Bluetooth link by obtaining and processing EEG signals, then predicting tasks and sending them to the target. So it remains only to properly stabilize the industrial limb fixation, and the prediction level can become wider according to our study results. Also, by making the control of the limbs unconsciously possible through motor imagery tasks. Moreover, BCI systems based on EEG signals contribute to the analysis and diagnosis of all complex brain-related diseases like epilepsy and schizophrenia [16–18]. In the order to predict all features for each disease according to their impact on various brain areas.

The human visual system is composed by eyes that translate photons reflected or emitted from any object into tiny electrical signals. These signals are transmitted using brain neural network to the occipital lobe, which occipital lobe or the visual cortex area is responsible for the processing and analysis of eye activities [19]. So, EEG signals are used in this work to characterize EEG tasks based on visual cortex and motor imagery activities [1,6]. EEG signals are chosen over other methods because of the high signal resolution acquired in real-time with a low material cost compared to other techniques. On the other hand, these EEG signals are characterized by the non-stationary nature [6,19,20],

* Corresponding author.
E-mail address: said.abenna@etu.fstm.ac.ma (S. Abenna).

https://doi.org/10.1016/j.compbiomed.2022.105931
Received 25 April 2022; Received in revised form 12 July 2022; Accepted 30 July 2022
Available online 3 August 2022

such as EEG signals' characteristics change randomly over time during every single EEG task, which implies a great inconvenience for the classification step.

The classification stage is one of the most important prediction systems steps, as it is the way to separate EEG clusters, which ends with the production of a prediction model [9,21]. So, there is a lot of research and studies that develop new methods that can separate clusters' data spatially, some of which aim to separate data by linearly or non-linearly curve, marge, neural networks, or decision tree [22,23]. From this algorithm, we mention some classifiers like linear discriminant analysis (LDA), k-nearest neighbours (k-NN), random forest (RF), support vector machine (SVM), extreme gradient boosting (XGB) [6,12], and convolutional neural network (CNN) [21,24]. But each method has various challenges, with three types of data, including linear separable clusters such as text and speech data, including also non-linear separable data like images, and non-separable data like EEG signals [6,25,26]. Thus, before classification, it must study and analyse the data complexity level and select or develop appropriate classifiers with good accuracy or improve the processing stages. On the other hand, accurate classifiers often lack classification speed due to algorithm complexity. Therefore, it is exceptional to develop some classifiers characterized by both speed and high accuracy level [6,24]. Then, in the absence of these optimal classifiers, many researchers sought to improve the accuracy level by combining optimization algorithms with fast classifiers like LDA, SVM, and LGBM [15].

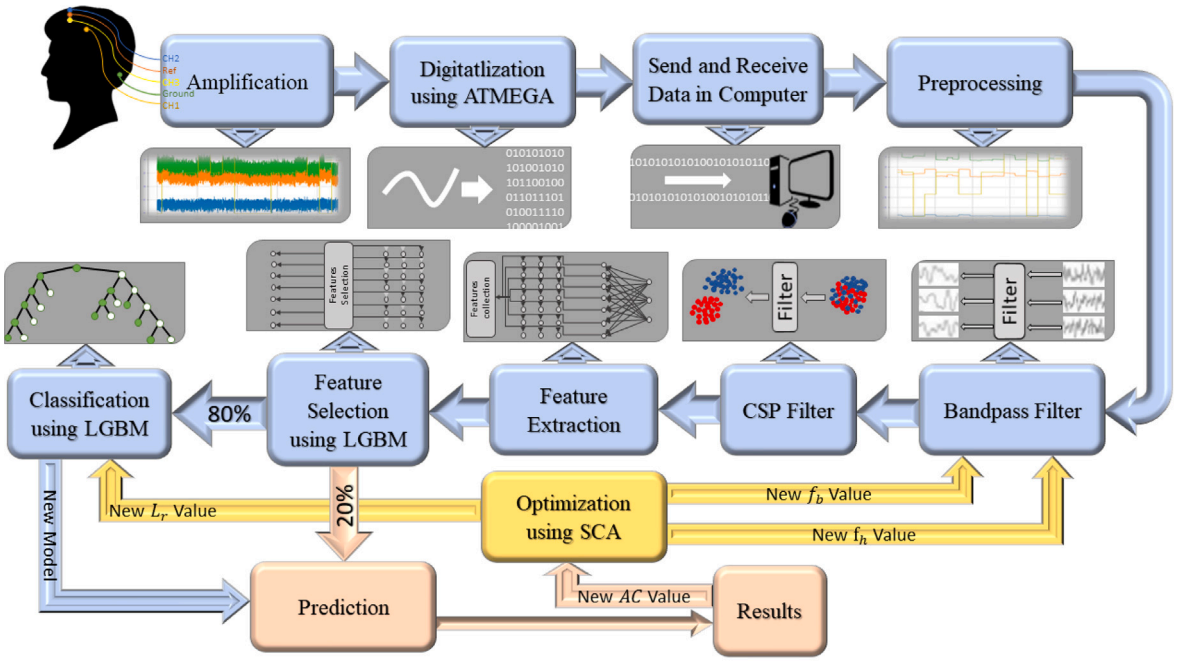
Moreover, the prediction quality can be improved from the signal processing stagesS by extracting as optimal features as possible, such as kurtosis, skewness, interquartile, and Hjorth [27–29], where these features contain specific temporal, spectral, and spatial characteristics. Then, several research works also use spatial filters such as independent component analysis (ICA) and CSP [6,9,21,30,31], these filters reduce the mutual information between all extracted features by enlarging the distance between classes' data, which increases the diversity of characteristics and then the separation between clusters spatially [6,32]. Thus, filtering using CSP is now one of the current foundations of BCI systems, where it aims to reduce the level of similarity between channels as well as between extracted features by maximizing the variance difference between two tasks' data. [14]. So the goal is not primarily to increase the accuracy level but to clean and remove similarities between features from different classes [6,33], CSP is therefore used to diagnose brain-related conditions, which distinguish the most active areas in the brain during each task.

On the other hand, raw EEG signals can be changed to ordinary features [27,29,34], which is able to increase accuracy level. To study the impact of each feature, the feature selection step can be used to auto-calculate the importance of each feature [6,9,35,36]. Thus, all non-important features are removed, then the accuracy levels increase and the size of the data will be reduced. The most important features of BCI have simple mathematical equations, in order to guarantee better system performance for BCI applications in real-time. So, to improve the EEG signal classification for real-time applications, our study introduced a new approach for the preprocessing step that makes it possible to eliminate the EEG signal's non-stationarity nature over time, in such a way that the signal characteristics of each task remain stable. In addition, we used a bandpass filter optimized automatically using SCA algorithm [37–39] to find the best bandwidths containing all characteristics. As a result, all optimal bandwidths belong to beta waves between 15 Hz and 30 Hz. Moreover, this work uses a CSP filter to minimize the correlation between all EEG channels. Therefore, all extracted features have been well stabilized in the prediction system, and LGBM algorithm is used for the classification and feature selection steps [6,24,33,40]. As a result, the entire system was applied to EEG data to detect baseline eye states such as the SAES acquired using our EEG device, UCI dataset [1,41,42], and motor imagery tasks using PhysioNet-MI dataset as illustrated in Fig. 1, to find that average accuracy values are greater than the related work's results.

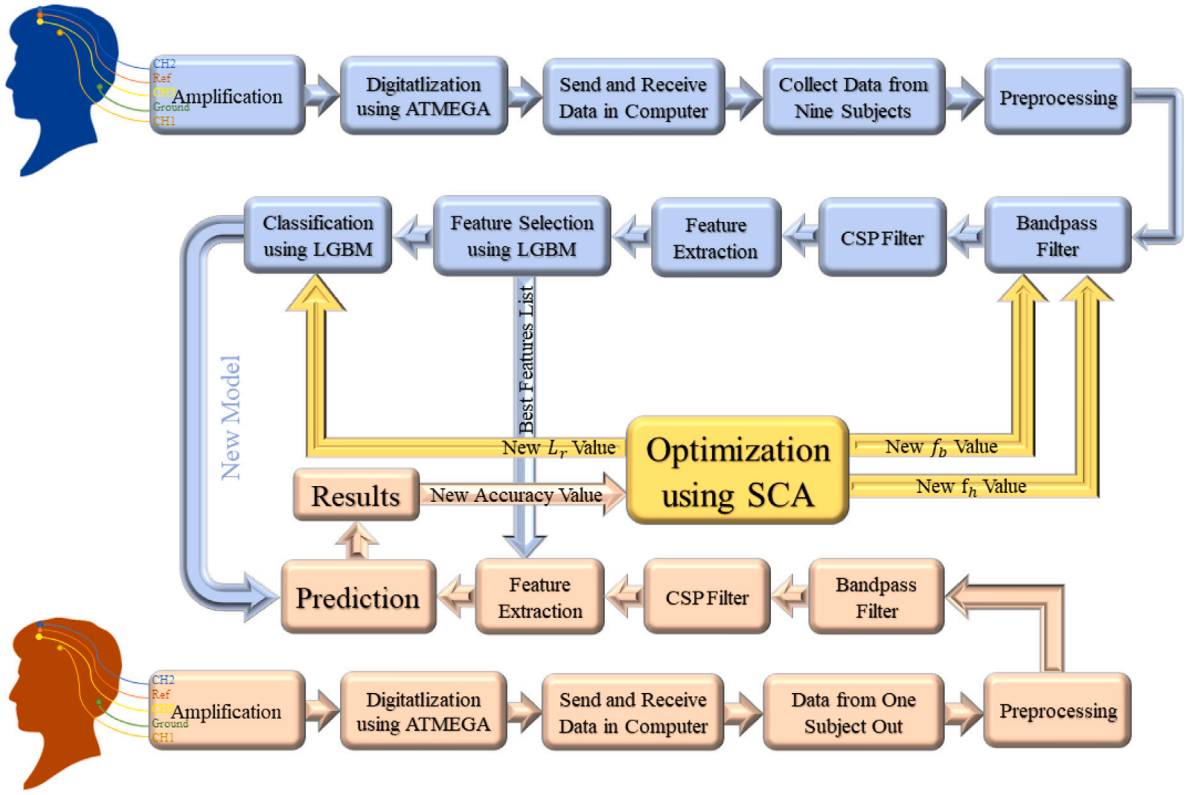
The rest of this paper is structured as Section 2 is a brief survey of relevant literature. Section 3 presents proposed EEG recordings, preprocessing, and bandpass filter optimized using SCA steps, as well as suggested spatial filter, feature extraction, feature selection procedure, and proposed LGBM-SCA classification to build prediction models. Section 4 presents and discusses all results found during each improvement step with a comparison with related work results. Section 5 presents and concludes our results obtained from this study and our probable future work.

