

Classification of Four-Class Motor-Imagery Data for Brain-Computer Interfaces



Ngan Thanh Luu, Anh Hoang Lan Duong, Dung Xuan Nguyen, Khiet Thi Thu Dang, Chau Nu Ngoc Pham, Thuong Hoai Nguyen, and Ha Thi Thanh Huong

Abstract Brain-computer interface (BCI) provides opportunities for patients with severe motor impairment to regain their motor controls by decoding and converting Electroencephalogram (EEG) signals into the motor control signal. Existing processing schemes, including artifact removal and band-pass filtering, usually cause leakage of EOG artifacts. Hence, the preprocessing order in this paper was optimized for artifact removal in event-related desynchronization. To address this challenge, we implemented an algorithm to classify four different motor imagery tasks, including the left hand, right hand, both feet, and tongue movements. Here, we use Graz dataset A published in the BCI competition IV in 2008. Band-pass filter (7–30 Hz) with 5th order Butterworth IIR combined with linear regression was applied to remove the artifact of the signal. Mutual Information Best Individual Features (MIBIF) and Linear Discriminant Analysis (LDA) algorithms were used for selecting features and classifying the classes. The mean training kappa of our classifier is 0.57, and the mean evaluated kappa is 0.41. Besides, the application of the Common Spatial Pattern enhanced the kappa score (from 0.33 to 0.41). Those results are comparable to those from the BCI competition (range from 0.29 to 0.57). Hence, our approach could be the foundation for further development of motor imagery based BCI research. For future works, we propose to create an algorithm for the automatic procedure by automatically choosing proper frequency bands for band power features.

Keywords EEG · BCI · Motor imagery · Common spatial pattern · Linear discriminant analysis · Mutual information

N. T. Luu · A. H. L. Duong · D. X. Nguyen · K. T. T. Dang · C. N. N. Pham · T. H. Nguyen · H. T. T. Huong (✉)
School of Biomedical Engineering, International University, Ho Chi Minh City, Vietnam
e-mail: htthuong@hcmiu.edu.vn

Vietnam National University-Ho Chi Minh City (VNU-HCMC), Ho Chi Minh City, Vietnam

1 Introduction

With the development of non-invasive techniques, electroencephalography (EEG) plays an important role in the research of neuroscience, cognitive science, and cognitive psychology, especially in the Brain-Computer Interface (BCI). EEG-based BCI system mainly relies on modulating the responses of large populations of neurons, either through a training period time or using external stimuli [1]. An important advantage of EEG is the extremely low risk to the subjects, compared to other invasive and semi-invasive techniques, such as electrocorticography (ECoG) [2]. Besides, EEG has an extremely high temporal resolution on milliseconds level, and its signals can be divided into several different frequency patterns (e.g. mu, beta) which can present different characteristics, patterns and access different functional brain states [3, 4]. Many factors are contaminating EEG signals, including both objective factors (e.g. environmental noises, faulty electrodes) and subjective factors [e.g. electrooculography (EOG)]. These factors result in various undesired artifacts, which reduce the classification accuracy in the BCI system, alter the characteristics of neurological phenomena, and make significant mistakes in controlling the BCI system [5, 6]. Therefore, it is important to apply the artifact removal and band-pass filtering to avoid causing leakage of EOG artifacts and gains meaningful EEG signals.

Through modulation of brain signals, such as motor imaginary (MI), BCI can be applied to both medical and non-medical fields. Applications in medicine, including cochlear implants for the deaf, deep brain stimulation for Parkinson's disease has become more and more popular in medical treatments [7, 8]. Besides, scientists and engineers have also investigated various non-medical BCI products such as lie detector, alertness monitor, games, e-learning system, etc. [9–12]. As an important role in BCI, EEG-based BCI systems have strong opportunities to help patients with severe motor impairment regain their motor controls such as cursor control, prosthetic control, and speller [13–15]. EEG in BCI also has applications in interacting with a Web browser, controlling robots as well as lie detector [16, 17]. Commonly, a BCI system works based on detecting changes in brain activity when the brain responds to voluntary or involuntary mental commands. Motor Imagery Brain-Computer Interface (MI-BCI) is the most popular used to detect EEG signals through imagined movements. When subjects imagine moving a particular part in the body, the neural activity in the sensorimotor cortex will be spatiotemporally similar to the activity during performing the real movement. Therefore, this type of neural response allows us to discriminate and match the EEG signals with each imagined body movement [18]. The biggest challenge in MI-BCI system is the difficulty in feature extraction, which needs making the selective features robust, informative, and discriminative. The motor imagery frequencies have both inter- and intra-subject variabilities, leading to low accuracy in this system [19, 20].

Related to EEG recording and motor imaginary system, Hans Berger has already described and discovered the oscillations in alpha waves during closed and open eyes in the 1930s [21]. This is a common phenomenon observed in the EEG experiments. A phenomenon like event-related desynchronization (ERD) and synchronization (ERS)

was introduced to investigate the dynamics of EEG oscillations [22]. By imagining the movement with EEG and EMG recording, these phenomena can be able to change the power of the mu (8–13 Hz) and beta (13–30 Hz) bands [23, 24].

