Biomedical Signal Processing and Control 79 (2023) 104210

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

An enhanced motor imagery EEG signals prediction system in real-time based on delta rhythm

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)
Delta waves
Data analysis
Feature extraction
Feature selection
Machine learning
Optimization

ABSTRACT

This work aims to develop a brain–computer interface (BCI) system based on electroencephalogram (EEG) signals, that is capable of remote controlling rehabilitation systems using wireless connections. This system can extract delta waves from raw EEG in real-time to predict motor imagery (MI) tasks. Where we built a simple acquisition device that acquires EEG signals using three dry electrodes, these non-invasive channels are positioned on the scalp surface at the occipital and central lobes. After the acquisition step, we amplify the signals and remove permanent noise during the preprocessing step. Then, in the feature extraction step, we extract possible features from each channel. Then, we select only some important features at the feature selection step, by the calculation of each feature's contribution score. In the classification phase using machine learning algorithms, we select the light gradient boosting machine (LGBM) algorithm enhanced by the multi-verse optimization (MVO) algorithm, which enables the building of optimum prediction models. Also, this work employed a data analysis phase. Where to evaluate the characteristics independent between features at each step, we analysed the data using the correlation matrix results. As well as, we analysed the data changes temporally and spatially between MI tasks at each step. Therefore, the classification results indicated that the system accuracy score is over 90%. While in related work, we have an accuracy value ranging between 79% and 89%. These comparative results show the best quality of our system proposed for this work-based delta wave.

1. Introduction

BCI system is an electronic system, that can help to bridge the gap between people and computers. Where BCI allow the control of computers through changes in brain signal activities using several technologies such as EEG signals. With the fast growth of cognitive neuroscience, EEG signals with a wealth of physiological and pathological information are obtaining more and more attention in recent years [1]. Also, EEG signal is commonly employed in BCI because of its low cost of maintenance and great temporal resolution [1,2]. On the other hand, MI-related BCI was found to get the user's intent [3]. Where MI is a psychological condition in which a person performs the same action over and over again. For example, MI-EEG recordings of the left and right hands can be used to move a target or a cursor on a computer screen for patients with lost motor function. This control method using MI tasks provides a new way of communicating. So, Fukunaga was the first to describe common spatial patterns (CSP) [1,4,5], which is a supervised learning approach for classifying MI-tasks based on an extension of principal component analysis (PCA). Such as, MI-EEG may be used to create a simple binary answer to assist amyotrophic lateral sclerosis (ALS) patients in responding to certain queries [3,6]. There are diverse MI tasks that cause different activation states in the brain's sensorimotor cortex [7], where MI may be gathered in a convenient, non-invasive, and cost-effective manner. As a result, numerous studies have worked on MI EEG signal feature extraction and categorization.

In recent years many studies use wavelet transform (WT) [7,8] for feature extraction algorithms when recording brain activity. The EEG is commonly employed since it is non-invasive and less expensive than alternative techniques like magnetoencephalography (MEG). The acquired signal must then be pre-processed to improve the data quality and remove unnecessary samples. After the preprocessing step, feature extraction is used to extract important information from each EEG task data. Finally, pattern recognition algorithms are used to classify the data using the vector of all features. Then, the predicted label is sent to the command unit for control using for example a wireless link [9–11]. Moreover, EEG signals are non-stationary, just like other biological signals. This suggests that the linear and non-linear classification approaches should be developed to obtain the most efficient system possible [11,12]. The most crucial step in this work for correctly

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

https://doi.org/10.1016/j.bspc.2022.104210
Received 29 October 2021; Received in revised form 4 September 2022; Accepted 10 September 2022
Available online 23 September 2022

separating the data of all clusters is the feature extraction step [4,13,14]. Instead of using raw EEG data, this phase seeks to highlight the most important characteristics of EEG tasks through feature variations in real-time. Then, data characteristics can be distributed individually in several classes according to the variance between features of tasks. Also, the preprocessing step is critical to improving the classification accuracy level of EEG data. Where, preprocessing has a variety of objectives, including channels' signals segmentation, signals normalization, artefact removal, and frequency bandwidths selection [11,15,16]. This helps to clarify signal changes for analysis and provides high-quality data for processing and classification. The classification of EEG data is the most important component of this work. Since the classification enables the creation of real-time expectation models for each subject's tasks. There is a high number of machines and deep learning algorithms that can be used to classify the EEG data. But, the suggested algorithms for the system must identify changes as fast and maximum accuracy levels, allowing for real-time decision-making easily [17]. Where Ding et al. [17] introduced a vast variety of current machine learning algorithms for EEG signals classification. From most algorithms, we suggest some popular classifiers like support vector machine (SVM) [18,19], logistic regression (LR) [20], and random forest (RF) [21].

This work uses the delta wave, which is extracted using a bandpass filter as illustrated in the paper of Abenna et al. [22], that removes all small signals variation (high frequencies) that degrade the system stability for every MI-task in real-time. Furthermore, this work extracted different features that characterize tasks from every channel. To ensure the best system flexibility, we applied the LGBM algorithm to select important features and classify the data during the training stage. In addition, this work uses the MVO algorithm to improve the LGBM's learning-rate parameter according to each subject as illustrated in Fig. 1. The rest of this paper will be organized as follows: Section 2 presents the related work with our work, Section 3 illustrates materials and algorithms used to develop our system, while Section 4 analyses the results of the EEG acquisition device, feature extraction, feature selection, classification, and optimization technique that improved our prediction system performance. Section 5 illustrate conclusion and future work.

2. Related work

According to previous research, ensembles and hybrid classifiers are more effective than single classifiers. Where, Kim et al. [23] offer the weighted difference of power spectral density (WDPSD) as a useful feature extraction approach for binary class MI-based BCI systems that may reflect the time, frequency, and spatial properties. The WDPSD method begins with the calculation of EEG signal PSD matrices for all channels. Following that, based on non-stationary and class separability, the best channel pair is picked from all available channel couples. Finally, a weight matrix is calculated, that reflects the matrix of PSD differences in the non-stationary of the chosen channel couple. The weight matrix is used in the PSD difference matrix, to obtain the robust and adaptable characteristics. To categorize upper and lower limb MI movements, Bose et al. [24] suggested a unique framework that combines spatial filtering-based CSP with the modified Stockwell transform (MST). The best classifier that can differentiate the MI classes, with the maximum accuracy is chosen from five frequently classifiers. Then, the variance in tasks was statistically tested using the analysis of variances (ANOVA) test. In the convolutional neural networks (CNN), Dai et al. [25] use a single convolution scalp, whereas the optimum convolution scale varies with each subject, which reduces the classification accuracy. Another difficulty is when training data is inadequate. Therefore, the classification step suffers. To tackle these difficulties, they created a hybrid-scale CNN (HS-CNN) architecture for MI-EEG classification, that included a data augmentation approach. Sun et al. [26] proposed a novel normalization strategy based on one contralateral electrooculography (EOG) channel, that maintains MI-related neuronal potentials while eliminating the duplication effect among EOG channels. It uses Hjorth features to train the sub-optimal weights of their normalization algorithm for MI classification of the validation data. Also, Luo et al. [27] used ensemble support vector learning (ESVL) based on SVM classifier. Where the SVM is used to establish the decision boundary with the biggest interclass margin. The distances between samples and the decision boundary are turned into posterior probabilities. It merged probabilities derived from various SVM classifiers, to create the prediction model. As a result, ESVL takes advantage of the benefits of many trained SVM classifiers, to create a more accurate prediction model based on posterior probabilities. Malan and Sharma [28] use a special multi-view feature selection method based on regularized neighbourhood component analysis to simultaneously optimize time windows and frequency bandwidths. It used CSP to extract spatial features from MI-related EEG data at various time frames and frequency bandwidths. Then improved the data using the feature selection method. Finally, SVM is trained to classify ideal CSP features to identify MI tasks. Also, Saa et al. [29] employ a time–frequency representation of the brain signal acquired from different brain areas, to extract essential characteristics. The capacity of PCA and sequential forward selection is used to re-present feature sets compactly while deleting non-important features. Finally, the classification step by using two machine learning-based methods LDA and SVM.

