Classification of left and right foot kinaesthetic motor imagery using common spatial pattern

Madiha Tariq, Pavel M Trivailo and Milan Simic
School of Engineering, RMIT University, Melbourne, VIC, Australia
E-mail: milan.simic@rmit.edu.au

Keywords: common spatial pattern (CSP), filter bank common spatial pattern (FBCSP), kinaesthetic motor imagery (KMI), brain–computer interface (BCI), supervised machine learning, EEG

Abstract

Background and objectives: Brain–computer interface (BCI) systems typically deploy common spatial pattern (CSP) for feature extraction of mu and beta rhythms based on upper-limbs kinaesthetic motor imageries (KMI). However, it was not used to classify the left versus right foot KMI, due to its location inside the mesial wall of sensorimotor cortex, which makes it difficult to be detected. We report novel classification of mu and beta EEG features, during left and right foot KMI cognitive task, using CSP, and filter bank common spatial pattern (FBCSP) method, to optimize the subject-specific band selection. We initially proposed CSP method, followed by the implementation of FBCSP for optimization of individual spatial patterns, wherein a set of CSP filters was learned, for each of the time/frequency filters in a supervised way. This was followed by the log-variance feature extraction and concatenation of all features (over all chosen spectral-filters). Subsequently, supervised machine learning was implemented, i.e. logistic regression (Logreg) and linear discriminant analysis (LDA), in order to compare the respective foot KMI classification rates. Training and testing data, used in the model, was validated using 10-fold cross validation. Four methodology paradigms are reported, i.e. CSP LDA, CSP Logreg, and FBCSP LDA, FBCSP Logreg. All paradigms resulted in an average classification accuracy rate above the statistical chance level of 60.0% (P < 0.01). On average, FBCSP LDA outperformed remaining paradigms with kappa score of 0.41 and classification accuracy of 70.28% ± 4.23. Similarly, this paradigm enabled discrimination between right and left foot KMI cognitive task at highest accuracy rate i.e. maximum 77.5% with kappa = 0.55 and the area under ROC curve as 0.70 (in single-trial analysis). The proposed novel paradigms, using CSP and FBCSP, established a potential to exploit the left versus right foot imagery classification, in synchronous 2-class BCI for controlling robotic foot, or foot neuroprosthesis.

1. Introduction

Brain–computer interface (BCI) is an augmented muscle-free communication channel between the human brain and output devices for assisting subjects with neuromotor disorders, spinal cord injuries (SCI) or amputated residual limbs [1–4]. It decodes a specific brain activity into computer command to control external device. Amongst the popularly used electro-encephalography (EEG)-based brain activity is event-related (de)synchronization (ERD/ERS) localized in the sensorimotor cortex [5–9]. The ERD/ERS features can be quantified via band-power changes that occur during any kinaesthetic motor imagery (KMI) task performed by the subject, e.g. imagination of limb movement (left-right hand or foot) [5, 7]. Frequency bandwidths, that reflect imaginary activity in EEG, lie in the mu and beta oscillatory activity, i.e. between 7 to ∼35 Hz.

In order to extract ERD/ERS EEG features for BCI, various methods have been introduced based on application requirements [1, 9–10]. According to [11], for a BCI that uses mu and beta rhythms, the selection of spatial filter can markedly affect its signal-to-noise ratio. The common spatial pattern (CSP) is one efficient method that has generally been used with oscillatory processes in the KMI feature extraction due to its simplicity, relatively high speed and robustness.
However, literature reflects that this method has been used with either hand motor imagery (MI), e.g. left versus right hand, or left-hand versus rest, or hand versus basic foot, or tongue movement imageries [12–15]. There was no evidence of left versus right foot discrimination task. Less literature on discrimination of lower-limbs compared to upper limbs is due to the location of lower-limbs representation area near the ‘mantelkante’ in the sensorimotor cortex [1, 7, 8]. It is located deep inside the mesial wall (within the inter-hemispheric fissure). In contrast to upper limbs, hip, knee, foot and toes share spatial proximity with each other that makes it difficult to detect them. This study therefore takes into account the left foot versus right foot dorsiflexion KMI tasks for establishing the basis of a 2-class BCI (that could generate two independent commands) to control 2 degrees of freedom (DOF) robotic foot.