Event-related desynchronization (ERD) is a neurological phenomenon that was introduced by Gert Pfurtscheller and colleagues in the 1970s [25]. This can be used in BCI to classify different motor imagery classes. In the case of a subject imagine about a limb movement, event-related desynchronization will happen. It leads to a drop in the mu and beta waves' power. Otherwise, the event-related synchronization (ERS) increases the power of mu and beta waves when the imaginary movement stops.

The band power is the most common methodology used to determine event-related desynchronization and synchronization. First, the signal is filtered in distinct bands such as 8–14 Hz, 24–30 Hz..., then the influence level of the ERD/ERS effect was represented with different frequency components (e.g. in 8–14 Hz, the ERD effect occurs the most clearly). In the same frequency bands, the effect of ERD/ERS does not perform similarly for each subject, so it is recommended to select the frequency bands for individual participants. Then, the signal will be converted from frequency bands into power values. It means that the ERD/ERS alters when the power changes [26]. In an experiment, we perform various trials to get the final results, so the resulting signal is the average value over trials. This above information describes how to detect and quantify the effect of ERD/ERS by the band power method.

In this paper, we used the BCI Competition IV 2008 dataset IV 2a that contains the EEG signal recording of 9 subjects. This dataset includes 18 files of data recordings (nine files for the training set and nine files for the evaluation set) and 18 files of true labels for the evaluation step. The dataset was divided into four classes that are four different motor imaginary tasks (left hand, right hand, both feet, and tongue) [27]. This work is aimed to build an algorithm that is utilized to distinguish four above different motor imagery tasks. Normally, multichannel electroencephalogram (EEG) signals give fuzzy images of brain activity due to its low signal-to-noise ratio (SNR) [28]. This is the reason why it is required to filter signals before usable in BCI applications. In our proposed algorithm, we applied numerous methods to detect and remove all unwanted factors to get clean EEG signals. First, EEG signals were filtered by bandpass filter in the frequency band between 7 and 30 Hz with 5th order Butterworth IIR. Besides, we also applied linear regression and down-sampling 62.5 Hz in preprocessing.

In comparison with other techniques such as wavelet-based filter, linear regression is a simpler method that allows us to easily predict the effect of the filtering [29]. Due to its efficiency and reliability, these methods are extremely helpful to the signal-to-noise ratio improvement. In the feature extraction, the common spatial pattern (CSP) method was introduced. CSP is the most efficient and common method in BCI that is used to extract the discriminability classes from EEG signals [30, 31]. This algorithm is a spatial filtering method that can be able to optimize the discriminability of two classes [30].

Furthermore, the number of features CSP components can be reduced by the CSP method; it leads to optimize differences between motor imagery classes. After that,

Mutual Information Best Individual Features (MIBIF) was used to select the best features for classification. The combination of CSP and MIBIF can raise the classification accuracy and Kappa score. In classification algorithms, linear discriminant analysis (LDA) was used to classify features and reduce the dimension of the feature vector.

2 Materials and Methods

In this work, we aim to translate the signal from the brain into a typical movement. To reach this goal, a set of processing phases is required to create the control signal, which comprises pre-processing, feature extraction, feature selection, and classification. An illustration of the processing scheme has been shown in Fig. 1 and the detail of each stage is represented in the following sections.

2.1 Data Acquisition and Dataset

The dataset used for this work is the dataset 2a from BCI Competition IV in 2008 [27]. This data set is provided by the Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology, (Clemens Brunner, Robert Leeb, Gernot Müller-Putz, Alois Schlögl, Gert Pfurtscheller).

The dataset 2a includes EEG data of 9 subjects. For each subject, two sessions were recorded on different days. There are six runs for each session, and each run consisted of 48 trials for four classes which means a single session during the experiment consisted of 288 trials. The cue of each trial was related to four different motor imaginary tasks of the left hand (class 1), right hand (class 2), foot (class 3), and tongue (class 4). The trial begins with a warning tone and a fixation cross displayed on the computer screen at the same time. After two seconds, a cue was shown as a small arrow pointing to the left, right, up, and down, and lasted for 1.25 s. The subjects were prompted to perform the imaginary movement task until the fixation cross disappeared from the screen at $t = 6$ s. After that, there was a short break. Figure 2 illustrates the paradigm of a single trial.