3. Materials and methods

3.1. EEG device

The work of this paper uses EEG data acquired using our EEG device [4]. This device uses just three non-invasive activated channels, these channels are placed at the level of C5, C6, and Oz on the scalp following the international system 10–20 as illustrated in Fig. 2(c). Also, we added an electrode at the level of the left ear, to disable interference with the mass. Another reference electrode is added at Cz, to eliminate additional skin signals. The EEG signals acquired using all electrodes are very weak and they are at the level of μV. So, this system uses the amplifier (TDA2822) to increase each signal amplitude to the order of V. Also, this system contains some components such as capacitors, resistances, and diodes to decrease the parasite effect (Fig. 2(a)). Therefore, amplifiers output better signal quality. After the amplification of EEG signals, this system uses a simple microcontroller (ATMGA328PU). This microcontroller is programmed and booted using the Arduino board, to digitize analogue signals. Then, all acquired signals sent to the server were created based on Python language (Fig. 2(b)). This server allows receiving data in real-time, to make records and predictions with a sampling rate of between 90 and 120 predictions per second.

Algorithm 1: Features extraction algorithm

```
- For every channel, get the raw EEG (X) obtained in real-time.;
- f_s is the system sampling frequency in real-time;
for each instance t do
    for each channel j do
        for each feature Feat do
            - Calculate the feature from every channel using:
              Feat[j][t] = Feat(X[j][t − f_s : t]);
        end
    end
    - Replace each Delta channel with their features extracted;
end
```

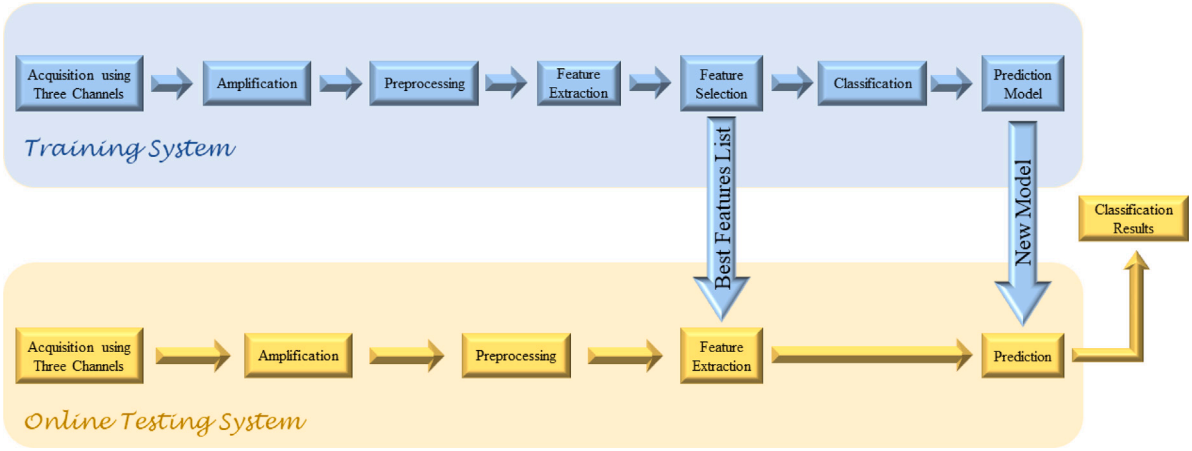
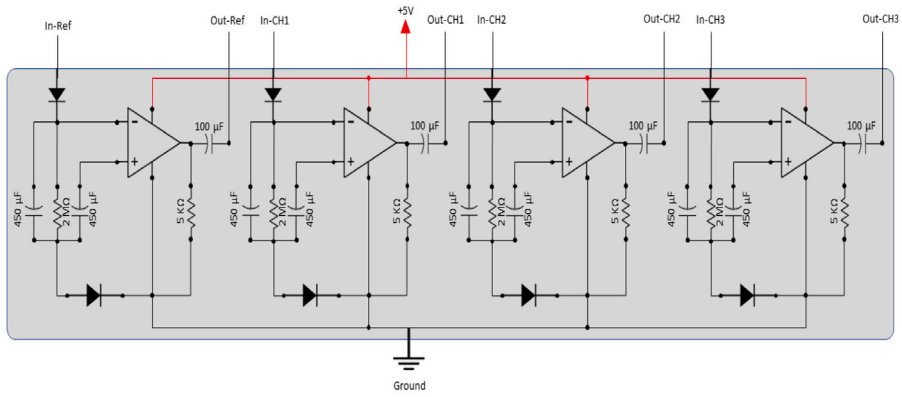
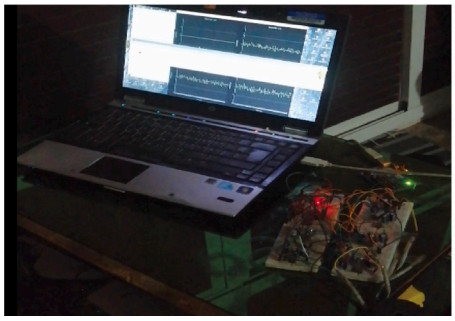


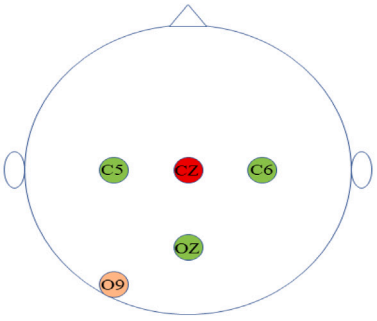
Fig. 1. Global design proposed for the MI prediction system in real-time based on delta wave.



(a) EEG signals amplifier



(b) EEG Device



(c) Electrodes locations

Fig. 2. A basic acquisition device for EEG signals was proposed for this work.

3.2. EEG data acquisition

The data used in this paper are acquired using our device presented in Section 3.1, and the signals are acquired by 3 channels and filtered by a bandpass filter to extract delta waves. These data are recorded from 7 subjects, such that each subject contains 100 trials of 100 samples, knowing that each subject is composed of 50% of the data for the imagination of the left-hand movement ('L'), and the remains 50% for the imagination of the right-hand movement ('R'). So, the sampling frequency during acquisition is in the order of 100 Hz. Also, the current frequency of the pay is 50 Hz.