2. Related work

The related work began by working on seriously and effectively baseline eye state EEG signal classification with Hamilton et al. [43] by building groups of enthusiastic learners, assuming that a system could be built with relatively accurate levels for K*-based systems but fast enough to be used within BCI. As a result, it was proposed to develop three machines by the combination of a rotational forest (RRF) based on RF, a rotational forest based on J48 trees and enhanced by adaptive boosting (ada(RJ48F)), and an RRF model set with K* model (RRF+K*). The results of this work show that optimal learners, such as K* algorithm is too slow to be used in BCI devices, while the ada model performed well within the time limits of classification and real-time control. Then, Reddy and Behera [44] develop the following designs: Multilayer neural networks with ReLU and dropout include deep learning networks based on unsupervised learning and dropout masks on deep neural networks (DNN). They compare their results with those of K*, RRF, K*+RRF and ada (RJ48F). In addition, utilizing t-distributed stochastic neighbour embedding (t-SNE) based visualizations and fitting elliptical contours to the features, they performed an in-depth analysis of binary class features. Also, Chhabra et al. [45] implements various machine learning algorithms on an EEG eye state dataset (UCI) to predict human baseline eye state based on previous EEG signals. Also, RF and DT were implemented for comparative analysis. The results show that RF obtained a high prediction accuracy of 92.43%. In simulating away, Ketu and Mishra [41] proposes a hybrid classification model for EEG signal-based eye state detection. To examine its appropriateness and correctness, this hybrid classification model was compared to other classic machine learning models (eight classification models and six state-of-the-art methodologies). This suggested classification technique creates a machine learning-based hybrid model for more predicted baseline eye state features based on raw EEG. Also, Hassan et al. [46] introduce an ANN-based technique in which multiple artificial neural networks (ANNs) are separately trained utilizing various sections of training data. The results of each ANN are then integrated with the results of another ANN to improve the prediction. Where the suggested ensemble technique has a short learning curve and produces extremely accurate baseline eye states classification. Inversely, Saghafi et al. [47] uses the maximum and minimum calculated from cross-channel to detect EEG data for a possible change in eye condition. When a possible change is detected, the last two seconds of the signal are analysed to extract meaningful information using a multi-variable experimental mode. After that, all features are classified using logistics regression (LR) and ANN to predict the baseline eye state. In addition, Zhou et al. [42] offers a prediction system with four essential steps: EEG signal preprocessing, feature extraction, feature selection, and classification. In begin, they use the symlets-8 wavelet to decompose the raw EEG and select the 5th-floor decomposition, from which EEG is subsequently denoised using the heuristic threshold approach. Where Zhou et al. [42] proposes a unique feature extraction method using an information accumulation strategy based on discrete wavelet transform (DWT). Then, they located ideal features using the CfsSubsetEval evaluator by finding ideal channel positions with good correlation coefficients, which is based on the BestFirst search technique for feature selection. Finally, they select RF as a classifier. We also have a work that can



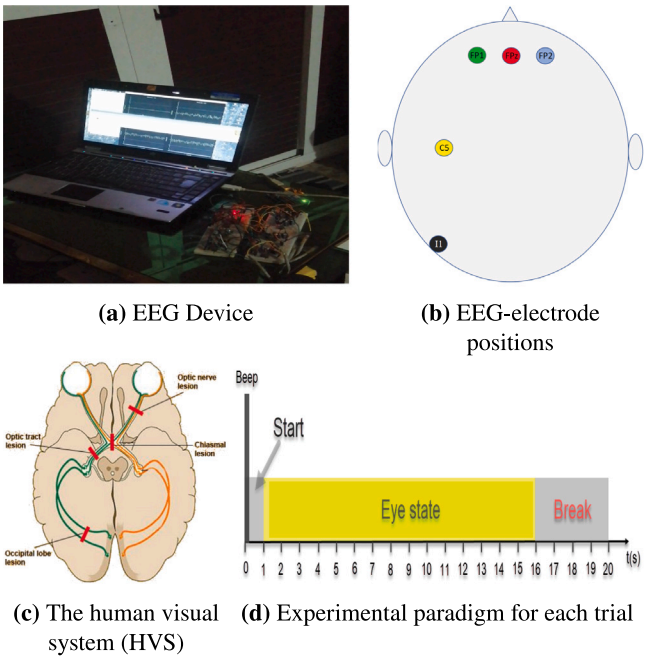
(a) Testing by trials, using 20% of each subject data



(b) Testing by subject, using cross-subject data

Fig. 1. General architecture for EEG prediction system based on two different testing-validation methods.

optimize the classifier parameters, where Hussien et al. [15] suggests using a modified grey wolf enhancer (MGWO) to pick the best EEG channels for BCI application. MGWO imitates the leadership of grey wolves and nature's hunting method, which considers meta-heuristics swarm intelligence algorithms and is integrated with two adjustments to balance exploration and exploitation: the first adjustment applies an exponential change in the number of iterations to increase the research space according to the exploitation, and the second adjustment is the transit process used to increase population diversity and enhance the capacity for exploitation. On the other hand, Kim et al. [48] suggests an algorithm that uses data-based multi-variable experimental mode degradation (MEMD) in order to obtain alpha and beta rhythms from non-linear EEG signals. Then, the complex unrelated CSP algorithm (SUTCCSP) is applied on extracted rhythms, so that complex data becomes unrelated and their common false contrast provides additional information about the power difference between the two rhythms.



(a) EEG Device **(b)** EEG-electrode positions

(c) The human visual system (HVS) **(d)** Experimental paradigm for each trial

Fig. 2. Proposed EEG recording system, electrode placement scheme, and experimental paradigms for SAES dataset acquisition.

The extracted features are classified using SUTCCSP which increases variations between layers by using different classification algorithms to separate left and right EEG motor imaging acquired from the Physionet database. Additionally, the SUTCCSP data on the power differential between the alpha and beta rhythms provides a crucial element for classifying the left- and right-hand motor imagery tasks. Additionally, when compared to traditional IIR filtering, MEMD has shown to be a more effective preprocessing technique for non-stationary and non-linear EEG signals. The RF classifier also performed well in classifying motor imagery tasks. For CNN application on signals such as motor imagery tasks, Dose et al. [21] offers a deep learning model based on CNN layers to learn generic features and minimize dimensions, while a fully connected (FC) conventional layer is utilized for classification. Together, they create a comprehensive end-to-end model that may be used with unprocessed EEG inputs. Additionally, it has been improved by being tailored to a certain subject. As a result, the learnt filter analysis sheds light on how a model like this determines a categorization. The globally chosen classifier achieved an accuracy of 80.38 percent for data containing two classifications. Using 5-fold cross-validation, the accuracy is examined. As a new approach, transient learning was used to adapt intra-subject classification and improve the average overall accuracy to 86.49%. In the same way, Lun et al. [49] suggests a deep CNN structure with separate time and spatial filters, which selects raw EEG signals for pair electrodes above the motor cortex area as hybrid samples without any prior processing or feature extraction processes. In the proposed structure, CNN was applied across 5 layers to learn EEG characteristics, a maximum assembly of 4 layers was used to reduce dimensions, and the FC layer was used for classification. Dropouts and batch normalization are also used to reduce the risk of overfitting. In the experiment, the use of EEG data from 10, 20, 60 and 100 subjects from PhysioNet dataset was used as a data source.

Our work aims to improve the accuracy of various EEG tasks at minimal hardware and software costs while resolving the classification slowness and the instability of results for different subjects. Thus, the main contributions of this study are:

- Removing the EEG's non-stationary nature for every subject during the preprocessing step;
- Detect the important frequency bandwidth for EEG signals during each classification in order to better analyse various brain activities;
- Use different test-validation to show the stability of the results after removing the EEG's non-stationary nature;
- Select important channels/features to study the prediction system stability when changing between subjects, tasks, and datasets (EEG device).

3. Materials and methods

3.1. EEG device

The data that has been used in this work has been acquired using our EEG device, such that this device is composed by 5 input dry electrodes, where that 3 active electrodes are placed at the levels of C5, FP2, and FP1, as illustrated in Fig. 2(b). A ground electrode was also inserted in I1 and a reference electrode was installed in FPz. EEG signals acquired from C5, FP1, FP2, and FPz electrodes are amplified using TDA2822, then digitized and multiplied by ATMEGA328PU and sent via a Bluetooth model to a computer with a sampling frequency of 120 Hz.

3.2. Datasets used in this work

3.2.1. EEG data acquisition: SAES dataset

The SAES dataset[1] is acquired by our EEG device, as data is acquired from 10 subjects according to the paradigm per trial in Fig. 5(d). Each subject is composed by 20 trials, with an average trial duration of 20 s per trial, the sampling frequency is at the level of 120 Hz, knowing that one second is for beep or stimulation, 15 s for each task, and 4 s for relaxation at the trial end. Each subject is composed by four labels with a percentage of 25% of the data per label, as follows:

- 'OE': when eyes are open;
- 'CE': when eyes are closed;
- 'CLE': when only the left eye is closed;
- 'CRE': when only the right eye is closed.

Then total samples acquired are at the level of 30 000 samples per subject, and during each classification, we subdivide the data using textsftrain_test_split[2] function into 80% for classification and to build all prediction models, and the rest 20% of the data is used for prediction during the testing step to evaluate the system performances. All codes are available on our online GitHub account.[3]

3.2.2. UCI eye state dataset

In this work, we also use UCI data[4] collected for the eye states detection. This dataset is used to compare the performance of our system with the literature. As UCI data, a NeuroHeadSet device with 14 acquisition electrodes (AF3, F3, FC5, T7, P7, F7, O1, O2, P8, T8, FC6, F4, F8, and AF4) is used. [1], the sampling frequency is at the level of 128 Hz, and this dataset size is 14 980 samples, where this dataset is composed only by two eye states: eyes are closed ('CE') and eyes are open ('OE').

[1] https://github.com/Abennasaid94/Eye_States/tree/main/SAES/Raw_EEG.
[2] https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html.
[3] https://github.com/Abennasaid94/Eye_States.
[4] https://archive.ics.uci.edu/ml/datasets/EEG+Eye+State.

3.2.3. PhysioNet-MI dataset

PhysioNet-MI data[5] is one of the most important EEG data [6], so it is much used to study motor imagery tasks, where this dataset is formed by 109 subjects, each subject acquired using BCI2000 device composed by 64 EEG channels (FC5, FC3, FC1, FCz, FC2, FC4, FC6, C5, C3, C1, Cz, C2, C4, C6, CP5, CP3, CP1, CPz, CP2, CP4, CP6, FP1, FPz, FP2, AF7, AF3, AFz, AF4, AF8, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FT7, FT8, T7, T8, T9, T10, TP7, TP8, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO3, POz, PO4, PO8, O1, Oz, O2, and Iz) during the activate of six various tasks, where each task takes 4s per trial with 2s additional for relaxation between trials. All six tasks are when close the eyes ('CE'), open the eyes ('OE'), imagine the right-hand movement ('R'), imagine the left-hand movement ('L'), imagine the movement of both hands ('LR'), and imagine the movement of both legs ('F'). Where, we select only first ten subjects from this dataset, Where during the test validation by trials for each subject, we use 80% of trials as training data for classification and the rest 20% of trials of each subject are used as testing data during the prediction step.

3.3. Real-time preprocessing

The preprocessing step is remained very important in BCI system, such that this step makes it possible to adapt the data acquired using the acquisition device with the prediction system requirements. In this step, we demultiplex the received data and then subtract the reference signal from the active electrodes ($V_i - V_{ref}$). Therefore, the real-time preprocessing, a way has been developed to minimize the EEG's non-stationarity nature effect, where this non-stationarity effect is a great constraint for classification, knowing that all EEG signals characteristics change over a long period, which implies a desynchronization between the tasks data in real-time. To minimize this effect, a methodology has been introduced that allows the average of the last seconds of the signal to be a reference point for new samples, after subtracting this average value from the last samples amplitude, as shown in the following Eq. (1):

$$\begin{aligned} X_P^c[t] &= X^c[t] - \text{mean}(X^c[t - n \times f_s : t]) \\ &= X^c[t] - \frac{1}{f_s} \sum_{i=t-n\times f_s}^{t} X^c[i] \end{aligned} \tag{1}$$

With X^c is data carrier for each electrode (c), and n is the number of seconds can be used to detect each task, such that we choose $n = 8$ in this work. Then, when there is no active task, the resulting signal converges to the origin ($0V$), but during an active task, the signal varies according to the activity potential. As a result, this equation makes it possible to improve the separation between different tasks characteristics over time, which shows that signals remain stationary.