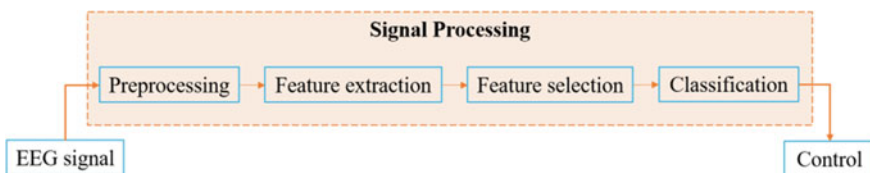


Fig. 1 The workflow of the experiment

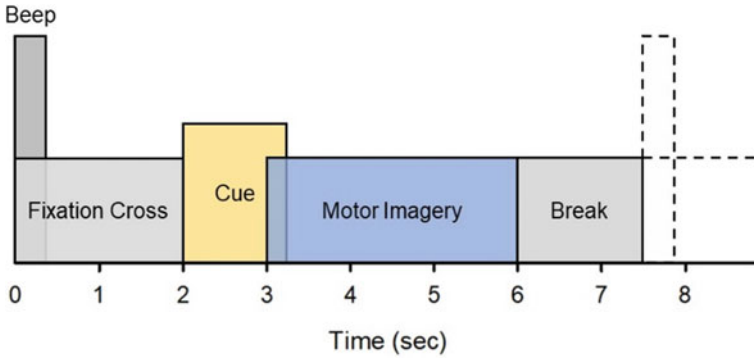


Fig. 2 Timing scheme of the experimental paradigm

Twenty-two EEG channels were recorded using Ag/AgCl electrodes. The specific challenge of this dataset is the eye movement artifact, so three monopolar electro-oculographic (EOG) channels were added for artifact processing purposes. The signals were sampled at 250 Hz and bandpass filtered between 0.5 and 100 Hz with 50 Hz notch filtered.

2.2 Preprocessing

Artifact Removal

As the output of this proposed method is the four different motor imagery classes, the frequency band of brain rhythm is within the ERD/ERS information. A 5th order Butterworth infinite impulse response (IIR) filter with the passband in the range 7–30 Hz to concentrate on the mu (8–13 Hz) and beta (13–30 Hz) rhythms. Additionally, three EOG components were used to utterly “subtract” the EOG artifacts in the EEG signals by linear regression [32]. We assume the regression model as follow:

$$X = E + kO \tag{1}$$

Accordingly, X is the recorded EEG which presented as a matrix of N channels $X = [x_1, x_2, x_3, \dots, x_N]$, E denotes the uncontaminated EEG signal without eye movement artifact, O denotes the pure EOG channel, and k denotes the weight of EOG artifact affecting on N component of EEG signal.

The dataset includes EOG channels for each subject (which electrodes are placed adjacent to the eye to avoid other artifacts). To compute the unknown k , a multiplication of (1) with O^T , where T denotes the recording time-points, yields:

$$\langle O^T . X \rangle = \langle O^T . E \rangle + k \langle O^T . O \rangle \tag{2}$$

As we can assume that the EEG without artifacts and EOG signal have no correlation, the $\langle O^T \cdot E \rangle$ is equal to zero, and coefficient k becomes:

$$k = \langle O^T \cdot X \rangle = \langle O^T \cdot O \rangle^{-1} \quad (3)$$

And the uncontaminated EEG signal can be found according to:

$$E = X - kO \quad (4)$$

This method is also called the “multiple least-squares approaches” since it can remove more than one EOG component. In Graz Dataset 2a, there are 3 EOG channel recordings for each subject, so the algorithm was applied respectively to each channel in order to choose k at the minimum mean square of E .

Resampling

This step lessens the time spent on processing the signal and limits small variations in the data. Regarding the Nyquist theorem [33], the highest frequency of the input signal accounts for half of the sampling rate. Thus, the filtered signal must be sampled more than 60 Hz. One-fourth of the original sampling rate was applied for our work.

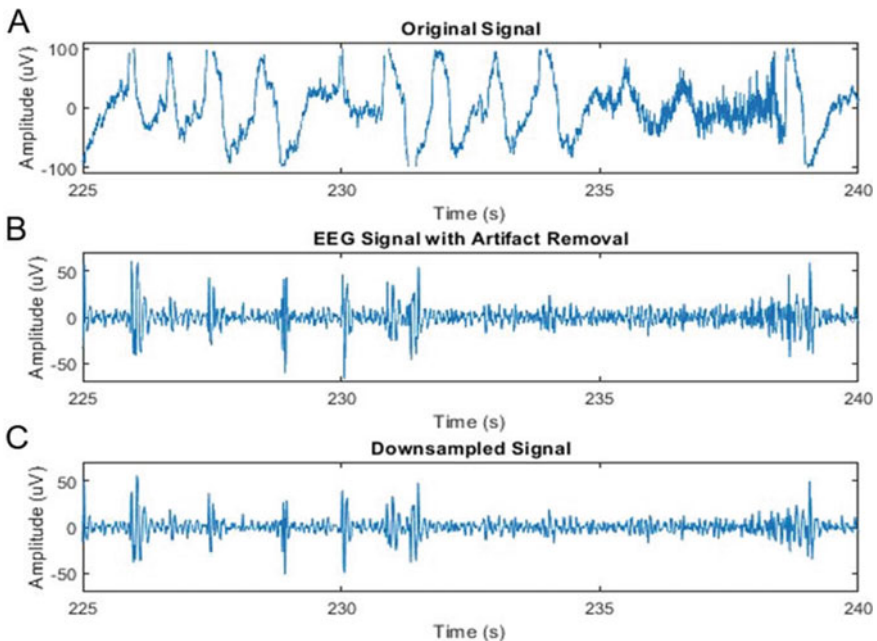


Fig. 3 An illustration of the preprocessing method. **a** Original EEG signal of the channel from subject 1. **b** The filtered signal. **c** The downsampled signal

A visualization of the preprocessing step is presented in Fig. 3. This figure shows the period time from 225 to 240 s of the first channel of subject 1.