3.3. Pre-processing

This step is very essential for this work because the preprocessing improves the quality of the acquired signals to have great results in real-time. Knowing that we eliminate the reference signal from each active channel by using the mathematical relationship $V_{Channel} - V_{Ref}$. This relationship obtains a signal emptied by the biopotential signals. Then all the signals must be standardized at the origin or the zero level. That is in other words when each subject is in relaxation or at rest, all EEG signals must be stabilized at the origin. But when each subject applies a task, all signals are influenced by these task activities [4]. For this work, we use the following relation at each moment t:

$$V'[t] = V[t] - \text{mean}(V[t - n \times f_s : t]) \quad (1)$$

Algorithm 2: LGBM algorithm

1- Classification;
- Initialize the training data $Train = \{(x_0, y_0), \dots (x_n, y_n)\}$;
- Initialize the great gradient data sample is α, while the minor gradient data sampling is β;
- Set the first node $\nu_0 = \underset{l}{\operatorname{argmin}}\,(L\,(y, l))$;

for *m=1* **to** *n-estimator* **do**
- Calculate the gradient absolute value: $y'_i = \left|\frac{\partial L(y_i, \nu(x_i))}{\partial \nu(x_i)}\right|_{\nu(x)=\nu_{m-1}(x)}$, $i = \{1, \dots, n\}$;
- $\mathrm{T} = \alpha \times \mathrm{len}(Train)$; $\mathrm{R} = \beta \times \mathrm{len}(Train)$;
- s=GetSortedIndeces(abs(y'));
- $\gamma = s[1 : \mathrm{T}]$, $\omega = \mathrm{RandomPick}(s[\mathrm{T} : \mathrm{len}(Train)], \mathrm{R})$;
- $\Psi = \gamma + \omega$;
- Calculate the gain: $G_j(d) = \frac{1}{n}\left(\frac{\left(\sum_{x_i \in \gamma_l} y'_i + \frac{1-\alpha}{\beta}\sum_{x_i \in \omega_l} y'_i\right)^2}{n'_l(d)} + \frac{\left(\sum_{x_i \in \gamma_r} y'_i + \frac{1-\alpha}{\beta}\sum_{x_i \in \omega_r} y'_i\right)^2}{n'_r(d)}\right)$;
- Build a new decision tree $\nu'_m(x)$ on set Ψ;
- Update $\nu_m(x) = \nu_{m-1}(x) + \nu_m(x)$;

end
return $\widetilde{\nu}(x) = \nu_{n-est}(x)$;
2- Features selection;
- Calculate the degrees of importance for each features;

for *each feature i* **do**
- Calculate the number of samples that reach the feature's node (mj);
- Using $\mathrm{IP}_j = \sum_{k=1}^{\widetilde{\nu}} freq_k \times (1 - freq_k)$, calculate the impurity value for every node j. With, $freq_k$ is the frequency of label k at a node, $\widetilde{\nu}$ is the number of unique labels;
- Calculate the importance $Degree_i$ using: $Degree_i = \sum_{j \in \mathrm{feature}} m_j \times \mathrm{IP}_j$;

end

Algorithm 3: MVO algorithm

- Initialize the population size and number of iterations (Iter);
- Initialize the sorted universes (SU), normalized inflation rate (NI), BlackHoleIndex at i and $r_1, r_2, r_3, r_4 = \mathrm{rand}([0, 1])$;
- Initialize all random universes x_i with $i \in \mathbb{N}$, WEP, TDR, low value (lb) and upper value (ub), and best universe;

while *True* **do**
- Calculate the fitness of universes;

for *each universe i* **do**
- Update WEP and TDR;

for *each Object j* **do**

if $r_1 < \mathrm{NI}(U(i))$ **then**
- WhiteHoleIndex= RouletteWheelSelection(-NI);
- U(BlackHoleIndex,j)= SU(WhiteHoleIndex,j);

end
if $r_2 <$ WEP **then**
if $r_3 < 0.5$ **then**
- U(i,j)= BestUniverse(j)+ TDR× ((ub(j)-lb(j))×r_4+lb(j));

else
- U(i,j)= BestUniverse(j)- TDR× ((ub(j)-lb(j))× r_4+lb(j));

end
end
end
end
end
return Best universe;

With $V[t]$ is the signal acquired at each moment (t), f_s is the system sampling frequency, and n is a positive integer. Such that when n is larger when the output signal is more sensitive to the task activated with a great delay in renormalizing the signal. But when n is small when the effect of each task is instantaneous and lasts for a few moments. So, in this work, we set the n value at $n = 12$. Furthermore, this system uses a bandpass filter to remove all unnecessary frequencies based on Abenna et al. [4,22] study, which improves signal stability against noisy sources.

3.4. Features extraction

The feature extraction step is the most important in this work for properly separating the data of all clusters. Where, this phase aims to highlight the characteristics of EEG-task through feature changes, rather than using raw EEG data. These features enable data properties to be distributed separately between different classes. Therefore, the classifier can distinguish between tasks data with an oversight compared to raw EEG. There are frequently EEG signal studies that have used a low number of channels, like sleep-stages recognition using only two channels. Where, based on Algo. 1, each feature is extracted from each channel at each moment (t). Where each feature is calculated from a real-time vector (queues) with a size of f_s samples. This vector contains samples for the last second acquired at each task. This means that the feature is connected at each moment with features specific to the same task. Therefore, we extract eleven distinct features from each channel, including Maximum, Minimum, Median, Mean, Kurtosis, Skewness, IQR, STD, and PFD [30–32], HM [4,33,34], and Spec. This system then extracted 33 separate features.

$$\mathrm{median}(X) = \begin{cases} \mathrm{sort}(X)\left[\frac{n+1}{2}\right] & \text{, if 'n' is an odd.} \\ \frac{\mathrm{sort}(X)\left[\frac{n}{2}\right] + \mathrm{sort}(X)\left[\frac{n}{2}+1\right]}{2} & \text{, if 'n' is an even.} \end{cases} \tag{2}$$

For filtering tasks, std is an excellent choice. The task with high std values are the excellent ones and lower std values indicate that the element is discriminative across classes. An element with a low standard value has almost equal values in all classes. This indicates it cannot tell the variance between several classes.

$$\mathrm{std}(X) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X[i] - \mathrm{mean}\,(X))^2} \tag{3}$$

The skewness (skew) of the data composition is a description of its symmetry requirement (X), it is defined as the anticipated value of the following equation:

$$\mathrm{skew}(X) = \frac{1}{n}\frac{\sum_{i=1}^{n}(X[i] - \mathrm{mean}\,(X))^3}{(\mathrm{std}(X))^3} \tag{4}$$

The gap between the first and third quartiles is called the inter-quartile range (iqr).

$$\mathrm{iqr}(X) = \mathrm{median}\left(\mathrm{sort}\,(X)\left[\frac{n}{2} :\right]\right) - \mathrm{median}\left(\mathrm{sort}\,(X)\left[: \frac{n}{2}\right]\right) \tag{5}$$

Petrosian fractal dimension (pfd) is expressed as follows:

$$\mathrm{pfd}(X) = \frac{\log(N)}{\log(N) + \log\left(\frac{N}{d/L}\right)} \tag{6}$$

where N is the number of samples in a segment, d is the waveform's planar diameter, and L is its overall length. The square root of the first derivative's variance divided by the signal's variance is the Hjorth mobility (HM), which reflects the mean frequency or the power spectrum's

standard deviation component.

$$\mathrm{HM}(X) = \sqrt{\frac{\mathrm{var}\left(\frac{\mathrm{dX}}{\mathrm{dt}}\right)}{\mathrm{var}(X)}} \tag{7}$$

$$\mathrm{Spec}(X) = \mathrm{mean}\left(\mathrm{abs}\left(\mathrm{fft2}\left(X\right)\right)\right) \tag{8}$$

where fft2 is the 2-dimensional Fourier transform.