As CSP finds spatial filters, which maximize the variance of the (projected) signal from one class, and minimize it for the other class, it offers a natural approach to efficiently estimate the discriminant information about KMI [16]. Adaptive spatial filter uses the log-variance features over single non-adapted frequency range (that may have multiple peaks), and in the signal, neither temporal structure (variations) is captured, nor the interactions between frequency bands [17]. Successful application of CSP mainly relies on the filter band selection (a wide filter band lies in the 8–35 Hz for KMI classification). However, the most effective frequency band is typically subject-specific that can hardly be determined manually [13]. In order to fix the filter band selection problem, the approaches proposed include simultaneous optimization of spectral filters within the CSP [18–20]; and selection of significant CSP features from multiple frequency bands [21, 22]. Filter bank common spatial pattern (FBCSP) was introduced for the selection of optimal filter bands, through estimation of the mutual information among CSP features in several fixed filter bands [22]. Consequently, we implemented the FBCSP as a further study, to optimize the subject-specific frequency band for CSP across participants.

The FBCSP is an extension of the CSP method. A set of CSP filters is learned for each of the several time/frequency filters, followed by the log-variance feature extraction, the concatenation of all features (over the chosen spectral filters), and finally machine learning. It could be very useful when oscillatory processes in different frequency bands (with different spatial topographies) e.g., mu, low beta and high beta, are jointly active. Their concerted reaction must be taken into account for the given prediction task [17]. In this study, since FBCSP’s feature space dimensionality was larger than in CSP followed by complex interactions, a more complex classifier than linear discriminant analysis (LDA) was additionally deployed to learn the appropriate model. However, with more flexibility comes a risk of overfitting, i.e. a tradeoff, therefore we compared its performance with the standard CSP performance. Since complex (relevant) interactions between mu and beta bands are seemingly rarely observed, the selection of time and frequency regions was critical.

We have focused on the optimal selection of discriminative ERD/ERS features from multiple frequency bands of mu and beta and the effective imagery time window using two feature selection algorithms, CSP and FBCSP, designed in BCILAB (MATLAB toolbox and EEGLAB plugin) [17]. The selected features were concatenated, and two machine learning models, i.e. logistic regression model and LDA, were trained on the selected features, in order to classify the left and right foot KMI tasks. The single-trial classification accuracies used in the training and testing data were validated using 10-fold cross validations for session-to-session transfer with all participants. After testing FBCSP with Logreg and LDA, the resulting accuracy rates were compared to the rates obtained from basic CSP algorithm with Logreg and LDA. The classification performance of each algorithm was statistically evaluated using Cohen’s kappa coefficient κ. With the highest average percentage accuracy of 70.28 ± 4.23, to discriminate between left and right foot KMI, FBCSP-LDA surpassed remaining algorithms, yielding a mean kappa value of 0.41 across all nine participants. This was followed by no experience of BCI protocol in advance, by any participant.

2. Materials and methods

2.1. Participants

This study involved nine healthy participants, with no history of neurological disorder, or any impairment, aged between 21–28 years, who voluntarily participated in the experiments. The participants had no BCI experience either. Ethics approval for the study was granted by the CHEAN (College Human Ethics Advisory Network) of RMIT University, Melbourne, Australia.

Participants were seated in a comfortable chair and were directed to watch a monitor (17") from a distance of approximately 1.5 m. To avoid the possibility of any proprioceptive signals due to muscle movement, a flat wooden sheet was placed underneath the feet of participants. Hence both legs were loosely fixed, with the knees flexed at 60° from full extension position, and ankles at neutral position. In the experiment, participants were directed to dorsiflex their foot approximately 25° for 1 s, analogous to the normal walking gait measurements [23].

2.2. Cortical activity recording

The EEG signal was recorded from 19 scalp electrodes, using neurofeedback BrainMaster Discovery 24E amplifier (BrainMaster Technologies Inc., Bedford, USA); referenced to the linked earlobes A1 and A2 [9]. To acquire EEG signal from the motor cortex, an electrocap with mounted electrodes (C3, C4, Cz, F3, F4, Fl, F2, C5, C6, P3, P4, and reference: linked earlobes)

MATLAB tool-

C3, C4, Cz, F3, F4, Fl, F2, C5, C6, P3, P4, and reference: linked earlobes)
F4, F7, F8, Fz, FP1, FP2, O1, O2, P3, P4, Pz, T3, T4, T5, T6), positioned according to the international 10–20 system [24] was used. Monopolar EEG was amplified and bandpass filtered in the frequency range of 1–100 Hz. All channels were sampled at 256 Hz and quantised with 24-bit resolution with ground electrode located near the forehead of participants. Experimental protocol was designed using OpenViBE designer tool that comes with integrated feature boxes [25, 26].