2.3 Feature Extraction

Common Spatial Pattern (CSP)

This algorithm can magnify the discriminative of classes by optimizing the variances of the filtered data [34]. Let W^T is a $N \times M$ matrix (M spatial filters for N samples) with T denotes recording time-points and $E(t)$ is the input EEG signal, the CSP model is as follow:

$$E_{CS}(t) = W^T E(t) \quad (5)$$

In terms of the spatial filter, covariance matrices R_c under the $c \in \{1,2\}$ conditions have been shown in the below equation. Though multiple conditions were used for this proposed method, two conditions were applied for the simplicity of the illustration.

$$R_c = \frac{1}{K} \sum_i Z_c^i (Z_c^i)^T \quad (6)$$

Z_c^i denotes the EEG data in the i -th trial at the condition c . Moreover, the CSP also estimates:

$$W^T R_c W = \tau_c \quad (7)$$

$$\tau_1 + \tau_2 = I \quad (8)$$

The sum of two diagonal matrices τ_1 and τ_2 is the identity matrix I . This result illustrates the sum variances must be 1. Thus, if the variance of class 1 is high, the variance of class 2 is accordingly low and vice versa. The results of CSP are estimated by the mean of diagonalization of covariance matrices for two classes.

Band Power and Time Domain Parameters

Band power features were extracted in this method by applying the *BioSig library* [35] to calculate the target bandwidths by bandpass filtering the signal. Our target frequency bands are 8–14, 19–24, and 24–30 Hz. The signal was filtered by the 5th order Butterworth IIR bandpass filter. The bandpass frequencies of this filter are our target bands. To prevent leakage effects that reduce the quality of the processing signal, we apply the smoothing window of 2 s by continue implementing the filter. Finally, natural logarithms of output signals are calculated to enhance to performance of linear classification.

2.4 Feature Selection

Feature selection is a critical step in the classification problem. Some extracted features are irrelevant to the output variable. They do not contribute much to model performance but take time for training. Furthermore, the high dimensionality of data will lead to complicated computation and overfitting results. Therefore, feature selection methods are applied to reduce dimensions of input while maintaining the essential information for classification.

Mutual Information Best Individual Features

Mutual information is a quantity that computes the relationship or correlation between two random variables [36]. The formula of mutual information between two variable X and Y is given by:

$$I(X, Y) = \sum_{x \in X} \sum_{y \in Y} P(x, y) \frac{P(x, y)}{P(x)P(y)} \quad (9)$$

In which, $P(X)$ and $P(Y)$ are the marginal distributions of X and Y .

The Mutual Information (MI) Function

The MI algorithms were developed by Räsänen [37].

The function MI estimates the mutual information and uses the number of variables ranking to score the weight of features based on their individual mutual information with four output classes. The more weight the feature has, the better the performance is when applied to the model.

From the previous step, 24 features corresponding to 3 bandwidths 8–14, 19–24, 2430 Hz are extracted. Mutual information of 24 features is calculated and weighted. Through the experiment, 15 largest-score features are collected.

2.5 Classification

Linear discriminant analysis (LDA) is implemented in our experiment.

Although multiple category classification was done in this work, an illustration for the principle of the LDA algorithm is binary classification due to its simplicity. Specifically, in order to discriminate two classes X and Y , LDA algorithms define a line w such that when the projection of each value of both classes onto the line satisfy two following conditions:

- Maximize the distance between the means of given classes.
- Minimize the variation within each class.

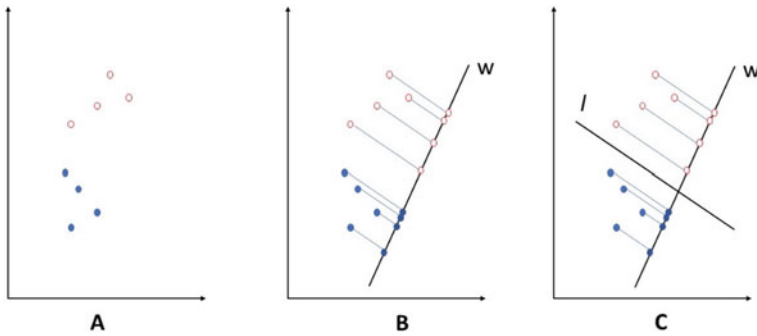


Fig. 4 The projection of red and blue data points representing two classes of the label in the dataset onto different vectors w , is done by *the LDA classifier*. **a** Scatter plot representing data of the two classes. **b** Projection of each data point onto new vector w . **c** Separation of data points of the two classes done by line l

By projecting each value onto line w , data of the two classes are presented in a new axis which is easier to establish the decision boundary l to discriminate them (Fig. 4).