Kurtosis is a dimensionless metric that characterizes the spike of a waveform and is a numerical measure that depicts the vibration signals distribution [35]. It can be written as:

$$\mathrm{Kurtosis}(X) = \frac{1}{n}\frac{\sum_{i=1}^{n}\left(X[i] - \mathrm{mean}(X)\right)^4}{\left(\mathrm{std}(X)\right)^4} \tag{9}$$

3.5. Classification and feature selection

Friedman established the gradient boosting machine (GBM) in 2011 to extend boosting to relapse issues, where the GBM is recommended for making additional characteristics with the least amount of prediction errors [22,36], and the GBM demonstrated is set to esteem that minimizes the loss function. The remaining esteem of the present model is estimated as the negative part of the loss function in each iterative preparation step, and an unused regression tree is built to coordinate the current remaining. When a current regression tree is incorporated in the going before the show [36], the remainder is revamped; at that point, the GBM show has advanced characteristics that had before been performed ineffectively. Because it allows choosing the significance of each include for the predicting framework for arrangement [4,22]. As well as to reduce the computation presumptions that the processors are subjected to, the include determination step is critical for reducing training thickness in any classification stage in real-time. The classification of EEG data is the most important aspect of this consideration since it allows for the building of real-time expectation models for activities engaged by each subject. To classify data, a variety of machines and deep learning algorithms can be used. But, we chose the LGBM classifier because it allows for building exceptional models with exceptionally proficient and a high classification speed during training. As well as simple control of key parameters using optimization algorithms, and it is based on (GBM) Algo. 2 [4,22]. Where for the LGBM algorithm, the loss changes of each class (y_i') are calculated during each model (v_m) created at each iteration, to reduce the error level to the lowest possible value. This means increasing the prediction accuracy to the maximum machine value during each cycle (m). Start cycle the dominant classes (γ) predicted by high accuracy values are selected, and the rest of the classes (ω) are randomly selected after. Which the new tree (v_m) is formed using the new tree (ψ) built according to the gain value (G_j), and then we update the model by adding the new model (v_m) to the ordinary model (v_{m-1}). Therefore, GBM trains a few models by adjusting the test weights at each stage of the preparation process. Then straightly combines these models to improve classification execution. In Algo. 2, we note the bottom that represents the calculation of the importance ratio features. Feature selection seeks to choose only the most important features from all, depending on the degree of importance, by calculating the value of each feature's importance degree during the models' development (v_m).

3.6. Optimization using MVO

The usage of MVO in this work is based on Algo. 3 [37] has become more significant in recent years to enhance any system and converge to a stable state and extreme in the shortest feasible period. Mirjalili et al. [37] employ the concepts of a white hole and black hole to analyse search spaces using MVO. On the other hand, MVO exploits the search spaces through wormholes, where each variable is compared to one of the things in the universe, and each solution to one of the universes. Moreover, Mirjalili et al. [37] apply an inflation rate to each solution that is proportional to the fitness function value of the solution. Through white or black hole tunnels, items can move between universes. According to Mirjalili et al. [37], white holes are more prevalent in universes with high inflation rates than in universes with low inflation rates, increasing the overall inflation rate. As a result, items can always be moved from a universe with strong inflation to one with low inflation. This can guarantee that over the iterations, the average inflation rates of all universes increase. As a result, MVO grants access to the LGBM classification's internal parameters, allowing the creation of high-expectancy models with high exactness values during testing. MVO algorithm is capable of adjusting inquiry and abuse more effectively and successfully. This avoids a large number of neighbouring configurations while allowing [38] to compare precision and convergence speed.

4. Results and discussion

4.1. Model evaluation

Terms of TP, FP, TN, and FN are the True/false prediction for positive/negative labels, that are used to describe various outcomes of the classification technique [4,22]. Accuracy (AC) is a commonly used metric for determining the classification capacity to distinguish between classes, and it is calculated using Eq. (10), where accuracy is one of the most important metric for assessing classifier execution, and it is defined as the number of test tests correctly classified by the algorithm.

$$\mathrm{AC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FN} + \mathrm{TN} + \mathrm{FP}} \tag{10}$$

To aid classifier performance comparisons, Cohen's Kappa (κ) measurement is employed. The measure κ can be useful for examining classifications that could be inferable by chance, and it is derived using Eq. (11):

$$\begin{cases} P_e = \frac{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})+(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}{(\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN})^2} \\ \kappa = \frac{\mathrm{AC}-P_e}{1-P_e} \end{cases} \tag{11}$$

Ordinary data recovery metrics such as precision (PR) (Eq. (12)), affectability (SE) (Eq. (13)), and the F1 degree are used to evaluate the test classification's execution [22,39].

$$\mathrm{PR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{12}$$

$$\mathrm{SE} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{13}$$

$$\mathrm{F1} = \frac{2 \times \mathrm{PR} \times \mathrm{SE}}{\mathrm{PR} + \mathrm{SE}} \tag{14}$$

One of the foremost exacting execution measurements for any prediction procedure is Mathew's correlation coefficient (MC) characterized by Eq. (15) [22].

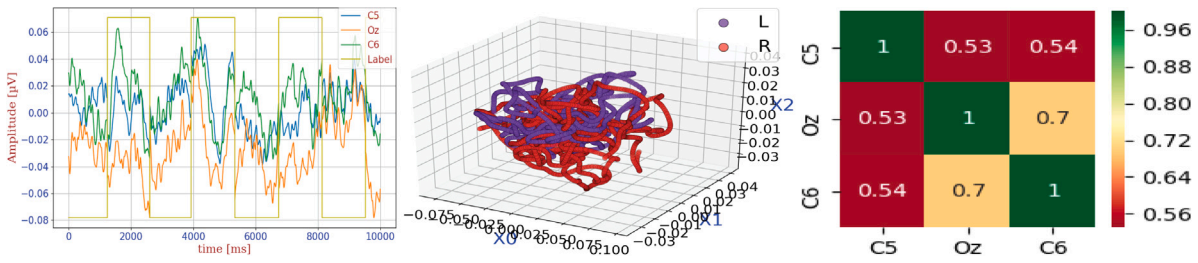
$$\mathrm{MC} = \frac{\mathrm{TP} \times \mathrm{TN} - \mathrm{FP} \times \mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}} \tag{15}$$

The following is the definition of the Jaccard-score (JS) between confusion matrix coefficients as illustrated Eq. (16):

$$\mathrm{JS} = \frac{\mathrm{FP}}{\mathrm{TN} + \mathrm{FN} + \mathrm{FP}} \tag{16}$$

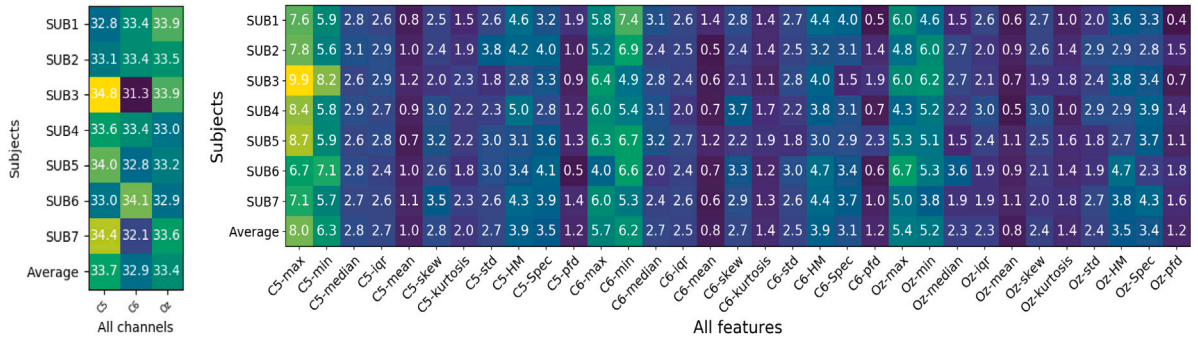
4.2. Delta waves acquisition

In this work, we extracted delta waves instead of using EEG data to predict EEG tasks, showing Fig. 3 part of the captured delta waves properties. Fig. 3(a) shows delta wave variations during the time, observing a near convergence of signal changes in each channel, especially with upper and lower ratios. Blue, orange, and green colour represent both C5, OZ, and C6. We observe MI labels change in yellow.