2.3. Foot motor tasks
Four cue-based sessions were performed without feedback. Each session comprised of 40 trials, with 20 trials for left foot and 20 trials for right foot KMI in a random order. This led to 80 repetitions for each foot KMI task. Prior to the four cue-based KMI sessions, a motor-task practice session, without imagery, was conducted for the participants, in which they dorsiflexed each foot approximately 25° for 1 s (nominal walking gait) post cue. Following this, the KMI sessions were conducted.

In the experimental paradigm as shown in figure 1, each trial was initiated with presentation of a fixation cross on screen for 3 s, used as reference period for processing of epochs. An audio beep of one second, right before the visual cue display, was incorporated in the first trial only, to let the participant pay attention. The temporal sequence of 1 trial is given in figure 1. Next, the visual cues were displayed for 2 s, followed by the display of a blank screen (black), 5 s in length, for MI task performance. This made a total of 10 s for each trial. Following this, a random (pause) interval of 1.5–3.5 s at the end of each trial was incorporated, to prevent fatigue. The visual cues in each trial reflected, either the right, or left foot dorsiflexion image with an arrow pointing in the respective direction. Both cues were displayed in a random order to avoid any adaptation. After recording, EEG signals were processed offline using MATLAB R2013b and BCILAB https://github.com/scnn/BCILAB.

2.4. Feature extraction using CSP and FBCSP
The filter bank common spatial pattern (FBCSP) has four stages involved in signal processing and machine learning, as illustrated in figure 2, adapted from [27]. First, a filter bank that decomposes EEG into multiple frequency pass bands using Chebyshev Type II filter is used. In this case a total of 3 bandpass filters are deployed, 8–12, 13–25, 28–32, covering ranges of mu and beta rhythms. Second stage involved spatial filtering using CSP algorithm. Third stage was the CSP feature selection, and finally the classification of these features based on the left versus right foot KMI tasks. The CSP projection matrix for each filter band, discriminative CSP features, and classifier model are computed from the labelled training data (2-class KMI tasks). Parameters, computed from the training phase, are then used for the testing phase, and finally for the prediction of the single-trial KMI task.

We have initially deployed the common spatial pattern (CSP) method for 2-class discrimination of foot KMI tasks. Based on literature [16] it was demonstrated that, for improving the signal-to-noise ratio, spatial filters overall are useful in single-trial analyses. CSP algorithm transforms the observed EEG signal as:

$$S_{b,j} = W_j^b E_{b,j}$$  \(1\)

where $E_{b,j} \in \mathbb{R}^{1 \times J}$ is the observed single-trial EEG signal from the $b$th bandpass filter (between 7–35 Hz) of the $j$th trial, $j = 1 \ldots n$, where $n$ is the number of training trials.

$W_b$ is the un-mixing matrix (CSP projection matrix) and $S_{b,j}$ is the recovered single-trial sources after spatial filtering, and $T$ denotes transpose operator. The CSP filter computes the un-mixing matrix $W_b$ in order to yield features that have optimal variances for discriminating the classes of measured EEG signal [12, 16, 28], in this case two classes. This is achieved by resolving the eigenvalue decomposition problem:

$$\sum_{k \in 1} W_k = (\Sigma_{b,1} + \Sigma_{b,2}) W_b$$  \(2\)

where $\Sigma_{b,1}$ and $\Sigma_{b,2}$ are the estimates of the covariance matrices of $b$-th bandpass filtered EEG signal based on two imagery tasks i.e. left and right foot movement.
The diagonal matrix $D_b$ consists of the eigenvalues of $\Sigma_{b,1}$, and the column vectors of $W_b^{-1}$ are the filters for CSP projections. For best results, most suitable contrast is provided by filters with the highest and lowest eigenvalues. It is therefore common to retain $e$ eigenvectors from both ends of the eigenvalue spectrum [16]. We used the MATLAB toolbox BCILAB https://github.com/sscn/BCILAB for algorithm implementation. Time window was kept [0 4], whereas for CSP algorithm we used the finite impulse response (FIR) filter for a frequency window of [7, 8, 29, 30].