A good model will induce a large distance between the two classes when projecting these data points two classes onto vector w . Therefore, a threshold can be defined to separate the two classes and used to predict new data points.

Regarding the implementation of LDA algorithm on the classification of four mentioned imaginary motor tasks, function *fitcdiscr* of MATLAB is used in classification tasks with:

- Bayesian optimization.
- Iterations: 30.

Through each time of iteration, the optimizer algorithm automatically changes the shape of the objective function model by changing the hyperparameters to obtaining the best fit to the observed data points. Therefore, the classification performance of the model is improved.

3 Evaluation Metric

Cohen's kappa coefficients were calculated to evaluate the performance of the classification model. Kappa scores are considered a robust evaluation of categorized discrimination against the problem as it not only takes into account the percentage of the agreement but also the possibility of chance agreement. Kappa statistics focus on the object-class distribution and ignore the number of features used to find out this distribution, so this measure is useful for imbalanced datasets.

Cohen's kappa equation [38] is defined as:

$$k = \frac{\text{Pr}(a) - \text{Pr}(e)}{1 - \text{Pr}(e)} \quad (10)$$

Pr(a) the relative observed agreement among raters.

Pr(e) the hypothetical probability of chance agreement that is calculated as the probabilities of each observer randomly seeing each category using observed data.

4 Result and Discussion

As mentioned in the evaluation metrics part, the kappa value is suggested to address several accuracy validation problems in BCI research. The different trials can result in different kappa values, so we took the average results of several trials to get a more representative kappa; also, the kappa value will be more general. This paper reached a 0.57 score of mean kappa when using the Graz dataset 2a to train this model. Additionally, we added the testing set from a similar dataset to evaluate our model, and we observed that the mean of evaluation kappa was 0.41, which approaches the training result. Consequently, our model is considered reliable. Table 1 summarizes the kappa coefficient of the training and evaluating set are represented.

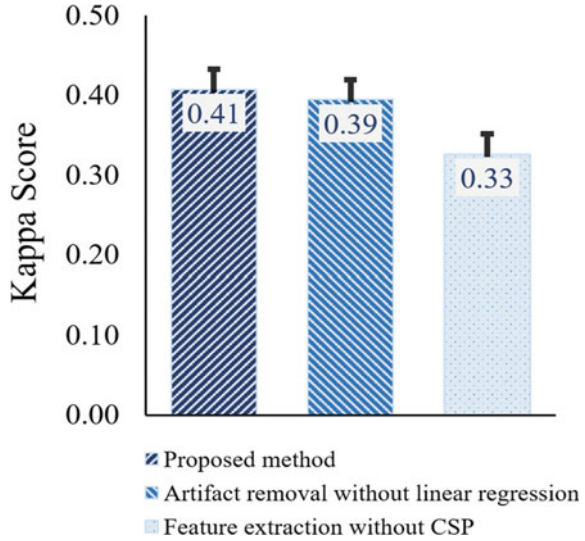
Improvement. To compare the efficiency of our method with traditional methods, we propose two other models. The first model is artifact removal without linear regression, and another is feature extraction without CSP. Figure 5 illustrates that the best model is applying our method ($k = 0.41$), while the model applying linear regression algorithm in artifact removal is less efficient ($k = 0.39$). This experiment also proves that the CSP which is used to extract the features improved the kappa score (from 0.33 to 0.41).

To recognize the role of linear regression and CSP on improving the kappa score, the Kruskal–Wallis test [39] was utilized to compare the difference between the mean

Table 1 The final training and evaluation kappa scores for each subject

Subjects	Training kappa	Evaluating kappa
1	0.72	0.70
2	0.50	0.11
3	0.78	0.52
4	0.35	0.21
5	0.38	0.10
6	0.38	0.18
7	0.76	0.66
8	0.74	0.59
9	0.47	0.61
Mean ± SEM	0.57 ± 0.061	0.41 ± 0.084

Fig. 5 Comparison of the effects of linear regression and CSP to kappa score



value of 3 statistical models. The first reason is that the population data has a small size with nine subjects per each group method and another reason is the non-normal distribution of this data, according to the Levene test ($p = 0.025$) [39, 40]. In the average ranks (see Table 2), the feature extraction without CSP was worth ranked bottom, which proves that the CSP plays an important role in improving our model. On the other hand, using the linear regression method just makes a small impact on the mean value compared to the method without using CSP. Furthermore, Fig. 5 shows that when applying both linear regression and CSP, the mean kappa score was enhanced. Nevertheless, no significant difference in the mean is observed among the three approaches ($p = 0.641$).

Confusion Matrix

To describe the effect of a 4-class classification problem, we use the confusion matrix which illustrates the relationship between the predicted classes (intended labels by user) and the true classes (known labels). Table 3 shows the proportion of samples from evaluation data predicted by the proposed model based on the ground truth, in which the bold ones illustrate the correct classifier. The higher the bold values of the confusion matrix, the more efficient model.