(a) EEG signals **(b)** 3D representation **(c)** Correlation matrix

Fig. 3. Example of Delta rhythm characteristics acquired in real-time for Hand MI tasks (Subject 10).



(a) Three channels **(b)** The importance degree for each extracted feature

Fig. 4. Features selection results for all channels/features for all subjects.

According to this Figure, we observe stability in signals compared to unstable raw EEG, which is characterized by a non-stationary nature. Fig. 3(b) also shows spatially data form, which shows the existence of a punctual continuity of the data for each task. Thus, it is easy to separate the data in real-time compared to the direct use of raw EEG data, they are generally non-stationary. Where samples are distributed in the form of interconnected lines, making predicting samples of each task possible compared to the natural random distribution of raw EEG. We also note a clear compression of data size in space at the scale of 0.01. This compression is due to the smallest delta waves amplitude as illustrated in Fig. 3(a). Also, in Fig. 3(b), we note a high concentration of data at the graph centre, due to the observed correlation levels between channels. So, Fig. 3(c) shows the correlation levels between all three channels with values of more than 0.5, these scores show the existence of a dependency between different channels' signals. Therefore, all channels' characteristic is linked by over 50%. Thus, less take advantage of EEG characteristics acquired from different scalp regions.

4.3. Features selection results

After extracting all features from delta waves from each channel, we select only the best features. Where Fig. 4 illustrates the feature selection results using our acquisition system based on delta rhythm for each subject. So, each value indicates the contribution ratio (in%) of each feature to recognize baseline MI tasks in each subject. Such as, Fig. 4(a) shows the importance of channels during direct classification using delta waves and Fig. 4(b) illustrates all importance degrees after the feature extraction step. Therefore, in Fig. 4(a), we notice that all importance degrees are at the same level between all subjects in each feature. This subfigure shows the convergence for all values to the range values between 31% and 35%, where the average result is represented at the bottom, while all values are almost equal between channels for the same feature. This equality is due to the prevailing similarity of delta rhythms signals variations. On the other hand, after extracting all features, we notice a clear variance between the importance degrees of features extracted from the same subject, but their importance remains semi-stable when changing the subject. Which indicates the stability of task characteristics over time. Therefore, in this figure, we notice that the best 10-features are found using min, max, HM, Spec, and skew measures from all three channels. Knowing that, when using delta waves, the difference is small between values of max and min, while signals are more synchronized with the activated task. This implies the possibility to separate between clusters linearly if the acquisition system quality has been increased.

4.4. Auto-reduce the number of extracted features

Fig. 5 aims to study the evolution of system accuracy by increasing the number of features extracted without employing of MVO optimizer, where features are selected according to their contribution degree to the MI tasks as illustrated in Fig. 4. So, features that impair performance are highlighted and removed. Such that, Fig. 5 shows that the accuracy value increases when the NFET remains between 1 to 5. Thus, we note that accuracy values exceed 80% after extracting the most important 3-features. Once the maximum values range from 85% to 86% when NFET values are between 6 and 18. Then, accuracy values begin to decrease when using more than 19-features to the order of 83%. Therefore, we conclude that some features weaken the system. So, since the maximum value of NFET is 15, the most important 15-features extracted, or even 6-features extraction can be satisfied because there is only a slight decrease in accuracy compared to the use of 15-features. But the lowest number of features is considered a less predictable time, for classification and therefore improvement. Therefore, using the results illustrated in Fig. 4(b) for each subject, the dominant features are max, min, HM, skew, median, std, iqr from each channel.

After extracting the best features for the system, Fig. 6 illustrates an example of temporal and spatial characteristics of signals, such as Fig. 6(a) illustrates an example of the variation of the features in real-time in synchronization with each task that has been activated (yellow). This figure shows that during task 'R', all features condense

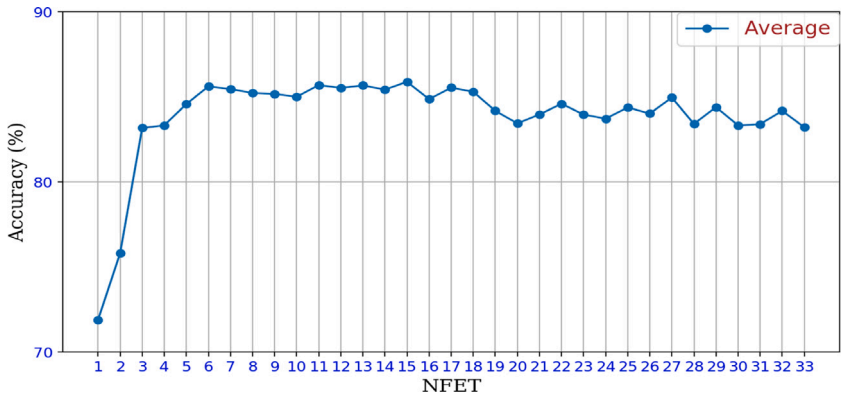
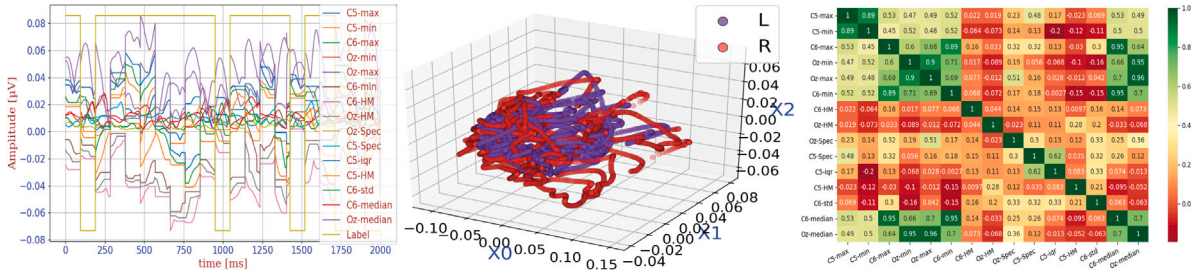


Fig. 5. Evaluation of the number of features needed for the prediction stage in real-time.