The CSP filter was applied for a left versus baseline and right versus baseline for each band, in the time segment starting after the cue presentation i.e. task performance duration of 5 s. Furthermore $c = 2$ eigenvectors from the top and from the bottom of the eigenvalue spectrum were retained. This method was implemented on the pre-processed training dataset, that yielded the un-mixing matrix $W_b \in \mathbb{R}^{s \times c}$ and source signals $S_{b,j} \in \mathbb{R}^{s \times t}$, where $s = 2 \times c \times 3$ (frequency bands) $\times 2$ (classes) is the number of sources i.e. the CSP projections, $c$ is the number of channels, the number of time samples is $t$ and $j = 1 \ldots n$, here $n$ is the number of trials of training sets.

When the spatial filtered signal $S_{b,j}$ from (1) uses $W_b$ from (2), it maximizes the difference in variance of the two classes of bandpass filtered EEG signal. The $m$ pairs of CSP features of $j$-th trial for $b$-th band-pass filtered EEG signal are given by:

$$v_{b,j} = \log \frac{\text{diag}(W_b^T E_b E_b^T W_b)}{\text{tr}[W_b^T E_b E_b^T W_b]}$$

where $v_{b,j} \in \mathbb{R}^{2m}$; $W_b$ signifies the first $m$ and the last $m$ columns of $W_b$; diag(.) returns the diagonal elements of the square matrix; tr[..] returns the sum of diagonal elements in the square matrix [27]. Consequently, the FBCSP feature vector for the $j$-th trial is formulated as:

$$v_j = [v_{1,j}, v_{2,j}, \ldots, v_{n,j}]$$

where $v_j \in \mathbb{R}^{1 \times (3 + 2m)}$, $j = 1, 2, \ldots, n$; $n$ represents the total number of trials in data.

The training data, that comprised extracted feature data, is given as (3) and the true class labels is denoted as (6), in order to make a difference from the testing and prediction data,

$$\bar{y} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_n \end{bmatrix}$$

$$\bar{Y} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_n \end{bmatrix}$$

where $\bar{Y} \in \mathbb{R}^{n \times (3 + 2m)}$; $\bar{y} \in \mathbb{R}^{n \times 1}$; and $\bar{y}_j$ and $\bar{y}_j$ are the feature vector and true class label respectively, from the $j$-th training trial, $j = 1, 2, \ldots, n_t$; where $n_t$ represents the total number of trials in training data [27].

2.5. Performance evaluation

This study uses the synchronous i.e. cue-based BCI protocol, therefore a traditional linear discriminant analysis (LDA) was used to measure classification accuracy, as recently reported [29] both for CSP and FBCSP features, respectively. However, in order to enhance the classification accuracy outcome of LDA,
we deployed the logistic regression (Logreg), a supervised machine learning model for both features.

We deployed the cross validation method to estimate the optimal parameters for the classifiers and avoid overfitting classifiers to the training data [31]. The k-fold cross validation estimates the true performance of machine learning model. Classification with each model, for correctly discriminated trials, was performed with 10-fold cross-validation. For each participant data, we partitioned all motor imagery trials into k = 10 folds of equal size, then using k − 1 part as a training set and checking the classification rate on one remaining part (testing set) for prediction accuracy. This is repeated for k times (folds). Consequently, the weight vectors and accuracy on each fold is estimated by calculating the average k classification rates obtained for k testing sets [32]. Feature scaling (regularization) was performed on the training and test sets. The mean and standard deviation of each classifier output was determined. Our proposed methodology resulted in four combinations of models, i.e. CSP-LDA, CSP-Logreg, FBCSP-LDA, and FBCSP-Logreg.

For machine learning based studies, the performance measure of classification model is an essential task. We utilized the area under the receiver operator characteristic curve (AUC-ROC curve). The ROC curve is plotted with the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings that illustrates the diagnostic ability of a binary classifier. It is a probability-curve and AUC signifies the degree of separability/distinguishing between classes [30]. In ROC, along x-axis lies the sensitivity, called true positive rate (TPR):

\[
TPR = \frac{TP}{TP + FN},
\]

where TP is the number of true positives and FN is the number of false negatives. Along y-axis lies 1-specificity, also termed the false positive rate (FPR):

\[
FPR = 1 - \text{Specificity} = \frac{FP}{TN + FP},
\]

where FP is the number of false positives and TN is the number of true negatives. With a higher AUC, model is better at predicting 0 s as 0 s and 1 s as 1 s, i.e. distinguishing between left foot KMI and right foot KML. Ideally, the AUC = 1.