Table 2 The Kruskal–Wallis test summarizes the mean rank of 3 methods

Method	Mean rank
Proposed method	15.33
Artifact removal without linear regression	14.67
Feature extraction without CSP	12.00

Table 3 Confusion matrix of the evaluation data

		Predicted class			
		Left hand	Right hand	Both Feet	Tongue
True class	Left hand	0.64	0.25	0.08	0.03
	Right hand	0.133	0.851	0.009	0.007
	Both feet	0.03	0.04	0.79	0.14
	Tongue	0.004	0.002	0.116	0.878

According to Table 3, tongue motor imagery got the best result with 87.8%. In contrast, the left-hand motor imagery just reached 64% accurate prediction; this class still was significantly confused back and forth with the right-hand class when being predicted (25% left-hand class was labeled right-hand motor imagery). Meanwhile, feet and right-hand motor imagery got considerable results, 79% and 85.1% sequentially. Overall, our model is considered good working in classifying four-class motor imagery (79% in the average).

Comparison with the BCI Competition IV (2008)

Compare with the BCI Competition IV (2008) on the same Graz dataset 2a, our work with removing eye artifact and applying Mutual Information Best Individual Features (MIBIF) and LDA algorithm has yielded promising results. Our kappa is higher than that of the 3rd prize in the competition, which is respectively 0.41 and 0.31 [41] (see Table 4). By using the Filter Bank CSP and Naive Bayes Parzen Window technique, K. K. Ang et al. won the first prize with the 0.57 kappa value [42]. Similar to the proposed method, Liu Guangquan et al. also applied the LDA algorithm, however, the combination of it with the Log variance of the best eight components and Bayesian is the reason why this kappa value holds more weight than the proposed method ($k = 0.52$) [43].

Table 4 The comparison kappa values of 9 subjects with the three best competitors

Subjects	Kai Keng Ang	Liu Guangquan	Wei Song	Proposed method
1	0.68	0.69	0.38	0.70
2	0.42	0.34	0.18	0.11
3	0.75	0.71	0.48	0.52
4	0.48	0.44	0.33	0.21
5	0.40	0.16	0.07	0.10
6	0.27	0.21	0.14	0.18
7	0.77	0.66	0.29	0.66
8	0.75	0.73	0.49	0.59
9	0.61	0.69	0.44	0.61
Mean	0.57	0.52	0.31	0.41

Code Deposit

Our code has been uploaded on the Github: https://github.com/nganluu0903/BCI_project.

Limitation and Future Steps

In this proposed method, the utilized techniques in signal processing procedure (including feature extractions, feature selection, and classification) are considered to be very common. Therefore, the future steps are to select and apply recent advantage techniques (e.g. recurrent neuronal network) to improve the quality of this study. Furthermore, there are various limitations in the evaluated dataset—Graz dataset A which was published in 2008. So, recent datasets or a self-constructed dataset are highly recommended to approach in the next work.

5 Conclusion

The EEG-based BCI is known as the non-invasive method that is growing significantly and bringing many advantages as well as challenges in EEG signal processing [44]. From recent studies, motor imagery based BCI (MI-BCI) has become a common approach; however, it is challenging to classify imaginary motor tasks [32].

In this paper, the proposed algorithm works on the event-related desynchronization (ERD) phenomenon to categorize four different motor imagery classes. We implemented band power features, time-domain parameters, and the common spatial patterns (CSP) to extract features while Mutual Information Best Individual Features was considered in feature selection. Besides, the linear discriminant analysis (LDA) classifier was utilized to classify four classes. The proposed algorithm was evaluated using the kappa score. The kappa value was obtained from the confusion matrix that evaluates the effectiveness of the algorithm and supports to compare the accuracy of models with other ones easily. Besides, the spatial filter is recommended to enhance the kappa score and increase the reliability of the model.

For future work, it is recommended to focus on raising the performance of the MIBCI algorithm by utilizing different techniques in encoding the EEG signal. In this paper, some factors influence the performance of the proposed algorithm. For instance, frequency band selection for each subject is necessary since the effect of the ERD/ERS phenomenon is different for each individual. To optimize this system quality, an automatic method for selecting the frequency bands should be considered before performing preprocessing. Furthermore, another method named the filter bank common spatial patterns (FBCSP) algorithm should be considered for feature extraction in separated smaller frequency bands.

Acknowledgements The Ministry of Science and Technology funds this research, Grant # KC-4.007/19-25.

Conflicts of Interest The authors have no conflict of interest to declare.