(a) Variation of the Features **(b)** 3D representation **(c)** Correlation matrix

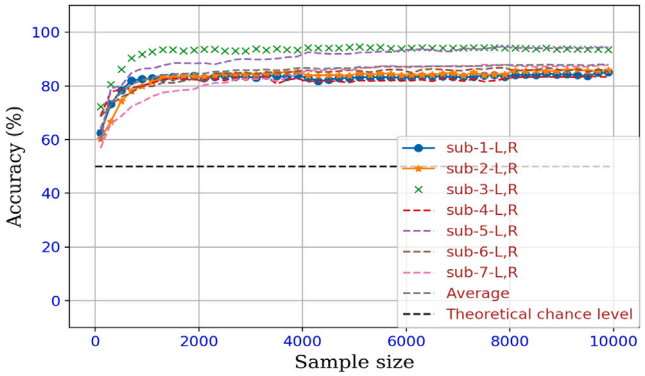
Fig. 6. Example of the best 15 features characteristics acquired in real-time for Hand MI tasks (Subject 1).

and approach the x-axis. On the other hand, Fig. 6(b) illustrates a spatial representation of EEG data after the application of the feature extraction step. Where the data expand spatially compared to the representation of Fig. 3(b), which increases the degree of non-linear separation between the two classes. In addition, Fig. 3(c) shows good independence between the features that were extracted except between the real-time measures of min, max and median of EEG signals, because of delta wave frequency characteristics. Also, links between the channels were illustrated in the correlation matrix of Fig. 3(c). But generally, the matrix in Fig. 6(c) shows great independence between the totality of features, which indicates the best quality of the system developed.

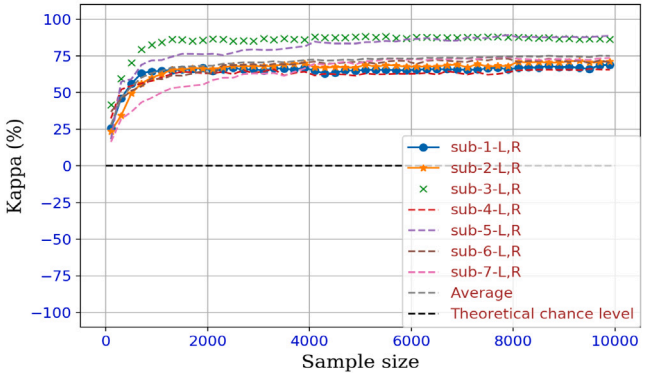
4.5. Classification results

4.5.1. Study of the system stability using experimental chance levels

Fig. 7 aims to demonstrate the stability of system accuracy levels when the volume of data increases over time, where Fig. 7(a) illustrates accuracy variations when extracting only the best 15-features per subject. To get these results, we take clusters data with ratios of 50%. At each moment, we classify the data 20 different times, then we calculate the average to find each accuracy value. So, the number of classifications per subject is 2000. At each time, we add 100 samples to the data on the x-axis. The same way is applied to get Kappa-score results in Fig. 7(b). As we note in the graph of Fig. 7(a), all results remain stable after using only 1000 samples, and accuracy values exceed 80%, which is fairly good. Such that, these results can be further increased when using more than 3 channels and more expensive devices. As is shown in the results of Abenna et al. [22] for an offline study. Moreover, this figure shows a convergence between the system performance for various subjects, which shows great stability for the extracted data during acquisition, except for two subjects were found with an accuracy value of over 95%.



(a) Accuracy



(b) Kappa

Fig. 7. Experimental chance levels for the resultants system.

Table 1
Classification results for baseline MI-EEG-based three-channel EEG using LGBM-MVO directly.

Subject	L_{rate}	Classification performances									
		AC (%)	*SE* (%)	*PR* (%)	*F1* (%)	κ (%)	*MC* (%)	*JS* (%)	*FM* (%)	*MI*	*ZOL*
SUB1	0.0102	57.65	59.31	58.55	57.08	16.75	17.85	40.10	53.23	0.02	847
SUB2	0.0364	68.20	67.08	67.81	67.21	34.64	34.89	50.89	58.01	0.06	636
SUB3	0.0624	64.70	66.46	67.31	64.55	31.48	33.76	47.70	56.00	0.06	706
SUB4	0.0926	59.90	60.94	60.81	59.88	21.05	21.75	42.74	52.71	0.02	802
SUB5	0.0003	58.80	58.86	58.66	58.51	17.37	17.52	41.44	52.04	0.02	824
SUB6	0.0402	76.80	75.74	76.62	76.01	52.13	52.35	61.51	65.54	0.14	464
SUB7	0.0612	58.25	58.31	58.17	58.04	16.37	16.49	40.94	51.73	0.01	835
Average	0.04	64.34	64.73	64.96	63.87	28.90	29.69	47.40	56.26	0.05	713.17

Table 2
Classification results for the best fifteen features extracted data using LGBM-MVO.

Subject	L_{rate}	Classification performances									
		AC (%)	*SE* (%)	*PR* (%)	*F1* (%)	κ (%)	*MC* (%)	*JS* (%)	*FM* (%)	*MI*	*ZOL*
SUB1	0.924	86.35	86.46	85.30	85.74	71.53	71.75	75.16	77.41	0.28	273
SUB2	0.919	88.05	89.21	87.28	87.72	75.57	76.47	78.18	79.62	0.33	239
SUB3	0.994	95.70	95.42	95.50	95.46	90.92	90.92	91.33	92.18	0.49	86
SUB4	0.961	86.25	86.06	85.25	85.59	71.21	71.31	74.94	77.31	0.28	275
SUB5	0.978	95.35	95.61	95.04	95.27	90.54	90.65	90.97	91.28	0.50	93
SUB6	0.990	90.50	91.11	90.21	90.40	80.85	81.32	82.49	83.02	0.39	190
SUB7	0.910	88.05	86.98	87.73	87.32	74.64	74.71	77.62	80.10	0.30	239
Average	0.96	90.37	90.64	89.76	90.03	80.10	80.40	82.18	83.47	0.38	192.67

4.5.2. Classification results using the global system

In the first phase of reviewing the results of this research, we calculated the accuracy of the classification using delta waves data extracted from each channel. Table 1 shows the results obtained when classifying the data of each subject with the average accuracy at the bottom. Knowing that the signals are acquired using three active channels. In this table, we note that the accuracy values are in the order of 60% because of the lower acquisition system quality, which uses basic and simple electronic components, that they are generally low cost compared to more expensive devices with large sizes like NeuroFax, BCI2000, NeuroHeadSet devices. These non-mobile devices restrict users' freedom. Through this table, we observe medium resolution values limited between 57% and 77%, and L_{rate} values are less than 0.1, which shows that the data are more spatially condensed between them. After extracting the best 15-features, Table 2 illustrates the classification results using the best 15 features. Where the data is divided into 80% as training data and the remaining 20% as testing data. Through the results of this Table, we note a significant rise in prediction accuracy to more than 86%, which allows for an increase in the system accuracy values between 85% and 95%, and kappa-score values between 71% and 91%, which shows an increase of 25% compared to the prediction of delta rhythm without feature extraction step. Moreover, after the use of the extracted features, the MVO optimizer found stable LGBM L_{rate} parameter values of more than 0.9, which implies that the data are well spatially expanded to the aim of being well-separated between clusters. Therefore, these results indicate the data classification stability reflected in the system's accuracy. On the other hand, Table 3 demonstrates the EEG signal classification results using other algorithms. Where training data for each subject is classified using each algorithm. Then, the prediction performances for classification using each algorithm are assessed, prompting us to observe the difference between classification algorithms' quality, where RF is the highest average accuracy of 86.56%. Due to Table 2 results, we note LGBM's superiority over the rest algorithms in Table 3, which confirms our choice validity of the best classifier.