3.4. Automatic optimization of the bandpass filter (OBPF)

Frequency filtering is an essential tool to process and eliminate the unnecessary EEG signals frequencies, in this sense, we used a bandpass filter (BPF) to find the best signals bandwidths corresponding to each classification. To determine these bandwidths, we combined SCA algorithm to cut-off frequencies parameters: higher (f_h) and lower (f_b). Where SCA algorithm generates new parameters of f_b and f_h according to the accuracy variation at each moment by the following steps:

1. Initialize accuracy value to 0%;
2. SCA generates new values for parameters f_b and f_h;
3. Filtering signals for each channel using Eq. (2);
4. Calculate new accuracy value based on LGBM algorithm;
5. SCA generates new values of f_b and f_h to converge at best possible accuracy value.

$$X_{\text{BP}}^c[t] = \text{BPF}\left(X_P^c[t - f_s : t]\right)[-1] \tag{2}$$

The above steps allow the filter to converge to the bandwidth that contains all characteristics corresponding to each EEG task. As the bandpass filter output is calculated using two libraries in python.scipy[6]: butter and lfilter of following Eqs. (3) and (4):

$$a, b = \text{butter}\left(\text{order} = 3, \left[\frac{2f_b}{f_s}, \frac{2f_h}{f_s}\right], \text{btype} = \text{'band'}\right) \tag{3}$$

$$X_{\text{BP}}^c[t] = \text{lfilter}\left(a, b, X_P^c[t - f_s : t]\right)[-1] \tag{4}$$

3.5. Common spatial pattern (CSP) in real time

Spatial filtering is an essential step in EEG signal processing systems, such that spatial filters are more often used to minimize the correlation coefficients between EEG channels in such a way increase the independence between EEG characteristics between different tasks. In this state, it is possible to improve the separation between classes and then increase the prediction system accuracy. In addition, these filters are more widely used to diagnose the brain activity distribution in different brain areas. Therefore, these filters trace the neural networks distribution between the eyes and the occipital lobe. From all spatial filters, we chose CSP filter because of its reliability to minimize correlation between channels compared to ICA and principal component analysis (PCA) algorithms. First, eigenvectors matrix is specified to be equivalent to the number of classes used ('OE', 'CLE', 'CRE', 'EC'). As for the data ($X_{\text{BP},i}^c$) of each class i, we calculate its covariance matrix (σ_i) and the covariance matrix (σ_i') for the rest data of other classes ($X_{\text{BP},i}'^c$) using Eqs. (5) and (6),

$$\sigma_i = \frac{X_{\text{BP},i}^c X_{\text{BP},i}^{cT}}{\text{trace}(X_{\text{BP},i}^c X_{\text{BP},i}^{cT})} \tag{5}$$

$$\sigma_i' = \frac{X_{\text{BP},i}'^c X_{\text{BP},i}'^{cT}}{\text{trace}(X_{\text{BP},i}'^c X_{\text{BP},i}'^{cT})} \tag{6}$$

Then, we calculate the eigenvalues matrix (E_i) and eigenvectors matrix (U_i) for the sum of covariance matrices ($\sigma_i + \sigma_i'$). Then we calculate the projector J_i using the following Eq. (7):

$$J_i = \left(\text{diag}(E_{i,l})^{-1}\right)^{\frac{1}{2}} U_{i,l} \tag{7}$$

After, we calculate the new common values (E_i') and vectors (U_i') between the matrices $J_i(\sigma_i J_i^T)$ and $J_i(\sigma_i' J_i^T)$. Then we determine the eigenvector $U_{i,l}'$ corresponding to the maximum eigenvalue $E_{i,l}'$. Finally, we calculate the vector CSP corresponding to the ith class by the real part of the vector.

$$W_i = \text{real}(U_{i,l}^T J_i) \tag{8}$$

Then, all that remains is to apply CSP on the filtered signal X_{BP} at each moment (t) by the following Eq. (9):

$$X_{\text{CSP}}^c[t] = \text{mean}\left(W_0 X_{\text{BP}}^{cT}\right) \tag{9}$$

3.6. Features extraction

The feature extraction is an essential and very important step for EEG signals classification, where the prediction system extract only temporal, frequency, and spatial characteristics that contain the totality of special characteristics for each EEG task. These features are more commonly used for the brain activation diagnosis. In this work, we extracted a set of 11 essential features such as variations in values

[5] https://physionet.org/content/eegmmidb/1.0.0.

[6] https://docs.scipy.org/doc/scipy/reference/generated/.

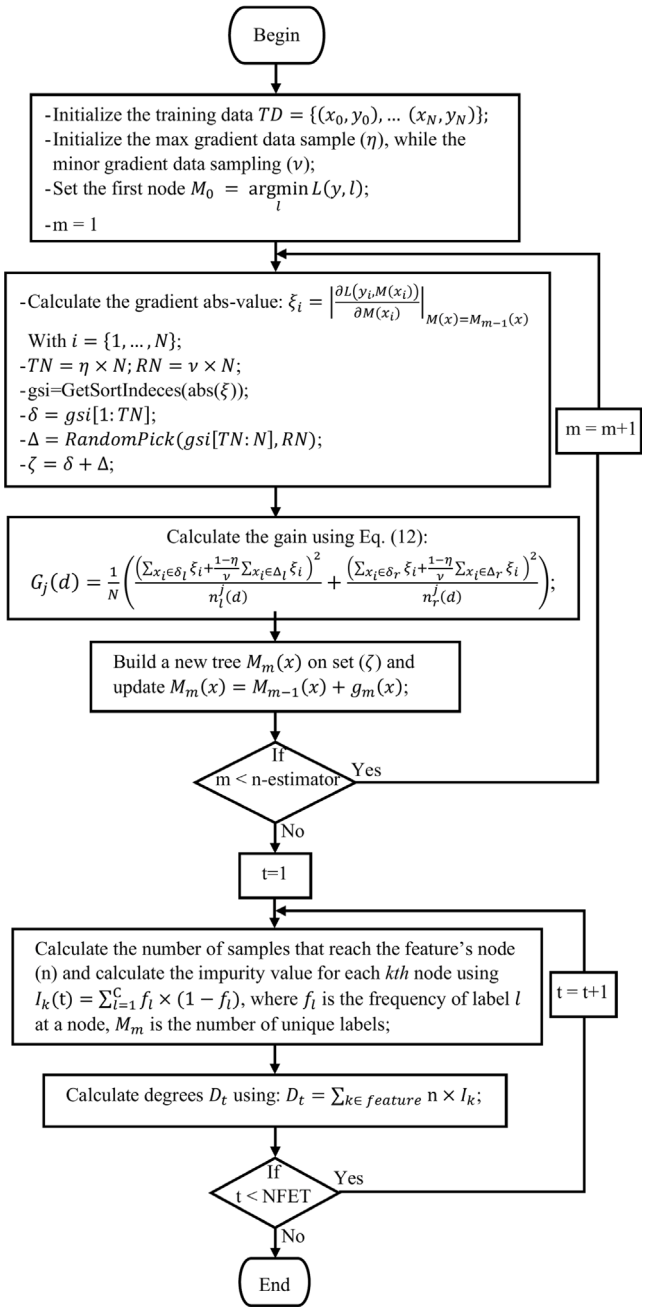


Fig. 3. Flowchart for LGBM algorithm.

of max, min, mean, median, kurtosis, ...etc., which are summarized in Table A.1. So, we extracted 11 features from each active electrode to generate 33 features, knowing that each feature updates at each moment from last-second samples ($X_{\mathrm{BP}}^{c}[t - f_s : t]$).

3.7. Classification using the combined LGBM with SCA algorithm

The classification step is the prediction system heart, as we generate in this step prediction models allow to separate between different classes' signals automatically, knowing that there is a large set of classification algorithms, and each of these classifiers is followed by a precise method. In the case of this work, we chose LGBM algorithm [6,12] as illustrated in Fig. 3 based on results of Table 3, which is based on the gradient boosting method, which allows building a prediction tree using the gradient boosted regression trees (GBRT) algorithm by calculating gradient descending for the loss function concerning each cluster data,

$$y_i' = -\frac{\partial \mathrm{Loss}\left(y_i, M_0(X_i)\right)}{\partial M_0(X_i)} \tag{10}$$

with the model is initialized by M_0 of Eq. (11),

$$M_0 = \underset{\delta}{\operatorname{argmin}}\left(\mathrm{Loss}(y_i', \delta)\right) \tag{11}$$

where, the gain for LGBM is (as illustrated in Fig. 3):

$$G_j(d) = \frac{1}{N}\left(\frac{\left(\sum_{x_i\in\delta_l}\xi_i + \frac{1-\eta}{\nu}\sum_{x_i\in\Delta_l}\xi_i\right)^2}{n_l^j(d)} + \frac{\left(\sum_{x_i\in\delta_r}\xi_i + \frac{1-\eta}{\nu}\sum_{x_i\in\Delta_r}\xi_i\right)^2}{n_r^j(d)}\right) \tag{12}$$

So, the LGBM's learning rate (L_r) is calculated by the following Eq. (13):

$$g_1 = \underset{\delta}{\operatorname{argmin}}\left(\mathrm{Loss}\left(y_i', M_0(X_i) + \delta\right)\right) \tag{13}$$

Finally, the new model is calculated by:

$$M_1 = M_0(X_i) + g_1 \tag{14}$$

After a number of iterations (m), the new model is generated using training data, then we evaluate the model performances using the prediction of testing data. To properly increase the generated model performances, we give SCA algorithm (introduced by Mirjalili [39]) the access to evaluate LGBM parameters following the algorithm illustrated in Fig. 4, specifically for the learning-rate parameter, where at each instance, SCA calculates get the accuracy using LGBM to will converge to the right L_r value. In such a way, SCA generate a set of agents, and each of these agents has a precise spatial position calculated using Eqs. (15) or (16). Where, if a value generated randomly is positive, the new agent position is calculated by the following Eq. (15):

$$L_r' = L_r + 2\left(1 - iter^{-1}\right) \times \sin(2\pi r_2) \times \left|r_3 \times (1 - \mathrm{AC}) - L_r\right| \tag{15}$$

Otherwise, if the value generated randomly is negative, the cosine is used instead of sinus by the following Eq. (16):

$$L_r' = L_r + 2\left(1 - iter^{-1}\right) \times \cos(2\pi r_2) \times \left|r_3 \times (1 - \mathrm{AC}) - L_r\right| \tag{16}$$

With r_2 and r_3 being random values. Then we classify the data using the new learning-rate value and we determine the new accuracy value. If the accuracy has been reduced, we return to the ancient learning-rate value and we re-establish SCA process. In the same way, we used SCA to optimize the bandpass filter parameters.