References

1. Rao RPN (2013) Brain-computer interfacing: an introduction. Cambridge University Press, Cambridge
2. Panoulas KJ, Hadjileontiadis LJ, Panas SM (2010) Brain-Computer Interface (BCI): types, processing perspectives and applications. In: Tsihrintzis GA, Jain LC (eds) Multimedia services in intelligent environments: integrated systems. Springer, Berlin, Heidelberg, pp 299–321
3. Gu X, Cao Z, Jolfaei A, Xu P, Wu D, Jung T-P, Lin C-T (2020) EEG-based Brain Computer Interfaces (BCIs): a survey of recent studies on signal sensing technologies and computational intelligence approaches and their applications
4. Zhang X, Yao L, Huang C, Gu T, Yang Z, Liu Y (2019) DeepKey: an EEG and gait based dual-authentication system. [arXiv:1706.01606](https://arxiv.org/abs/1706.01606) [cs]
5. Jiang X, Bian G-B, Tian Z (2019) Removal of artifacts from EEG signals: a review. *Sensors* 19:987. <https://doi.org/10.3390/s19050987>
6. Mannan MMN, Kamran MA, Kang S, Jeong MY. Effect of EOG signal filtering on the removal of ocular artifacts and EEG-based brain-computer interface: a comprehensive study. <https://www.hindawi.com/journals/complexity/2018/4853741/>
7. Deep brain stimulation: BCI at large, where are we going to? https://www.researchgate.net/publication/51596032_Deep_brain_stimulation_BCI_at_large_where_are_we_going_to
8. Roy D, Hash L (2016) Advisor: study of brain computer interfacing (BCI) with cochlear implant as an example
9. Wang H, Chang W, Zhang C (2016) Functional brain network and multichannel analysis for the P300-based brain computer interface system of lying detection. *Expert Syst Appl* 53:117–128. <https://doi.org/10.1016/j.eswa.2016.01.024>
10. Maksimenko VA, Runnova AE, Zhuravlev MO, Makarov VV, Nedayvozov V, Grubov VV, Pchelintceva SV, Hramov AE, Pisarchik AN (2017) Visual perception affected by motivation and alertness controlled by a noninvasive brain-computer interface. *PLoS ONE* 12:e0188700. <https://doi.org/10.1371/journal.pone.0188700>
11. Finke A, Lenhardt A, Ritter H (2009) The MindGame: a P300-based brain-computer interface game. *Neural Netw* 22:1329–1333. <https://doi.org/10.1016/j.neunet.2009.07.003>
12. Marchesi M, Riccò B (2013) BRAVO: a brain virtual operator for education exploiting brain-computer interfaces. Presented at the Apr 27
13. Li Y, Long J, Yu T, Yu Z, Wang C, Zhang H, Guan C (2010) An EEG-based BCI system for 2-D cursor control by combining mu/beta rhythm and P300 potential. *IEEE Trans Biomed Eng* 57:2495–2505. <https://doi.org/10.1109/TBME.2010.2055564>
14. Guger C, Harkam W, Hertenæs C, Pfurtscheller G (1999) Prosthetic control by an EEG-based brain-computer interface (BCI). In: Proceedings AAATE 5th European conference for the advancement of assistive technology
15. Chen X, Chen Z, Gao S, Gao X (2014) A high-ITR SSVEP-based BCI speller. *Brain-Comput Interfaces* 1:181–191. <https://doi.org/10.1080/2326263X.2014.944469>
16. Bensch M, Karim AA, Mellinger J, Hinterberger T, Tangermann M, Bogdan M, Rosenstiel W, Birbaumer N (2007) Nessi: an EEG-controlled web browser for severely paralyzed patients. *Comput Intell Neurosci* 2007:e71863. <https://doi.org/10.1155/2007/71863>
17. Chae Y, Jeong J, Jo S (2012) Toward brain-actuated humanoid robots: asynchronous direct control using an EEG-based BCI. *IEEE Trans Rob* 28:1131–1144. <https://doi.org/10.1109/TRO.2012.2201310>
18. dos Santos EM, Cassani R, Falk TH, Fraga FJ (2020) Improved motor imagery brain-computer interface performance via adaptive modulation filtering and two-stage classification. *Biomed Signal Process Control* 57:101812. <https://doi.org/10.1016/j.bspc.2019.101812>
19. A review on motor imagery signal classification for BCI. <https://www.cscjournals.org/library/manuscriptinfo.php?mc=SPIJ-284>
20. Hersche M, Benini L, Rahimi A (2020) Binary models for motor-imagery brain-computer interfaces: sparse random projection and binarized SVM. Presented at the 2nd IEEE international conference on artificial intelligence circuits and systems (AICAS 2020)