4.5.3. Comparison with literature review

Table 4 illustrates a comparative study of our results with related work, that they classified the signals acquired for tasks of hand's MI ('L' or 'R'), using just 3 channels. As in this table, we note that the best accuracy value was found in the work of Kim et al. [23], which found the highest accuracy value of 89.36% in the literature using WDPSD for the feature extraction step and the SVM algorithm for classification. Such that, this work allows to increase the separation and reduce the non-stationary between MI tasks based on existing frequency characteristics for each channel. This methodology is very advanced to study the energy levels in each channel for each task. This feature was also used in our work under the name of 'Spec', with an average value of importance degrees for the prediction step at 13% as illustrated in Fig. 4(b). This shows, that this study can be unenforceable using our real-time acquisition system. Also, the second important accuracy value result is 88.85% was found by Bose et al. [24] using the PPV-LSSVM method, such this work mixes the CSP filter and the MST transformation for well separated between clusters' data, and LSSVM algorithm for classification. This method remains classic like most methods that use the CSP filter as a kernel to eliminate the correlation between all channels. The third good work was done by Dai et al. [25], to find an accuracy value of 87.6% using HS-CNN architecture. This architecture generates prediction models based on different features in all scales mixed compared to the classical state. These generated prediction models based on each scale features separately from others, which allows to miss a large set of characteristics in the data. This recent approach is better but is not sufficient to obtain high levels of accuracy. We can also notice in this table the method of Luo et al. [27], which allows classifying the data of MI with a kappa value of 72% using ESVL classification. This method allows generating a large set of interclass margins based on SVM, which increases the separation level between clusters as the classification approach is based on SVM with Kernel-RBF. Thus, the data is filtered using CSP previously, which improves the independence between features. Also, the CNN algorithm is used by Dose et al. [40], to classify the EEG signals acquired using the BCI2000 device. This generated dataset is named as PhysioNet MI-EEG. Such that, this work found an accuracy value of 79.2%, which is an average result and realist, that this method is capable of real-time portable applications.

5. Conclusion

This work shows the efficiency and importance of delta rhythm for predicting real-time MI tasks. Where the use of a basic signal acquisition system, shows that our preprocessing and prediction system can be found high recognition results by using advanced acquisition system

Table 3
Overall classification results for optimal fifteen features using other classifiers.

Classifier	Classification performances									
	AC (%)	*SE* (%)	*PR* (%)	*F1* (%)	κ (%)	*MC* (%)	*JS* (%)	*FM* (%)	*MI*	*ZOL*
RF	86.56	86.98	85.95	86.17	72.44	72.92	76.28	78.08	0.31	268.83
XGB	84.80	85.32	84.21	84.40	68.95	69.52	73.35	75.31	0.27	304.00
DT	80.77	81.37	80.42	80.33	61.00	61.78	67.75	70.70	0.22	384.50
KNN	79.33	79.35	78.97	78.62	57.53	58.31	65.41	69.18	0.20	413.50
MLP	72.78	74.32	73.41	72.56	46.26	47.72	57.62	62.77	0.14	544.33

Table 4
Comparative results with related work.

Work	*AC* (%)	κ (%)	Method
Dose et al. [40]	79.20	nan	CNN
Monesi and Sardouie [41]	nan	62.7	ECSP
Saa et al. [29]	82.5	nan	SFS-SVM
	83.92	nan	SFS-LDA
	82.58	nan	PCA-SVM
	81.23	nan	PCA-LDA
Malan and Sharma [28]	84.5	nan	RNCA-CSP-SVM
Kim et al. [23]	89.36	nan	WDPSD-SVM
Bose et al. [24]	88.85	nan	PPV-LSSVM
Dai et al. [25]	87.6	nan	HS-CNN
Sun et al. [26]	nan	72.0	CCNM-SVM
Luo et al. [27]	nan	71.0	ESVL
Luo et al. [42]	85.0	nan	STDF-SVM
This work	90.37	80.1	Delta-LGBM-MVO

architectures, as delta rhythms are more stable and less noisy allowed to quantify signal states easily. Also, both steps of feature extraction and feature selection show great importance to improve the separation between classes and increase the accuracy level between 60% and 85%. Moreover, the LGBM algorithm shows its reliability to generate great prediction models, especially after the learning-rate parameter optimization using the MVO algorithm. Finally, the resulting system is more stable with accuracy values in order of 90% as the use of a simple acquisition system. Therefore, this system is characterized by simple and effective methods to predict complex MI tasks with accuracy ratios of more than 90%. That allows for improving the system performance in future versions. Also, this work is characterized by simulating all stages' work for the use of the system in real-time applications, making it easy for all steps to integrate with other acquisition devices with the same results. Moreover, the use of the lowest number of channels is necessary to study the system efficiency as a wireless device. That allows reducing signal processing requirements. Then reducing the device's energy consumption to the lowest level. On the other hand, it is clear that the acquisition device is still simple that resulting the loss of many MI characteristics. This means the prediction system's ability to exceed its current accuracy in future works. Therefore, in the future, we hope to develop and manufacture highly acquisition systems to predict different MI tasks. Also, we will look for new and effective approaches to get advanced features that can identify particular MI activities with high quality in real-time acquisition.

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.

Data availability

I have shared the link to the data used in the manuscript.

References

[1] 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.

[2] A. Asif, M. Majid, S.M. Anwar, Human stress classification using EEG signals in response to music tracks, Comput. Biol. Med. (ISSN: 1879-0534) 107 (August) (2019) 182–196, http://dx.doi.org/10.1016/j.compbiomed.2019.02.015.

[3] K. Venkatachalam, A. Devipriya, J. Maniraj, M. Sivaram, A. Ambikapathy, S.A. Iraj, A novel method of motor imagery classification using eeg signal, Artif. Intell. Med. (ISSN: 18732860) 103 (June 2019) (2020) http://dx.doi.org/10.1016/j.artmed.2019.101787.

[4] S. Abenna, M. Nahid, H. Bouyghf, B. Ouacha, EEG-based BCI: a novel improvement for EEG signals classification based on real-time preprocessing, Comput. Biol. Med. (2022) http://dx.doi.org/10.1016/j.compbiomed.2022.105931.

[5] S.B. Lee, H.J. Kim, H. Kim, J.H. Jeong, S.W. Lee, D.J. Kim, Comparative analysis of features extracted from EEG spatial, spectral and temporal domains for binary and multiclass motor imagery classification, Inform. Sci. (ISSN: 0020-0255) 502 (2019) 190–200, http://dx.doi.org/10.1016/j.ins.2019.06.008.

[6] E.G.M. Pels, E.J. Aarnoutse, S. Leinders, Z.V. Freudenburg, M.P. Branco, B.H.V.D. Vijgh, T.J. Snijders, T. Denison, M.J. Vansteensel, N.F. Ramsey, Stability of a chronic implanted brain-computer interface in late-stage amyotrophic lateral sclerosis, Clin. Neurophysiol. (ISSN: 1388-2457) 130 (10) (2019) 1798–1803, http://dx.doi.org/10.1016/j.clinph.2019.07.020.

[7] X. Tang, W. Li, X. Li, W. Ma, X. Dang, Motor imagery EEG recognition based on conditional optimization empirical mode decomposition and multi-scale convolutional neural network, Expert Syst. Appl. (ISSN: 0957-4174) 149 (2020) http://dx.doi.org/10.1016/j.eswa.2020.113285.