3.8. Feature selection

The feature selection step is a way to minimize the number of features [1,6,12] that can be extracted without degrading the prediction system accuracy, knowing that this step reduces the size of data classified during training or predicted data during testing, so this step increases SCA algorithm convergence speed towards the best optimal parameters. In this sense, the feature selection process calculate different features importance degrees using LGBM algorithm, where LGBM build the prediction model in the tree form, then we calculate the frequency (f_i) of each label y_i for each jth node of each feature (F) in the tree, then we calculate the impurity value of each node by following Eq. (17):

$$I_k = \sum_{i=1}^{C}(1 - f_i)f_i \tag{17}$$

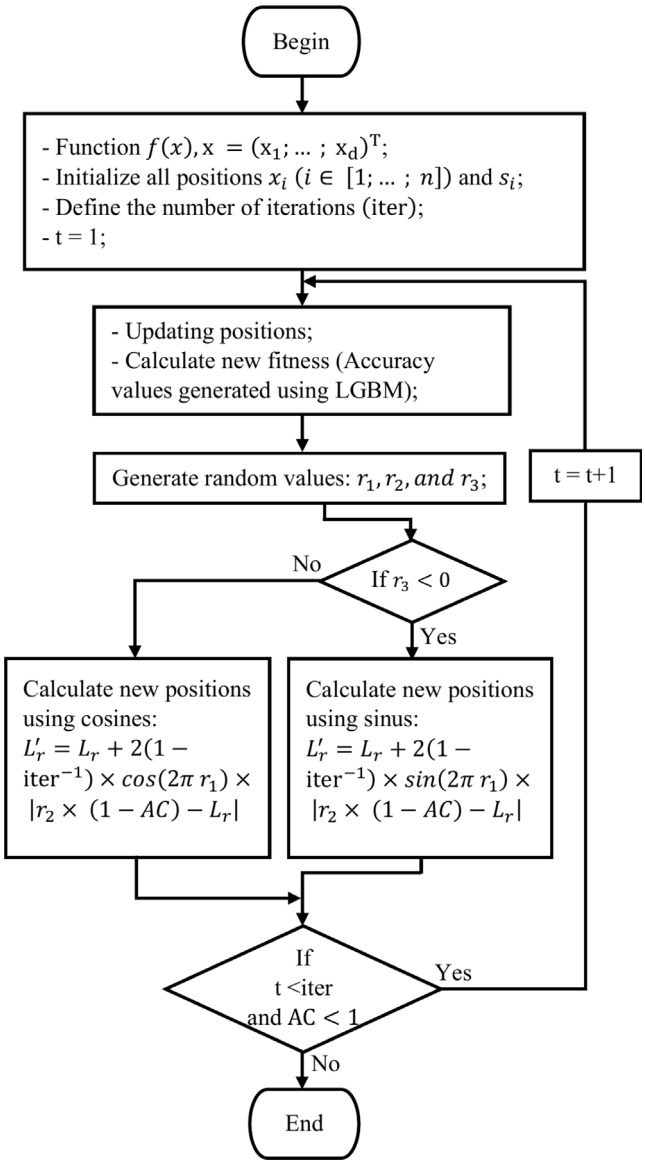


Fig. 4. Flowchart for SCA algorithm.

where C is the number of classes used. Finally, the importance degree (D) of each feature is calculated by the following Eq. (18):

$$D = \sum_{k \in F} nI_k \tag{18}$$

where n is the number of samples built by nodes of each feature (F). After calculating all features importance degrees, we organize these features from the most important to the least. Then we evaluate the number of features (NFET) that can be chosen for classification according to the prediction accuracy values, as illustrated in Fig. 7.

3.9. Model evaluation

All prediction made during the testing is extracted as some of the generated model performances during classification. These performances are calculated from many true predictions (TP) and false predictions (TN) for the main class, as well as the true predictions (FP) and false predictions (FN) for the rest classes [1,6,12]. As in this work, we used four very famous performances scientifically: the accuracy of Eq. (19), the Kappa-score of Eq. (20), the F1-score of Eq. (22), and the sum of errors of Eq. (23).

$$\text{AC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}} \tag{19}$$

$$\kappa = \frac{\text{AC} - P_e}{1 - P_e} \tag{20}$$

where,

$$P_e = \frac{(\text{TP} + \text{FP})(\text{TP} + \text{FN}) + (\text{TN} + \text{FP})(\text{TN} + \text{FN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})^2} \tag{21}$$

$$\text{F1} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{ZOL}} \tag{22}$$

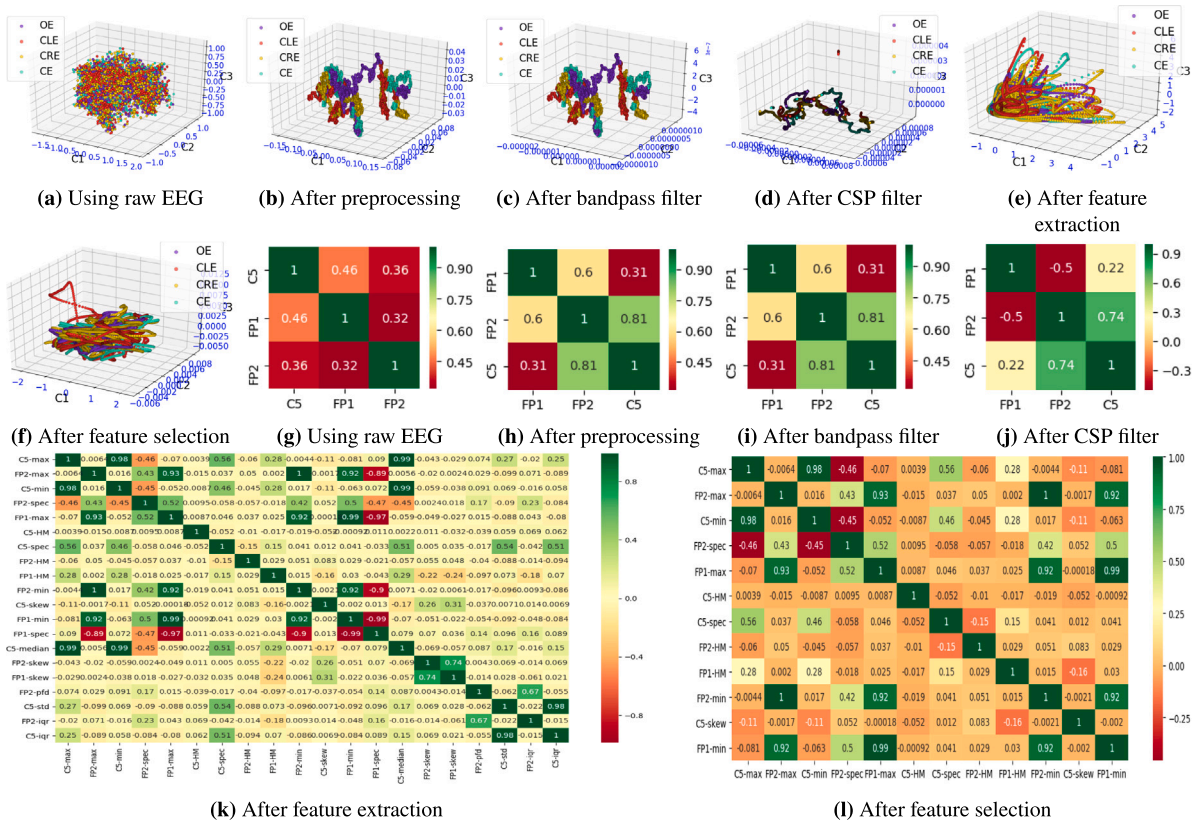
$$\text{ZOL} = \text{FP} + \text{FN} \tag{23}$$

4. Results and discussion

All steps of preprocessing, band-pass filtering, spatial filtering, feature extraction and selection, classification, and optimization was carried out in python language on HP-PC with 6 Gb of RAM and a Windows-10 exploitation system.

4.1. EEG data analysis

Fig. 5 illustrates spatial representations and correlation matrices of EEG data deformation during each global system step, such as Fig. 5(a) presents the data after the acquisition step using our EEG device. Where in Fig. 5(a), we notice that the data are randomized and non-separable, which shows that these signals predictions will have poor accuracy in the order of 50%. To improve all signals' quality, we applied Eq. (1) presented in a real-time preprocessing step to make the signals more stable, which gives the new data form in Fig. 5(b). So, Fig. 5(b) shows that each cluster sample is glued in a linear form. This form shows that samples closest to each sample are from the same class. Therefore, the preprocessing step plays an important role in improving the system's accuracy. Furthermore, the correlation matrix in Fig. 5(h) shows that there is an electrical link between signals of electrodes FP2 and C5 and also between FP2 and FP1. On the other hand, electrodes FP1 and C5 have no electrical connection. These bindings are normal because of the nervous system distribution for visual signals in the brain, as illustrated in Fig. 2(c). Also, Fig. 5(c) shows a data representation after optimizing the bandpass filter, such that the data shape remains similar to that of Fig. 5(b) and also the correlation matrix remains stable after filtering, which shows that the bandpass filter does not affect the signals' spatial characteristics, therefore all EEG characteristics remain in the optimal beta wave bandwidths obtained using SCA. So, to improve the prediction quality, our work uses the CSP filter to increase the variance level between all EEG channels. In this way, the data form is returned as more linear and more condensed. Also, the correlation matrix in Fig. 5(j) shows that there is only one strong correlation between FP2 and C5 electrodes. This connection represents the nerve that binds between the right eye and the left brain part to the occipital lobe, which shows that the right eye movement has a greater effect than the left eye. Therefore, both the "OE vs. CRE" and "CLE vs. CE" classifications have greater accuracy value classifications than "OE vs. CLE" and "CRE vs. CE". This situation can be reversed by moving the channel C5 position to C6. So all that remains is to extract temporal, frequency, and spatial features from each active channel. As Fig. 5(e) shows that the data are distorted into a linear form, which increases the independence between EEG tasks, we also notice that the data have an origin represented at the left of this figure, which indicates the existence of mutual characteristics between features. These characteristics are determined using the correlation matrix in Fig. 5(k), such as we notice a set of groups show a strong correlation between features such as (C5-max, C5-min, C5-median), (C5-iqr, C5-std), and (FP2-max, FP2-min, FP1-max, FP1-min). On other



(a) Using raw EEG (b) After preprocessing (c) After bandpass filter (d) After CSP filter (e) After feature extraction

(f) After feature selection (g) Using raw EEG (h) After preprocessing (i) After bandpass filter (j) After CSP filter

(k) After feature extraction (l) After feature selection

Fig. 5. Improvement of EEG data spatial form from the acquisition to the feature selection stage (SAES: subject 9).

hand, we notice all medium correlations groups such as (C5-max, C5-spec), (C5-min, C5-spec), (FP2-spec, FP2-min), (FP2-min, FP1-spec), (FP1-max, FP2-spec), (FP2-spec, FP1-max), (C5-median, C5-spec), (C5-spec, C5-max), (C5-spec, C5-min), (C5-spec, C5-iqr), (C5-spec, C5-std), (FP1-skew, FP2-skew), and (FP2-pfd, FP2-iqr). Note that strong correlations are between variations of 'max' and 'min' and between variations of 'iqr' and 'std'. But the totality of medium correlation is between 'spec' values and other variations of 'iqr', 'std', 'max', 'min', and 'median', and they are strongly correlated to each other, which shows temporal and spatial characteristics are related by spectral domain. As in real-time applications, we cannot use all 33 features for prediction. Therefore, our work includes a part on feature selection, which makes it possible to minimize the number of features (NFET) as much as possible according to the system accuracy value, as shown in Fig. 5(f) that illustrates the best 12 extracted features spatial representation, such that the data shapes remain linear, as in the case of Fig. 5(e), and a correlation matrix between features has been shown in Fig. 5(l), in such a way that the best features are the variations of 'max', 'min', 'spec', and 'HM', where Hjorth-Mobility variations are independent with other features. This means that strong correlation coefficients do not have much effect on EEG classification, as several researchers think.

4.2. Features selection results

Fig. 6 illustrates the matrix of importance degrees in (%) for each feature according to each subject and each classification, where these degrees are sorted in a decreasing direction from the left (yellow) to the right side (purple). This figure shows that each feature's importance degree is independent of the type of classification. In this matrix, we note that each feature extracted from FP2 electrode is more important than those extracted from FP1 and C5 electrodes because FP2 belongs to a different area than the other electrodes, and that degrees are few dependent on subject type because of the brain activity quality variation between various subjects. Also, this figure shows that several features (NFET) can be minimized based on the results in this matrix. So, all features of kurtosis, skewness, and inter-quartile can be eliminated if the accuracy value can be increased, so there are only 24 essential features left to detect eye state. Fig. 7 shows an evaluation of the number of features (NFET) that can be selected according to system accuracy values using LGBM algorithm with $L_r = 0.64$ (without using SCA). So, Fig. 7(a) illustrates accuracy value evaluation for each binary classification case, and Fig. 7(b) for the multiclass case. In the binary case, we notice that accuracy values are low when using a lower NFETs values than 5 features (where if NFET = 4, we select only the data of the four optimal features during each classification according to the importance degrees from Fig. 6, which are the variation of min and max extracted from FP2 and FP1 channels). But when NFETs are between 9 and 25, the average accuracy value (in brown) remains between 89.5% and 90%, and when NFETs exceed 26 features, accuracy decreases to less than 89.5%. While the effective NFET for all classifications is between 9 and 25, according to this subfigure, each classification has a specific maximum NFET value in triangle sign, where the classification "OE vs. CLE" (purple) has an optimal value of NFET = 9, "OE vs. CRE" (red) has a value of NFET = 21, "OE vs. CE" (yellow) has a value of NFET = 25, "CLE vs. CRE" (blue) has a value of NFET = 15, "CLE vs. CE" (green) has a value of NFET = 22, and "CRE vs. CE" (pink) has a value of NFET = 11. These NFET maximum values show that there is a precise number of NFETs corresponding to the best accuracy values during each classification, so when using false NFET values, the accuracy is decreased, which shows that the evaluation of NFET is a very interesting step. Similarly to the multiclass case in Fig. 7(b), the optimal NFET values are between 14 and 33, with a maximum accuracy value of NFET = 25. Thus, when the average precision values for each classification or the average in brown are almost constant at the end of the NFET value range, the minimum value of NFET is chosen because of the size of features extracted for training (classification) and testing data (prediction). This process automatically increases the speed of classification, optimization, and prediction steps.