In addition to this, we statistically evaluated the performance of the classifiers, as a measure of distinctiveness between the two classes, by using Cohen’s kappa coefficient \( \kappa \) [33, 34]. In our study of 2-class problem, the evaluation of the classifier is defined by its confusion matrix \( H \), that describes the relationship between the true classes and observed output of the classifier. Given \( H \), the classification accuracy \( ACC \) (overall agreement) is given as:

\[
ACC = \frac{1}{N} \sum_{i} H_{ii}
\]

The chance expected agreement is given as:

\[
P_c = \frac{\sum_{i=1}^{N} n_{ii} n_{ii}}{N^2},
\]

where \( N = \sum_{i=1}^{N} H_{ij} \) is the number of samples, \( H_{ii} \) are the elements of confusion matrix \( H \) on the main diagonal, whereas \( n_{ii} \) and \( n_{io} \) are the sums of each column and row, respectively. Therefore, the estimate of kappa coefficient \( \kappa \) is given as:

\[
\kappa = \frac{P_c - P_e}{1 - P_e},
\]

with the chance probability \( P_e = \frac{1}{M} \) [35].

2.5.1. Test-statistic and family-wise error rate correction

In order to statistically evaluate and compare the outputs of proposed models, two independent samples t-test were conducted on the two groups of each feature (CSP and FBCSP), across participants. The \( p \)-values were used for comparisons, to direct towards the statistically significant features. Multiple comparison corrections were done using the Bonferroni correction, for \( p \)-values adjustment.

The observed \( p \)-values obtained from LDA and Logreg classifier models were corrected for CSP and FBCSP features, respectively as shown in the schematic below.

Results from LDA model were compared to Logreg model for CSP feature; similarly results from LDA were compared to those from Logreg for FBCSP. To calculate the family-wise error rate (12) is used, from [36].

\[
\alpha_{FW} = 1 - (1 - \alpha_{PC})^c,
\]

where \( \alpha_{FW} \) is the family wise error rate, \( \alpha_{PC} \) is the indicated per comparison error rate, and \( c \) is the number of comparisons performed. In this research, \( \alpha_{PC} = 0.05 \), with two statistical analyses conducted on the same sample of data, \( c = 2 \), the Bonferroni correction is given by (13).

\[
\frac{\alpha_{PC}}{c}
\]

Any observed \( p \)-value less than the corrected \( p \)-value of 0.025 is confirmed to be statistically significant.
3. Results

During the experimental trials, all participants successfully performed the tasks, as per instructions. There was no report from any participant about fatigue or anxiety during experiment.

3.1. Common spatial pattern scalp projections

As the anatomical properties of cortical folding between people are different [37, 38], there is a reason for maximum discrimination power for the ERD/ERS characteristic of foot movement and MI during experiment are not strictly located beneath electrode positions C3, Cz and C4 [14]. For this reason, the CSP method generates subject-specific spatial filters that are optimized for discrimination among the two experimental tasks. It spatially filters the raw EEG channels to smaller time-series, whose variances are optimized to discriminate the two classes, i.e. left foot KMI and right foot KMI.

For each participant a 3-pair set of CSP scalp projections were generated, however to save space, the figure 3 illustrates a 3-pair set of CSP scalp projections for one participant, P01. Each CSP filter contains 2 patterns that illustrate how the signal projects to scalp through training data generated by FBCSP. During right and left foot KMI tasks, CSP pattern 1 reflects the time invariant EEG source distribution vectors i.e. the sensorimotor area activation around channel C3, this confirms the cortical lateralization of ERD/ERS. Similarly, the CSP pattern 3 elicits the contralateral dominance at channel C4, whereas, the pattern 2 is focused at the vertex (Cz). The CSP patterns 1 and 3 are concentrated in the contralateral hand area representation of the cortex. On the contrary CSP pattern 2 is centrally-focused around the vertex of sensorimotor cortex which is the foot area representation, as established by [5].

Figure 3. A set of common spatial patterns (CSPs) filters of participant P01, where CSPs are optimized for the discrimination of right and left foot kinaesthetic motor imageries with respect to the reference period.