21. Lemm S, Müller K-R, Curio G (2009) A generalized framework for quantifying the dynamics of EEG event-related desynchronization. *PLoS Comput Biol* 5:e1000453. <https://doi.org/10.1371/journal.pcbi.1000453>
22. Schlögl A, Supp G (2006) Analyzing event-related EEG data with multivariate autoregressive parameters. *Prog Brain Res* 159:135–147. [https://doi.org/10.1016/S00796123\(06\)59009-0](https://doi.org/10.1016/S00796123(06)59009-0)
23. Toro C, Deuschl G, Thatcher R, Sato S, Kufta C, Hallett M (1994) Event-related desynchronization and movement-related cortical potentials on the ECoG and EEG. *Electroencephalogr Clin Neurophysiol/Evoked Potentials Sect* 93:380–389. [https://doi.org/10.1016/0168-5597\(94\)90126-0](https://doi.org/10.1016/0168-5597(94)90126-0)
24. Percio CD, Infarinato F, Iacoboni M, Marzano N, Soricelli A, Aschieri P, Eusebi F, Babiloni C (2010) Movement-related desynchronization of alpha rhythms is lower in athletes than non-athletes: a high-resolution EEG study. *Clin Neurophysiol* 121:482–491. <https://doi.org/10.1016/j.clinph.2009.12.004>
25. Klimesch W, Pfurtscheller G, Mohl W, Schimke H (1990) Event-related desynchronization, ERD-mapping and hemispheric differences for words and numbers. *Int J Psychophysiol* 8:297–308. [https://doi.org/10.1016/0167-8760\(90\)90020-E](https://doi.org/10.1016/0167-8760(90)90020-E)
26. Pfurtscheller G, Lopes da Silva FH (1999) Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiol* 110:1842–1857. [http://doi.org/10.1016/S1388-2457\(99\)00141-8](http://doi.org/10.1016/S1388-2457(99)00141-8)
27. Brunner C, Leeb R, Müller-Putz G, Schlögl A, Pfurtscheller G (2008) BCI competition 2008—Graz data set A. Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology, 16
28. Singla R, Devgan A, Gogna P, Batra A (2014) Fixation of delayed union or non-union posterior cruciate ligament avulsion fractures. *J Orthop Surg*. <https://doi.org/10.1177/230949901402200118>
29. Fogarty DP, Deering AL, Guo S, Wei Z, Kautz NA, Kandel SA (2006) Minimizing image-processing artifacts in scanning tunneling microscopy using linear-regression fitting. *Rev Sci Instrum* 77. <https://doi.org/10.1063/1.2390633>
30. Lu H, Plataniotis KN, Venetsanopoulos AN (2009) Regularized common spatial patterns with generic learning for EEG signal classification. In: 2009 annual international conference of the IEEE engineering in medicine and biology society, pp 6599–6602
31. Lotte F, Guan C (2011) Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms. *IEEE Trans Biomed Eng* 58:355–362. <https://doi.org/10.1109/TBME.2010.2082539>
32. Wang D, Miao D, Blohm G (2012) Multi-class motor imagery EEG decoding for brain-computer interfaces. *Front Neurosci* 6:151
33. Srinivasan R, Tucker DM, Murias M (1998) Estimating the spatial Nyquist of the human EEG. *Behav Res Methods Instrum Comput* 30:8–19. <https://doi.org/10.3758/BF03209412>
34. Rao RPN, Scherer R (2010) Statistical pattern recognition and machine learning in brain-computer interfaces. In: *Statistical signal processing for neuroscience and neurotechnology*. Elsevier, Amsterdam, pp 335–367
35. Vidaurre C, Sander TH, Schlögl A (2011) BioSig: the free and open source software library for biomedical signal processing. *Comput Intell Neurosci* 2011
36. Learned-Miller EG (2013) Entropy and mutual information
37. Pohjalainen J, Räsänen O, Kadioglu S (2013) Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits. *Comput Speech Lang* 29. <http://doi.org/10.1016/j.csl.2013.11.004>
38. Cohen J (2016) A coefficient of agreement for nominal scales. *Educ Psychol Meas*. <https://doi.org/10.1177/001316446002000104>
39. Stephanie: non parametric data and tests (Distribution Free Tests). <https://www.statisticshowto.com/parametric-and-non-parametric-data/>
40. Stephanie: Levene test for equality of variances. <https://www.statisticshowto.com/levenetest/>
41. Tangermann M, Müller K-R, Aertsen A, Birbaumer N, Braun C, Brunner C, Leeb R, Mehring C, Miller KJ, Müller-Putz GR, Nolte G, Pfurtscheller G, Preissl H, Schalk G, Schlögl A, Vidaurre

- C, Waldert S, Blankertz B (2012) Review of the BCI competition IV. *Front Neurosci* 6. <http://doi.org/10.3389/fnins.2012.00055>
42. Ang KK, Chin ZY, Zhang H, Guan C (2008) Filter bank common spatial pattern (FBCSP) in brain-computer interface. In: 2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence), pp 2390–2397
43. Liu G. BCI competition IV: results. http://www.bbc.de/competition/iv/results/ds2a/LiuGuan_guan_desc.txt
44. Machado S, Araújo F, Paes F, Velasques B, Cunha M, Budde H, Basile LF, Anghinah R, Arias-Carrión O, Cagy M, Piedade R, de Graaf TA, Sack AT, Ribeiro P (2010) EEG-based brain-computer interfaces: an overview of basic concepts and clinical applications in neurorehabilitation. *Rev Neurosci* 21:451–468. <https://doi.org/10.1515/REVNEURO.2010.21.6.451>