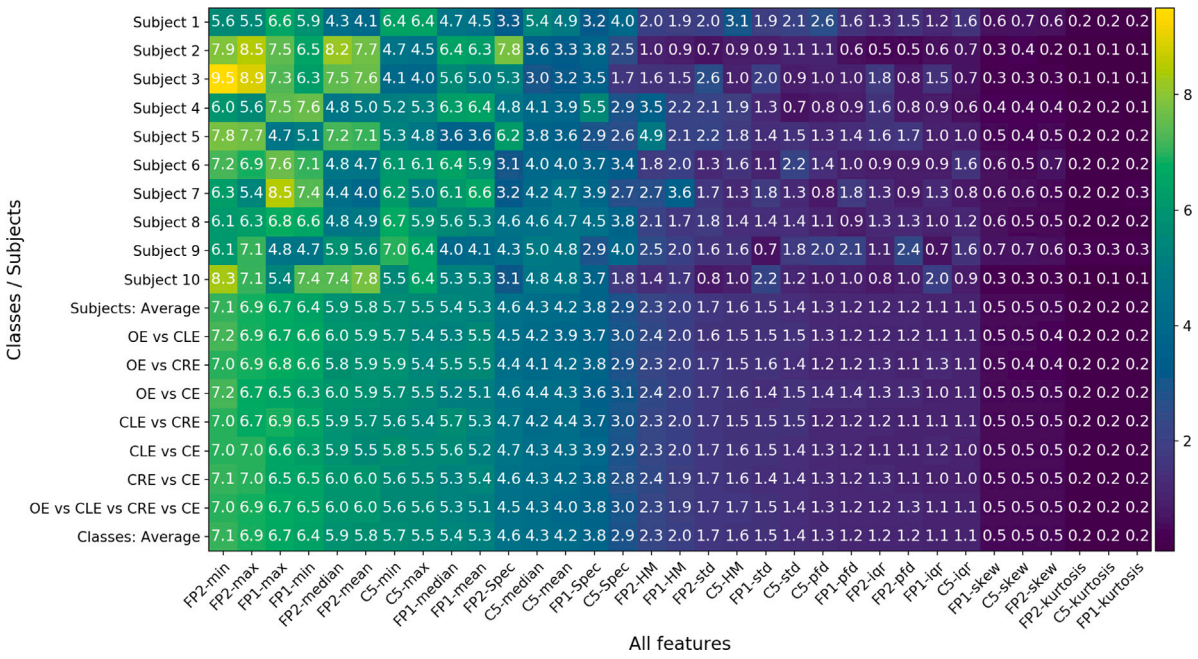
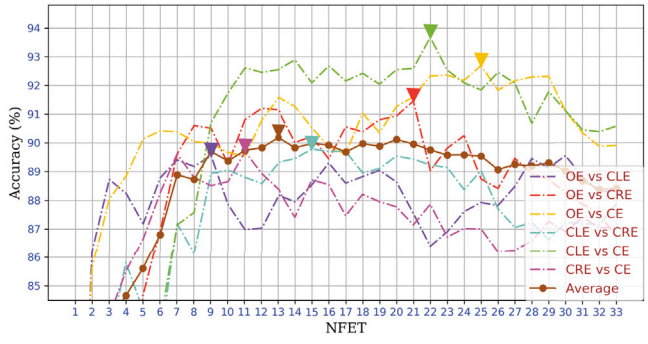
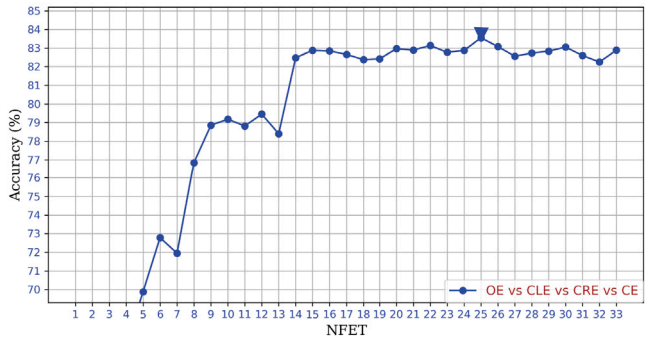


Fig. 6. Matrix representative of all importance degrees for all 33 features extracted from each subject and for each classification (SEAS dataset).



(a) For binary classifications



(b) For multiclass classification

Fig. 7. Evaluation of NFET value for all classifications based on the matrix of importance degrees.

4.3. Classification results during each system step

Table 1 shows an example of bandpass filter optimization results in lower frequency bandwidth than 30 Hz, such that SCA search optimal values of f_b and f_h allow finding the best system accuracy state during the binary classification "OE vs. CE" using subject-10 data. As this table shows, the accuracy value increased from 92.28% to 94.97%

Table 1
Example of the bandpass filter optimization results for the classification 'OE vs. CE' using SCA algorithm (SAES: subject-10).

f_b	f_h	AC (%)	T_o (s)
12.574863	14.429753	92.38	8.63
11.540733	26.773369	93.47	9.99
17.928963	18.989627	94.08	19.04
15.552641	27.133325	94.29	36.83
6.591234	24.055781	94.90	85.31
14.597291	21.073237	94.97	112.20

in 2 min and an optimal bandwidth between 14.6 Hz and 21.07 Hz, which belongs to the beta waves area. In the same way, for all other classifications and subjects using SAES dataset, the best bandwidths belong to beta waves as shown in Table 2. So, Table 2 shows the classification performances during different prediction system steps, as illustrated in Fig. 1(a), where according to Fig. 1(a), the optimization process is started from the bandpass filter step, so that in Table 2, T_o for raw EEG and preprocessing is only the sum between the prediction time (T_p) and the classification time (T_c) which is lower than 3 s as illustrated in Table 3. Before optimization and using raw EEG that was acquired using 3 channels. Although it is impossible to predict any task accurately using raw signals, the system accuracy level remains at 50% for all binary classifications and 28.74% for the multiclass case. However, after applying the preprocessing approach and during bandpass filter optimization, accuracy values are sharply raised to the order of 90% in the binary case and to the level of 77% in the multiclass case, showing the possibility of predicting EEG signals with a high level of accuracy. Then, SCA found that all signal characteristics belong to beta wave bandwidth (between 15 Hz and 30 Hz). After extracting all 11 features from each EEG channel, SCA found that the best LGBM parameter values (L_r) are between 0.88 and 0.92, which shows that data are spatially expanded and all classes are easily separable between them, as the system accuracy is increased by up to 92% when using all 33 features and bandwidths that have been obtained previously. The fifth step in this table corresponds to the optimization of all steps at the same time using SCA, such that SCA searches for the combination of all best parameters that allows us to predict EEG tasks with optimal accuracy values. So, we notice almost the stability between the results of

Table 2
Average test performances (leave 20% of trials out for each subject as a testing data) during each improvement step and using the global prediction system (using SAES dataset).

Steps	Classes	System parameters				System performances				
		f_b	f_h	NFET	L_r	AC (%)	F1 (%)	κ (%)	ZOL	T_o (s)
Before improvement	OE vs. CLE	nan	nan	3	0.64	53.16	52.40	9.00	1395.44	<3
	OE vs. CRE	nan	nan	3	0.64	53.51	52.47	8.39	1384.67	<3
	OE vs. CE	nan	nan	3	0.64	53.80	52.69	8.82	1369.22	<3
	CLE vs. CRE	nan	nan	3	0.64	51.51	50.77	5.43	1446.89	<3
	CLE vs. CE	nan	nan	3	0.64	52.49	51.49	6.23	1410.67	<3
	CRE vs. CE	nan	nan	3	0.64	50.26	49.51	3.14	1476.89	<3
	OE vs. CLE vs. CRE vs. CE	nan	nan	3	0.64	28.74	28.47	6.33	4238.33	<3
	Average	nan	nan	3	0.64	52.45	51.55	6.84	1413.96	<3
Improvement using real time preprocessing	OE vs. CLE	0.5	30	3	0.64	90.19	88.65	77.43	253.11	<3
	OE vs. CRE	0.5	30	3	0.64	94.05	92.79	85.63	159.22	<3
	OE vs. CE	0.5	30	3	0.64	89.42	88.37	77.06	262.33	<3
	CLE vs. CRE	0.5	30	3	0.64	87.72	86.14	72.60	354.89	<3
	CLE vs. CE	0.5	30	3	0.64	90.64	90.00	80.17	252.89	<3
	CRE vs. CE	0.5	30	3	0.64	88.32	87.49	75.23	322.78	<3
	OE vs. CLE vs. CRE vs. CE	0.5	30	3	0.64	77.98	76.60	69.68	1181.56	<3
	Average	0.5	30	3	0.64	90.06	88.91	78.02	267.54	<3
Optimization of the bandpass filter	OE vs. CLE	20.94	27.35	3	0.64	89.12	88.15	76.40	323.78	117.40
	OE vs. CRE	15.42	24.19	3	0.64	95.00	94.33	88.69	148.67	62.68
	OE vs. CE	20.53	27.84	3	0.64	91.24	90.42	80.96	259.67	59.03
	CLE vs. CRE	14.84	23.55	3	0.64	86.10	84.92	70.10	414.11	74.61
	CLE vs. CE	19.48	29.00	3	0.64	92.88	92.10	84.26	211.22	84.10
	CRE vs. CE	15.12	27.27	3	0.64	87.61	86.49	73.11	367.78	51.98
	OE vs. CLE vs. CRE vs. CE	18.73	27.82	3	0.64	73.65	72.67	64.04	1566.67	52.62
	Average	17.72	26.53	3	0.64	90.33	89.40	78.92	287.54	74.97
Optimization of the LGBM classifier during feature extraction	OE vs. CLE	20.94	27.35	33	0.88	92.62	91.73	83.59	242.44	48.90
	OE vs. CRE	15.42	24.19	33	0.91	95.57	95.27	90.58	115.56	31.13
	OE vs. CE	20.53	27.84	33	0.90	94.32	94.16	88.39	139.00	57.98
	CLE vs. CRE	14.84	23.55	33	0.92	86.33	85.51	71.58	480.22	36.63
	CLE vs. CE	19.48	29.00	33	0.90	91.78	91.37	82.79	287.33	35.23
	CRE vs. CE	15.12	27.27	33	0.88	93.20	93.06	86.21	178.22	28.50
	OE vs. CLE vs. CRE vs. CE	18.73	27.82	33	0.88	82.39	81.79	75.97	834.78	36.14
	Average	17.72	26.53	33	0.90	92.30	91.85	83.86	240.46	39.73
Optimization of the whole global system using optimal NFET values	OE vs. CLE	20.84	24.70	11.00	0.88	95.19	94.52	89.08	141.89	1617.98
	OE vs. CRE	20.63	25.04	16.89	0.93	96.80	96.29	92.60	94.22	1169.12
	OE vs. CE	19.14	24.60	15.33	0.90	93.35	92.38	84.83	195.22	1387.67
	CLE vs. CRE	21.01	27.65	12.44	0.88	91.46	90.56	81.31	252.56	1270.90
	CLE vs. CE	19.74	25.66	13.22	0.90	92.30	91.24	82.55	226.89	1018.97
	CRE vs. CE	20.83	22.98	13.56	0.87	87.69	86.72	73.93	362.44	1224.00
	OE vs. CLE vs. CRE vs. CE	19.04	23.94	16.44	0.91	83.13	82.67	76.89	999.00	2852.07
	Average	20.36	25.11	13.74	0.89	92.80	91.95	84.05	212.20	1281.44

Table 3
Comparative of average test performances for the classification (OE vs. CE) using CNN, LGBM, RF, and XGB during the improvement using real-time preprocessing step (SAES dataset).