[8] K.A. Khan, P.P.M. Shanir, Y.U. Khan, O. Farooq, A hybrid local binary pattern and wavelets based approach for eeg classification for diagnosing epilepsy, Expert Syst. Appl. (ISSN: 0957-4174) 140 (2020) http://dx.doi.org/10.1016/j.eswa.2019.112895.

[9] 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.

[10] B. Ouacha, H. Bouyghf, M. Nahid, S. Abenna, DEA-based on optimization of inductive coupling for powering implantable biomedical devices, Int. J. Power Electron. Drive Syst. (ISSN: 2088-8694) 13 (3) (2022) 1558–1567, http://dx.doi.org/10.11591/ijpeds.v13.i3.pp1558-1567.

[11] S.R. Varas, R.A. Vazquez, Evaluating the effect of the cutoff frequencies during the pre-processing stage of motor imagery EEG signals classification, Biomed. Signal Process. Control (ISSN: 17468108) 54 (2019) 101592, http://dx.doi.org/10.1016/j.bspc.2019.101592.

[12] D. Yang, L. Wang, W. Hu, C. Ding, W. Gan, F. Liu, Trajectory optimization by using EMD- and ICA-based processing method, Meas.: J. Int. Meas. Confed. (ISSN: 0263-2241) 140 (2019) 334–341, http://dx.doi.org/10.1016/j.measurement.2019.03.063.

[13] 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, 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.

[14] C. Wei, L.-l. Chen, Z.-z. Song, X.-g. Lou, D.-d. Li, EEG-based emotion recognition using simple recurrent units network and ensemble learning, Biomed. Signal Process. Control (ISSN: 1746-8094) 58 (2020) 101756, http://dx.doi.org/10.1016/j.bspc.2019.101756.

[15] E. Dagdevir, M. Tokmakci, E. Dagdevir, 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.

[16] V. Noreika, S. Georgieva, S. Wass, V. Leong, 14 Challenges and their solutions for conducting social neuroscience and longitudinal EEG research with infants, Infant Behav. Dev. (ISSN: 0163-6383) (June) (2020) 101393, http://dx.doi.org/10.1016/j.infbeh.2019.101393.

[17] 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: 15732517) 251 (2019) 156–161, http://dx.doi.org/10.1016/j.jad.2019.03.058.

[18] S. Abenna, M. Nahid, H. Bouyghf, An enhanced EEG prediction system for motor cortex-imagery tasks using SVM, in: 10th International Conference on Innovation, Modern Applied Science & Environmental Studies, EDP Sciences, 2022, http://dx.doi.org/10.1051/e3sconf/202235101026.

[19] M. Hearst, S. Dumais, E. Osuna, J. Platt, B. Scholkopf, Support vector machines, IEEE Intell. Syst. Appl. 13 (1998) 18–28, http://dx.doi.org/10.1109/5254.708428.

[20] R. Fan, K. Chang, C. Hsieh, X. Wang, C. Lin, LIBLINEAR: a library for large linear classification, J. Mach. Learn. Res. 9 (2008) 1871–1874, http://dx.doi.org/10.5555/1390681.1442794, https://dl.acm.org/doi/10.5555/1390681.1442794.

[21] A. Liaw, M. Wiener, Classification and regression by RandomForest, Forest 23 (2001).

[22] 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 (103102) (2022) http://dx.doi.org/10.1016/j.bspc.2021.103102.

[23] C. Kim, J. Sun, D. Liu, Q. Wang, S. Paek, An effective feature extraction method by power spectral density of EEG signal for 2-class motor imagery-based BCI, Med. Biol. Eng. Comput. (2017) http://dx.doi.org/10.1007/s11517-017-1761-4.

[24] R. Bose, K. Samanta, S. Chatterjee, S. Bhattacharyya, A. Khasnobish, Motor imagery classification enhancement with concurrent implementation of spatial filtration and modified stockwell transform, 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.

[25] G. Dai, J. Zhou, J. Huang, N. Wang, HS-CNN: A CNN with hybrid convolution scale for EEG motor imagery classification, J. Neural Eng. (2019) http://dx.doi.org/10.1088/1741-2552/ab405f.

[26] L. Sun, Z. Feng, B. Chen, N. Lu, A contralateral channel guided model for EEG based motor imagery classification, Biomed. Signal Process. Control 41 (2018) 1–9, http://dx.doi.org/10.1016/j.bspc.2017.10.012.

[27] J. Luo, X. Gao, X. Zhu, B. Wang, N. Lu, J. Wang, Motor imagery EEG classification based on ensemble support vector learning, Comput. Methods Programs Biomed. 193 (2020) http://dx.doi.org/10.1016/j.cmpb.2020.105464.

[28] 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.

[29] J.F.D. Saa, H.J.V. Lora, D.J.M. Vásquez, Classification of imaginary motor task from electroencephalographic signals: A comparison of feature selection methods and classification algorithms, Rev. Mex. Ingen. Biomed. (2018) http://dx.doi.org/10.17488/RMIB.39.1.8.

[30] E.A.F. Ihlen, Introduction to multifractal detrended fluctuation analysis in matlab, Front. Physiol. 3 (June) (2012) 1–19, http://dx.doi.org/10.3389/fphys.2012.00141.

[31] F.S. Bao, X. Liu, C. Zhang, PyEEG : An open source python module for EEG / MEG feature extraction, Comput. Intell. Neurosci. 2011 (2011) http://dx.doi.org/10.1155/2011/406391.

[32] A. Petrosian, Kolmogorov complexity of finite sequences and recognition of different preictal EEG patterns, in: Proceedings Eighth IEEE Symposium on Computer-Based Medical Systems, 1995.

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

[34] 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.

[35] 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.

[36] M. Gong, Y. Bai, J. Qin, J. Wang, P. Yang, S. Wang, Gradient boosting machine for predicting return temperature of district heating system : A case study for residential buildings in tianjin, J. Build. Eng. (ISSN: 2352-7102) 27 (2020) http://dx.doi.org/10.1016/j.jobe.2019.100950.

[37] S. Mirjalili, S.M. Mirjalili, A. Hatamlou, Multi-verse optimizer: a nature-inspired algorithm for global optimization, Nat. Comput. Appl. (30) (2015) http://dx.doi.org/10.1007/s00521-015-1870-7.

[38] Y. Tikhamarine, D. Souag-gamane, A.N. Ahmed, O. Kisi, A. El-shafie, Improving artificial intelligence models accuracy for monthly streamflow forecasting using grey wolf optimization (GWO) algorithm, J. Hydrol. (ISSN: 0022-1694) (2019) 124–435, http://dx.doi.org/10.1016/j.jhydrol.2019.124435.

[39] Z. Kastrati, A.S. Imran, Performance analysis of machine learning classifiers on improved concept vector space models, Future Gener. Comput. Syst. (ISSN: 0167739X) 96 (2019) 552–562, http://dx.doi.org/10.1016/j.future.2019.02.006.

[40] 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. (2018) http://dx.doi.org/10.1016/j.eswa.2018.08.031.

[41] M.J. Monesi, S.H. Sardouie, Extended common spatial and temporal pattern (ECSTP): A semi-blind approach to extract features in ERP detection, Pattern Recogn. (2019) 18940–18950, http://dx.doi.org/10.1016/j.patcog.2019.05.039.

[42] J. Luo, Z. Feng, N. Lu, Spatio-temporal discrepancy feature for classification of motor imageries, Biomed. Signal Process. Control 47 (2019) 137–144, http://dx.doi.org/10.1016/j.bspc.2018.07.003.