Classifier	AC (%)	F1 (%)	κ (%)	T_c (s)	T_p (s)
RF	92.05	91.17	82.51	1.94	0.09
LGBM	90.65	89.72	79.67	2.69	0.01
XGB	89.95	88.99	78.34	5.36	0.02
CNN	69.82	47.00	8.96	26.21	0.18

sequential optimization (1st to 4th step) and parallel optimization (5th step in Table 2). Considering that bandwidths belong to beta waves, average NFET values are between 11 and 17 for all classifications. Also, LGBM's L_r values are between 0.87 and 0.93, which means that the most important thing about this step is that accuracy has been increased to 83.13% in the multiclass state, which means that this system has been developed to allow detecting baseline eye states easily using only our EEG device.

4.4. Comparison of LGBM performance with most popular EEG signal classifiers

Table 3 shows a comparison between the performances of the most popular classifiers after the baseline eye state classification using SAES dataset in order to choose the best algorithm, where accuracy is the important qualifier for any comparison between various classifiers during testing, also the time consumed by the classification step using training data (T_c) [6,12], and the duration for the prediction using testing data (T_p) by each classifier. Through this table, we note that accuracy values changed from one algorithm to another, with accuracy results ranked successively at 92.05%, 90.65%, 89.95%, and 69.85% collected when using RF, LGBM, XGB, and CNN [12,21,36,50], respectively. It should also be noted that during CNN classification, we use 5 epochs for the optimization using Adam. The model is built using 3 layers of Conv1D, 3 MaxPooling1D, one layer of BatchNormalization, and 4 layers of Dropout. At the model end, we use Flatten and 2 Dense-FC layers. But the best algorithm that can be used in this system, as shown Table 3 is RF (AC = 92.05%). However, by the preference virtue of using optimizers to improve the classification performance, LGBM and XGB algorithms are considered to be preferred because they have key parameters that directly control the regression tree (GBRT) inversely to RF, which has no key parameters to improve the accuracy level. From the most important parameters for GBM algorithm, we have the number of estimators that determines the maximum classification development, which often exceeds the value of 50. That is also the number of trees, which is indicated by m in Table 3. Also, we have L_r parameter, which controls the weight of new trees and dropped trees inversely, such as when decreasing L_r value, the weight and dropped ratios increase, and the number of estimators must be increased Table 3. Also, the

Table 4
Average performances for the classification using raw EEG (PhysioNet and UCI datasets).

Dataset	Tasks	NEEG	AC (%)	F1 (%)	κ (%)	ZOL
UCI	OE vs. CE	14	73.70	70.04	40.97	781
	OE vs. CE	3	52.86	49.74	4.25	1400
PhysioNet	L vs. R	64	65.03	64.88	30.22	4118.22
	L vs. R	3	51.60	51.42	3.52	5699.67
	OE vs. CE	3	62.72	60.97	24.89	1455.44
	LR vs. F	3	50.77	50.72	2.17	5797.56

number of leaves that control the tree model complexity, where this parameter must be lower than the square of max-depth (as default = 13) to avoid overfitting. So, in this work, SAES dataset is classified using default parameters, where the number of leaves is 175, the learning rate value is 0.64 and the number of estimators is 100. The issue of classification and prediction speeds has also resolved the choice because LGBM outperforms other algorithms. This makes the prediction system fast, and then the acquisition device frequency can correspond to the global system's prediction speed when adding the prediction system. The prediction speed using the NeuroHeadSet device (UCI) will therefore be 128 predictions per second. Also, using the BCI2000 device (PhysioNet), the prediction speed will be 160 predictions per second.

4.5. Test of the system using other datasets

Table 5 contains the classification results using UCI and PhysioNet datasets by the system, where the system uses bandpass and CSP filters, classification and best features selection values to improve real-time prediction accuracy as illustrated in Fig. 1(a). We note that accuracy increases to more than 99% instead of 73.7% in Table 4, when using all 14 channels' data for the "OE vs. CE" classification with an optimal value of NFET = 19 and an accuracy value of 99.29%. So, after reducing the number of electrodes (NEEG) to only 3 channels (AF3, AF4, and FC5), the accuracy value increased to 99.48%, where the optimal value of NFET is 5 features. Furthermore, this table shows that L_r values are dependent only on the acquisition system and not on the prediction system. In the same way, optimal bandwidth remains at the beta waves level when using UCI dataset. On the other hand, when using PhysioNet data after reducing the NEEG value from 64 to only 3 channels (AF3, AF4, and FC5), accuracy values also rose to more than 95% instead of 62.72% using raw EEG in Table 4, then the system efficiency that adapts to other different EEG signal acquisition devices. As in all tasks, we observe stability in the filter range between 20 Hz and 27 Hz. This range also belongs to beta waves, as evidence that the most important visual cortex and motor imagery features belong to this wave after the real-time preprocessing step. Also, when using NFET values at 12 and 13, the data contains only four ordinary features due to max, min, median, and mean variations extracted from each channel, which means that the remaining seven features can be eliminated while using the system in real-time permanently. On the other hand, we note that the values T_o are proof of the system's ability to improve its quality within 3 min above. The duration T_o includes the number of optimization cycles referred in Fig. 4 by (t), where each cycle includes all steps of preprocessing, processing, NFET evaluation, classification, and prediction. That is, each cycle is estimated at 120 s of the optimization period, so most of this time goes to the NFET evaluation stage, as confirmed by the results in Tables 1 and 2.

4.6. Test of the system using cross-subject

Table 6 contains SAES and PhysioNet data classification results, using single-subject data (using LeaveOneOut[7] function) to test LGBM-built models, and other subjects' data is used as training data for the classification step (Fig. 1(b)), indicating that the model is tested using strange data as illustrated in Fig. 1(b). According to this table, we note that all accuracy values are more than 94% to 99%. Also, parameter optimization and feature selection results are similar or higher than those obtained during test validation using trials' data, where for each subject we allocate 20% of its trials for prediction and model evaluation. This certainly confirms EEG's non-stationary nature elimination, as previously confirmed by feature selection results in Fig. 6, where importance degree results remain stable when subjects change, which means that the signal is not affected by the confusion occurring inside and outside of different subjects' brains. The more results are shown, the system accuracy is not affected by the quality of the classification tasks or the performance of the acquisition devices, where the system's ability to work perfectly with any subject or acquisition device as it was (perfect system). On the other hand, as expected, the optimization time (T_o) is doubled when using cross-subject during testing due to the consumed data amount doubling, where training data was doubled by 9 times during the classification and NFET evaluation, and testing data was doubled 5 times during the prediction step, and 10 times during the preprocessing and processing stages (9 subjects as training data + 1 subject out as testing data) as illustrated in Fig. 1(b). This means that the improvement will cost at least 6 min compared to test results obtained by trials.

4.7. Comparative results with related work

Table 7 shows a comparison between our prediction system results and literature works, such that accuracy results in related works are lower than 98.5% found by Ketu and Mishra [41]. Then, our work builds a new improved EEG prediction system based on essential steps from the acquisition to the prediction step, where the parameters of all steps are adaptable using SCA. This work found accuracy values higher than 99% using UCI dataset, as the strengths of this system are real-time preprocessing and self-optimization of the bandpass filter. In such a way, these two steps are developed to make samples of the same class correlate with each other, which shows good separation between clusters' data, as illustrated in Figs. 5(b) and 5(c). While remaining steps like feature extraction and feature selection are used to stabilize the quality of the classifications, which is shown in Table 2. Therefore, Table 7 contains EEG signal classification results for baseline eye state and motor imagery in related works, where each work uses the UCI or PhysioNet dataset to show the effectiveness of its own method. Then our research results in Table 7 are the best ever, whether the dataset, the number of channels, or the test-validation method. This shows the flexibility of our prediction system components. For accuracy values exceeding only 90% by classifying raw EEG signals in related work, it is not entirely possible to use testing data either by trials or by subject, because of EEG's non-stationary nature, but the testing data can be circumvented using shuffle[8] function, because shuffle function enables to mix between different tasks data (samples not trials), which evenly distributes important EEG characteristics between training and testing data, this helps the prediction model knowing tasks with high accuracy using only raw EEG, but when using new subjects as testing data, the accuracy can decrease to the right level (maybe 50%) because of the originality of non-stationary nature in EEG signals. However, this procedure used in much-related work to shuffle samples before classification is unrealistic and does not simulate the use of the prediction device to recognize EEG-task using new testing data in real-time, as confirmed by the accuracy values obtained when using cross-subject data in both works of Dose et al. [21] and Lun et al. [49], where the accuracy level is decreased from 95.76%, 86.49%, and 79.2% to 66.45%, 80%, and 73% using data acquired from 64

[7] https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.LeaveOneOut.html.

[8] https://scikit-learn.org/stable/modules/generated/sklearn.utils.shuffle.html.

Table 5
Average test performances (leave 20% of trials out for testing) for the global prediction system using other available datasets like UCI-eye-state and PhysioNet-MI.

Dataset	NEEG	Tasks	System parameters				System performances				
			f_b	f_h	NFET	L_r	AC (%)	F1 (%)	κ (%)	ZOL	T_o (s)
UCI	14	OE vs. CE	23.59	29.85	19	0.96	99.29	99.02	98.03	19	6821.39
	3	OE vs. CE	21.93	27.65	5	0.96	99.48	99.28	98.57	14	633.72
PhysioNet	3	OE vs. CE	21.43	25.50	12.67	0.88	97.78	97.14	94.33	68.89	222.45
	3	L vs. R	20.23	25.26	13.00	0.92	96.53	96.32	92.65	96.67	156.27
	3	LR vs. F	19.74	24.41	12.11	0.88	95.19	94.94	89.93	133.78	156.81

Table 6
Test performances for the result prediction system using cross-subject (by leave-one-subject-out for testing: LOSOXV), with NEEG = 3.

Dataset	Tasks	System parameters				System performances				
		f_b	f_h	NFET	L_r	AC (%)	F1 (%)	κ (%)	ZOL	T_o (s)
PhysioNet	LR vs. F	17.59	19.86	10	0.85	99.50	99.43	98.86	20	501.47
	L vs. R	15.12	27.15	12	0.86	92.05	91.37	82.86	318	497.69
	OE vs. CE	22.82	26.27	10	0.89	94.45	93.88	87.81	222	498.03
SAES	OE vs. CE	17.94	20.58	11	0.80	97.04	96.98	93.96	129	460.90
	OE vs. CLE	20.82	26.70	14	0.89	99.71	99.70	99.41	17	556.71
	OE vs. CRE	17.03	21.77	17	0.90	95.76	95.75	91.51	226	407.00
	CLE vs. CRE	15.37	20.57	14	0.89	98.33	98.31	96.63	89	402.69
	CLE vs. CE	18.41	23.10	10	1.00	94.61	94.47	88.96	235	356.16
	CRE vs. CE	16.83	22.19	11	0.87	94.99	94.99	90.00	188	426.08

Table 7
Comparative results with related work for binary eye and motor imagery tasks classifications.

Work	Dataset	Test-validation	Tasks	NEEG	AC (%)	Method
Chhabra et al. [45]	UCI	By trials	OE vs. CE	14	92.43	RF
Hamilton et al. [43]	UCI	By trials	OE vs. CE	14	97.4	K*-RRF
	UCI	By trials	OE vs. CE	14	97.3	K*
	UCI	By trials	OE vs. CE	14	97.2	ada(RJ48F)
	UCI	By trials	OE vs. CE	14	95.1	RRF
Hassan et al. [46]	UCI	By trials	OE vs. CE	14	89.2	ANN
Hussien et al. [15]	UCI	By trials	OE vs. CE	14	96.50	KNN-bMGWO
Ketu and Mishra [41]	UCI	By trials	OE vs. CE	14	98.5	Hybrid Model
Reddy and Behera [44]	UCI	By trials	OE vs. CE	14	94.92	MLN
Saghafi et al. [47]	UCI	By trials	OE vs. CE	14	88.2	Detlta + Theta-LR
	UCI	By trials	OE vs. CE	14	82.4	Detlta + Theta-ANN
	UCI	By trials	OE vs. CE	14	70.6	Detlta + Theta-SVM
Zhou et al. [42]	UCI	By trials	OE vs. CE	14	91.9	Delta-RF
Dose et al. [21]	PhysioNet	By trials	L vs. R	64	86.49	CNN
	PhysioNet	By trials	L vs. R	14	82.66	CNN
	PhysioNet	By trials	L vs. R	3	79.20	CNN
	PhysioNet	By subject	L vs. R	64	80.38	CNN
	PhysioNet	By subject	L vs. R	14	76.66	CNN
	PhysioNet	By subject	L vs. R	3	73.20	CNN
Kim et al. [48]	PhysioNet	By trials	L vs. R	14	80.05	SUT-CCSP-RF
Lun et al. [49]	PhysioNet	By trials	L vs. R	64	95.76	CNN
	PhysioNet	By subject	L vs. R	64	66.45	CNN
This work	UCI	By trials	OE vs. CE	14	99.29	OBPF-LGBM-SCA
	UCI	By trials	OE vs. CE	3	99.48	OBPF-LGBM-SCA
	PhysioNet	By trials	L vs. R	3	96.53	OBPF-LGBM-SCA
	PhysioNet	By subject	L vs. R	3	92.05	OBPF-LGBM-SCA
	PhysioNet	By trials	OE vs. CE	3	97.78	OBPF-LGBM-SCA
	PhysioNet	By subject	OE vs. CE	3	94.45	OBPF-LGBM-SCA
	SAES	By trials	OE vs. CE	3	93.35	OBPF-LGBM-SCA
	SAES	By subject	OE vs. CE	3	97.04	OBPF-LGBM-SCA

and 3 channels, respectively. Conversely to our work, which started by removing the EEG's non-stationary in the preprocessing step, where accuracy values increased significantly from 50% to more than 92% in both test-validations by trails and by subject, in a simulated way to the use of the system in real-time, that automatically decreases the features' variance between all subjects during training with all new subjects used during testing according to the results in Table 6.

Also, by comparing performance results of CNN with other machine learning algorithms in Table 3, we observe a high difference in accuracy of 20%, so CNN algorithm is not preferred to classify EEG tasks in this work, whether baseline eye state or motor imagery, so the selection of the classification algorithm phase is important to build the best-integrated prediction systems. Also, the feature extraction stage is missed in many works due to a lack of necessary experience to play with EEG features, and then discover distinctive features for each high-quality EEG task. In this work, we conceive that the variations of min, median, max and mean contain the most important EEG task characteristics after removing the non-stationary nature, and all related works without exception decrease their results at the moment when changing the test-validation method from trials to subject, because of the vulnerability of related methods. Inversely, ours is not affected by the change in the test-validation method. This shows the robustness of the system formed during this research, highlighting the ability of our system to process all EEG classifications, whether for control

Table A.1
Equations used to extract all features from each channel (c) after the use of CSP filtering.

Feature	Equation
Mean	$\text{mean}(X^c_{\text{CSP}}[t]) = \frac{1}{f_s}\sum_{i=t-f_s}^{t}\left(X^c_{\text{CSP}}[i]\right)$
Kurtosis	$\text{Kurtosis}(X^c_{\text{CSP}}[t]) = \frac{1}{f_s}\frac{\sum_{i=t-f_s}^{t}\left(X^c_{\text{CSP}}[t]-\text{mean}(X^c_{\text{CSP}}[t])\right)^4}{\left(\text{std}(X^c_{\text{CSP}}[t])\right)^4}$
Standard Deviation	$\text{std}(X^c_{\text{CSP}}[t]) = \sqrt{\frac{1}{f_s}\sum_{i=t-f_s}^{t}\left(X^c_{\text{CSP}}[i]-\text{mean}(X^c_{\text{CSP}}[t])\right)^2}$
Hjorth Mobility	$\text{HM}(X^c_{\text{CSP}}[t]) = \sqrt{\frac{\text{var}\left(\frac{dX^c_{\text{CSP}}[t-f_s:t]}{dt}\right)}{\text{var}(X^c_{\text{CSP}}[t-f_s:t])}}$
Skewness	$\text{skew}(X^c_{\text{CSP}}[t]) = \frac{1}{f_s}\frac{\sum_{i=t-f_s}^{t}\left(X^c_{\text{CSP}}[i]-\text{mean}(X^c_{\text{CSP}}[t])\right)^3}{\left(\text{std}(X^c_{\text{CSP}}[t])\right)^3}$
Median	$\text{median}(X^c_{\text{CSP}}[t]) = \text{sort}(X^c_{\text{CSP}}[t-f_s:t])\left[\frac{f_s+1}{2}\right]$, if n=len(X) is an odd number
	$\text{median}(X^c_{\text{CSP}}[t]) = \text{sort}(X^c_{\text{CSP}}[t-f_s:t])\left[f_s/2\right] + \text{sort}(X^c_{\text{CSP}}[t-f_s:t])\left[1+f_s/2\right]/2$, if n=len($X^c_{\text{CSP}}$) is an even number
Spectrum	$\text{Spec}(X^c_{\text{CSP}}[t]) = \text{mean}\left(\text{abs}\left(\text{fft2}(X^c_{\text{CSP}}[t-f_s:t])\right)\right)$
Petrosian Fractal Dimension	$\text{pfd}(X^c_{\text{CSP}}[t]) = \frac{\log(N)}{\log(N)+\log\left(\frac{N}{d/L}\right)}$, With N is the number of samples in a segment, d is the waveform's planar diameter, and L is its overall length
Inter-Quartile	$\text{iqr}(X^c_{\text{CSP}}[t]) = \text{median}(\text{sort}(X^c_{\text{CSP}}[t-f_s:t])[t-\frac{f_s}{2}:t]) - \text{median}(\text{sort}(X^c_{\text{CSP}}[t-f_s:t])[t-f_s:t-\frac{f_s}{2}])$

by or to diagnose brain-related diseases in future work with fewer channels. In addition, when increasing the size of training data provides an addition to improving the quality of all models that are equipped, this automatically increases the accuracy level, with the detriment of optimization time.

5. Conclusion

This work shows that it is possible to predict baseline eye states and motor imagery tasks using EEG signals with high quality, such that all signals are acquired using only 3 active dry electrodes positioned on the scalp. Then, the main contribution of our work is at the level of the pre-processing step to minimize EEG's non-stationary effect, such that this step allows suppressing all instabilities in amplitude and frequency over time that have been generated without any task activation. Moreover, our work uses SCA algorithm to get the best bandwidths that contain all characteristics of EEG tasks. Also, to increase the variance between channels and then to eliminate most mutual information between features, our system employed CSP algorithm in real-time to eliminate the totality of correlations between features. In addition, the feature extraction step plays an important role in increasing the presence of as many signal characteristics as possible. This step found accuracy values greater than 92%, and to show adaptation between system stages, the entire prediction system was applied to UCI dataset to find accuracy values greater than 99.2%, which is a much higher value than related work results that have a maximum accuracy of 98%. Therefore, these results show the robustness of the system built in this work. In our future work, we will develop new approaches and applications for real-time diagnostic and classification of several EEG tasks, using smaller devices.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix. Most performances and features equations used for data analysis and prediction results

See Table A.1

References

[1] S. Abenna, M. Nahid, A. Bajit, Brain-computer interface: A novel EEG classification for baseline eye states using LGBM algorithm, in: M. S., B. B. (Eds.), Digital Technologies and Applications, Vol. 211, Springer Science and Business Media LLC, ISBN: 978-3-030-73882-2, 2021, pp. 189–198, http://dx.doi.org/10.1007/978-3-030-73882-2_18, Chapter 18.

[2] R. Chatterjee, A. Datta, D.K. Sanyal, Ensemble learning approach to motor imagery EEG signal classification, in: Machine Learning in Bio-Signal Analysis and Diagnostic Imaging, Elsevier Inc., ISBN: 9780128160862, 2019, pp. 183–208, http://dx.doi.org/10.1016/B978-0-12-816086-2.00008-4, Chapter 8.

[3] X. Wang, W. Liu, P. Toiviainen, T. Ristaniemi, F. Cong, Group analysis of ongoing EEG data based on fast double-coupled nonnegative tensor decomposition, J. Neurosci. Methods (ISSN: 0165-0270) 330 (June 2019) (2020) 108502, http://dx.doi.org/10.1016/j.jneumeth.2019.108502.

[4] E. Dagdevir, M. Tokmakci, Determination of effective signal processing stages for brain computer interface on BCI competition IV data set 2b : A review study, IETE J. Res. (2021) http://dx.doi.org/10.1080/03772063.2021.1914204.

[5] W.O. Tatum, G. Rubboli, P.W. Kaplan, S.M. Mirsatari, K. Radhakrishnan, D. Gloss, M. Cook, S. Beniczky, Clinical neurophysiology clinical utility of EEG in diagnosing and monitoring epilepsy in adults, Clin. Neurophysiol. J. 129 (2018) 1056–1082, http://dx.doi.org/10.1016/j.clinph.2018.01.019.

[6] S. Abenna, M. Nahid, A. Bajit, Motor imagery based brain-computer interface: improving the EEG classification using delta rhythm and LightGBM algorithm, Biomed. Signal Process. Control 71 (11) (2022) http://dx.doi.org/10.1016/j.bspc.2021.103102.

[7] N.S. Malan, S. Sharma, Time window and frequency band optimization using regularized neighbourhood component analysis for multi-view motor imagery EEG classification, Biomed. Signal Process. Control (ISSN: 1746-8094) 67 (February) (2021) 102550, http://dx.doi.org/10.1016/j.bspc.2021.102550.

[8] C.D.G. Virgilio, J.H.A. Sossa, J.M. Antelis, L.E. Falcón, Spiking neural networks applied to the classification of motor tasks in EEG signals, Neural Netw. (ISSN: 1879-2782) 122 (2020) 130–143, http://dx.doi.org/10.1016/j.neunet.2019.09.037.

[9] S.Z. Zahid, M. Aqil, M. Tufail, M.S. Nazir, Online classification of multiple motor imagery tasks using filter bank based maximum-a-posteriori common spatial pattern filters, Irbm (ISSN: 18760988) 1 (2019) http://dx.doi.org/10.1016/j.irbm.2019.11.002.

[10] D.J. Mcfarland, J.R. Wolpaw, EEG-based brain-computer interfaces, Curr. Opin. Biomed. Eng. (ISSN: 2468-4511) 4 (2017) 194–200, http://dx.doi.org/10.1016/j.cobme.2017.11.004.

[11] M. Miao, W. Hu, H. Yin, K. Zhang, Spatial-frequency feature learning and classification of motor imagery EEG based on deep convolution neural network, Comput. Math. Methods Med. 2020 (2020) http://dx.doi.org/10.1155/2020/1981728.

[12] S. Abenna, M. Nahid, H. Bouyghf, Alcohol use disorders automatic detection based BCI systems: A novel EEG classification based on machine learning and optimization algorithms, Int. J. Inf. Sci. Technol. (ISSN: 2550-5114) 6 (1) (2022) 14–25.

[13] N. Sperelakis, Cable properties and propagation of action potentials, in: Cell Physiology, Elsevier Inc, 2012, pp. 325–343, http://dx.doi.org/10.1016/B978-0-12-387738-3.00018-4, Chapter 18.

[14] A. Ahmadi, A. Khorasani, V. Shalchyan, M.R. Daliri, State-based decoding of force signals from multi-channel local field potentials, IEEE Access (ISSN: 2169-3536) 8 (2020) 159089–159099, http://dx.doi.org/10.1109/ACCESS.2020.3019267.

[15] H.R. Hussien, E.-S.M. El-Kenawy, A.I. El-desouky, EEG channel selection using a modified grey wolf optimizer, Eur. J. Electr. Eng. Comput. Sci. 5 (1) (2021) 17–24, http://dx.doi.org/10.24018/ejece.2021.5.1.265.

[16] I.E. Kutepov, V.V. Dobriyan, M.V. Zhigalov, M.F. Stepanov, A.V. Krysko, T.V. Yakovleva, V.A. Krysko, Informatics in medicine unlocked EEG analysis in patients with schizophrenia based on Lyapunov exponents, Inform. Med. Unlocked (ISSN: 2352-9148) 18 (2020) 100289, http://dx.doi.org/10.1016/j.imu.2020.100289.

[17] M. Sharma, S. Patel, U.R. Acharya, Automated detection of abnormal EEG signals using localized wavelet filter banks, Pattern Recognit. Lett. (ISSN: 0167-8655) 133 (2020) 188–194, http://dx.doi.org/10.1016/j.patrec.2020.03.009.

[18] M. Sollychin, B.N. Jack, A. Polari, A. Ando, G.P. Amminger, C. Markulev, P.D. McGorry, B. Nelson, T.J. Whitford, H.P. Yuen, S. Lavoie, Frontal slow wave resting EEG power is higher in individuals at ultra high risk for psychosis than in healthy controls but is not associated with negative symptoms or functioning, Schizophr. Res. (ISSN: 15732509) 208 (2019) 293–299, http://dx.doi.org/10.1016/j.schres.2019.01.039.

[19] R. Bose, K. Samanta, S. Chatterjee, S. Bhattacharyya, A. Khasnobish, Motor imagery classification enhancement with concurrent implementation of spatial filtration and modified stockwell transform, in: K. Pal, A. Khasnobish, I. Banerjee, H.-B. Kraatz, S. Bag, U. Kuruganti (Eds.), Bioelectronics and Medical Devices, Elsevier Ltd, ISBN: 9780081024201, 2019, pp. 793–818, http://dx.doi.org/10.1016/B978-0-08-102420-1.00038-8, Chapter 31.

[20] A. Borowicz, Using a multichannel Wiener filter to remove eye-blink artifacts from EEG data, Biomed. Signal Process. Control (ISSN: 1746-8108) 45 (2018) 246–255, http://dx.doi.org/10.1016/j.bspc.2018.05.012.

[21] H. Dose, J.S. Møller, H.K. Iversen, S. Puthusserypady, An end-to-end deep learning approach to MI-EEG signal classification for BCIs, Expert Syst. Appl. 114 (2018) 532–542, http://dx.doi.org/10.1016/j.eswa.2018.08.031.

[22] M.A.S. ALTobi, G. Bevan, P. Wallace, D. Harrison, K.P. Ramachandran, Fault diagnosis of a centrifugal pump using MLP-GABP and SVM with CWT, Eng. Sci. Technol. Int. J. (ISSN: 2215-0986) 22 (3) (2019) 854–861, http://dx.doi.org/10.1016/j.jestch.2019.01.005.

[23] X. Ding, X. Yue, R. Zheng, C. Bi, D. Li, G. Yao, Classifying major depression patients and healthy controls using EEG, eye tracking and galvanic skin response data, J. Affect. Disord. (ISSN: 1573-2517) 251 (March) (2019) 156–161, http://dx.doi.org/10.1016/j.jad.2019.03.058.

[24] Y.U.N. Ju, G. Sun, Q. Chen, M.I.N. Zhang, A model combining convolutional neural network and LightGBM algorithm for ultra-short-term wind power forecasting, IEEE Access 7 (2019) 28309–28318, http://dx.doi.org/10.1109/ACCESS.2019.2901920.

[25] M. Deriche, S. Arafat, M. Siddiqui, Eigenspace time frequency based features for accurate seizure detection from EEG data, IRBM (ISSN: 1959-0318) 1 (2019) 1–11, http://dx.doi.org/10.1016/j.irbm.2019.02.002.

[26] P. Krauss, C. Metzner, N. Joshi, H. Schulze, M. Traxdorf, A. Maier, A. Schilling, Neurobiology of sleep and circadian rhythms analysis and visualization of sleep stages based on deep neural networks, Neurobiol. Sleep Circadian Rhyth. (ISSN: 2451-9944) 10 (March) (2021) 100064, http://dx.doi.org/10.1016/j.nbscr.2021.100064.

[27] S. Abenna, M. Nahid, H. Bouyghf, Sleep stages detection based BCI: A novel single-channel EEG classification based on optimized bandpass filter, in: R. Saidi, B. El-Bhiri, Y. Maleh, A. Mosallam, M. Essaaidi (Eds.), Advanced Technologies for Humanity, Vol. 110, Springer, Cham, 2022, pp. 96–105, http://dx.doi.org/10.1007/978-3-030-94188-8_10, ISBN 978-3-030-94187-1; 978-3-030-94188-8. Chapter 10.

[28] B.O. Hjorth, Technical contributions eeg analysis based on time domain properties, Electroencephalogr. Clin. Neurophysiol. (1970) 306–310.

[29] D. Jiang, Y.-n. Lu, Y. Ma, Y. Wang, Robust sleep stage classification with single-channel EEG signals using multimodal decomposition and HMM-based refinement, Expert Syst. Appl. 121 (2019) 188–203, http://dx.doi.org/10.1016/j.eswa.2018.12.023.

[30] Y. Guo, Y. Zhang, Z. Chen, Y. Liu, W. Chen, EEG classification by filter band component regularized common spatial pattern for motor imagery, Biomed. Signal Process. Control (ISSN: 17468108) 59 (2020) http://dx.doi.org/10.1016/j.bspc.2020.101917.

[31] H. Li, T. Liu, X. Wu, Q. Chen, Research on bearing fault feature extraction based on singular value decomposition and optimized frequency band entropy, Mech. Syst. Signal Process. (ISSN: 10961216) 118 (2019) 477–502, http://dx.doi.org/10.1016/j.ymssp.2018.08.056.

[32] P. Gaur, K. Mccreadie, R. Bilas, H. Wang, G. Prasad, An automatic subject specific channel selection method for enhancing motor imagery classification in EEG-BCI using correlation, Biomed. Signal Process. Control (ISSN: 1746-8094) 68 (December) (2021) http://dx.doi.org/10.1016/j.bspc.2021.102574.

[33] H. Zeng, C. Yang, H. Zhang, Z. Wu, J. Zhang, G. Dai, F. Babiloni, W. Kong, A lightgbm-based EEG analysis method for driver mental states classification, Comput. Intell. Neurosci. 2019 (2019) http://dx.doi.org/10.1155/2019/3761203.

[34] N. Michielli, U.R. Acharya, F. Molinari, Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals, Comput. Biol. Med. (ISSN: 0010-4825) 106 (January) (2019) 71–81, http://dx.doi.org/10.1016/j.compbiomed.2019.01.013.

[35] C.-t. Lin, A.P. Javier, M. Prasad, A hierarchical meta-model for multi-class mental task based brain-computer interfaces, Neurocomputing (ISSN: 0925-2312) (2019) http://dx.doi.org/10.1016/j.neucom.2018.07.094.

[36] S. Yang, B. Li, Y. Zhang, M. Duan, S. Liu, Y. Zhang, X. Feng, R. Tan, L. Huang, F. Zhou, Selection of features for patient-independent detection of seizure events using scalp EEG signals, Comput. Biol. Med. 119 (January) (2020) http://dx.doi.org/10.1016/j.compbiomed.2020.103671.

[37] S. Abenna, M. Nahid, A. Bajit, BCI: Classifiers optimization for EEG signals acquiring in real-time, in: 2020 6th IEEE Congress on Information Science and Technology (CiSt), IEEE, 2021, http://dx.doi.org/10.1109/CiSt49399.2021.9357209.

[38] M. Diykh, S. Abdulla, K. Saleh, R.C. Deo, Fractal dimension undirected correlation graph-based support vector machine model for identification of focal and non-focal electroencephalography signals, Biomed. Signal Process. Control 54 (101611) (2019) http://dx.doi.org/10.1016/j.bspc.2019.101611.

[39] S. Mirjalili, SCA: A Sine cosine algorithm for solving optimization problems, Knowl.-Based Syst. 96 (2016) 120–133, http://dx.doi.org/10.1016/j.knosys.2015.12.022.

[40] H. Pan, Z. Li, C. Tian, L. Wang, Y. Fu, X. Qin, F. Liu, The lightgbm-based classification algorithm for Chinese characters speech imagery BCI system, Cognit. Neurodyn. (ISSN: 1871-4099) 2020 (2022) http://dx.doi.org/10.1007/s11571-022-09819-w.

[41] S. Ketu, P.K. Mishra, Hybrid classification model for eye state detection using electroencephalogram signals, Cognit. Neurodyn. 16 (2022) 73–90, http://dx.doi.org/10.1007/s11571-021-09678-x.

[42] Z. Zhou, P. Li, J. Liu, W. Dong, A novel real-time EEG based eye state recognition system, in: Communications and Networking, Vol. 262, Springer, Cham, 2019, pp. 175–183, http://dx.doi.org/10.1007/978-3-030-06161-6_17.

[43] C.R. Hamilton, S. Shahryari, K.M. Rasheed, Eye state prediction from EEG data using boosted rotational forests, in: 2015 IEEE 14th International Conference on Machine Learning and Applications, 2015, http://dx.doi.org/10.1109/ICMLA.2015.89.

[44] T.K. Reddy, L. Behera, Online eye state recognition from EEG data using deep architectures, in: 2016 IEEE International Conference on Systems, Man, and Cybernetics, 2016, http://dx.doi.org/10.1109/SMC.2016.7844325.

[45] G. Chhabra, E. Studies, V. Sapra, E. Studies, R. Sharma, E. Studies, K. Joshi, Eye state classification based on EEG signals, Des. Eng. (ISSN: 0011-9342) (2021) 2544–2551.

[46] M.M. Hassan, M.R. Hassan, S. Huda, M.Z. Uddin, A. Gumaei, A. Alsanad, A predictive intelligence approach to classify brain-computer interface based eye state for smart living, Appl. Soft Comput. 108 (2021) 73–90, http://dx.doi.org/10.1016/j.asoc.2021.107453.

[47] A. Saghafi, C.P. Tsokos, M. Goudarzi, H. Farhidzadeh, Random eye state change detection in real-time using EEG signals, Expert Syst. Appl. (ISSN: 09574174) 72 (2017) 42–48, http://dx.doi.org/10.1016/j.eswa.2016.12.010.

[48] Y. Kim, J. Ryu, K.K. Kim, C.C. Took, D.P. Mandic, C. Park, Motor imagery classification using mu and beta rhythms of EEG with strong uncorrelating transform based complex common spatial patterns, Comput. Intell. Neurosci. 2016 (2016) http://dx.doi.org/10.1155/2016/1489692.

[49] X. Lun, Z. Yu, T. Chen, F. Wang, Y. Hou, A simplified CNN classification method for MI-EEG via the electrode pairs signals, Front. Human Neurosci. 14 (September) (2020) 1–14, http://dx.doi.org/10.3389/fnhum.2020.00338.

[50] H.K. Lee, Y.-S. Choi, Application of continuous wavelet transform and convolutional neural network in decoding motor imagery brain-computer interface, Entropy 21 (1199) (2019) 1–11, http://dx.doi.org/10.3390/e21121199.