Table 1. The 10-fold cross-validation performance of misclassification rate using CSP and FBCSP with linear discriminant analysis (LDA) and logistic regression (Logreg) classifiers.

| Participant | CSP LDA (mcr (%)) | Logreg (mcr (%)) | FBCSP LDA (mcr (%)) | Logreg (mcr (%)) |
|-------------|-------------------|------------------|---------------------|------------------|
| P01         | 27.50             | 25.00            | 22.50               | 30.00            |
| P02         | 32.50             | 37.50            | 25.00               | 32.50            |
| P03         | 35.00             | 40.00            | 30.00               | 35.00            |
| P04         | 37.50             | 40.00            | 27.50               | 35.00            |
| P05         | 30.00             | 37.50            | 30.00               | 40.00            |
| P06         | 32.50             | 35.00            | 30.00               | 37.50            |
| P07         | 37.50             | 37.50            | 35.00               | 40.00            |
| P08         | 35.00             | 40.00            | 32.50               | 35.00            |
| P09         | 37.50             | 40.00            | 35.00               | 37.50            |
| Average     | 33.89             | 36.94            | 29.72               | 35.83            |
| S.D.        | 3.56              | 4.81             | 4.23                | 3.31             |

* Over chance level of 2-class discrimination, 57.50% (p < 0.05).
* Over chance level of 2-class discrimination, 60.00% (p < 0.01).

3.2. Classification accuracy and KMI task discrimination

While in general, the CSP scalp projections clearly revealed the discrimination of left-right foot imageries, there were cases where the projections did not exhibit strong left-right difference. Nevertheless, even if a slight left-right difference is shown by the BCI user, it is probable to enhance the difference and increase the control accuracy of BCI using machine learning [39].

Table 1, illustrates the misclassification rate (mcr) for nine participants using two feature vectors, and applying two different machine learning models on each feature vector individually. We began with the CSP features, in order to compare the results with FBCSP features. The CSP features were used for training and testing LDA model first. Following this, the
Logreg was trained and tested. Both models resulted in prediction of misclassification rates (in percentage) for each participant. Similar approach was used for the FBCSP feature vector. In all four cases, models were cross-validated using 10-folds, for training and testing data. Majority of the participants performed above the statistical chance level of $p < 0.01$ with both classifiers. Remaining performances exceeded the chance level of 57.5% for $p < 0.05$. Participant, P01 performed the best amongst all with the lowest mcr in case of FBCSP-LDA $= 22.50\%$ and in case of CSP-Logreg $= 25.00\%$. The average mcr for nine participants with CSP-LDA was $33.89 \pm 3.56$, with CSP-Logreg it was $36.94 \pm 4.81$, with FBCSP-LDA it was $29.72 \pm 4.23$, and with FBCSP-Logreg came out to be $35.83 \pm 3.31$. This implies that average classification accuracies of all models are clearly well above the chance level of a 2-class discrimination BCI problem which, according to [40], should be 57.5% ($p < 0.05$) or 60.0% ($p < 0.01$) for a total of 80 trials.

The area under ROC curve (AUC) for each participant is shown in figure 4, where x-axis denotes the FPR and y-axis denotes the TPR. The dark blue curve represents the CSP-Logreg output, green represents FBCSP-Logreg, yellow-chartreuse curve reflects CSP-LDA, and maroon curve represents FBCSP-LDA. In all graphs the grey line signifies the 50% chance level for the binary classifier. For ideal detection AUC should be 1. As discussed earlier, the chance level of a 2-class discrimination BCI problem should be above or equal to 57.5% ($p < 0.05$) or 60.0% ($p < 0.01$). In each case it is evident that the participants obtained the four respective AUCs above the chance level. From table 2, it can be realized that participant P01 exhibited maximum AUC of 0.74 with CSP-Logreg, followed by AUC = 0.73 with CSP-LDA. The maximum average AUC in case of CSP was with LDA, i.e. 0.62 ± 0.06, in case of FBCSP, it was with LDA as well, i.e. 0.64 ± 0.04. FBCSP overall exhibited maximum average AUC amongst the four deployed models, although the average AUC difference amongst all the models was not much.

Table 2 also projects the kappa statistic scores. During study, it is observed that all kappa scores are above chance level ($\kappa = 0$). The average scores range

| Participant | CSP LDA | CSP Logreg | FBCSP LDA | FBCSP Logreg |
|-------------|---------|------------|-----------|--------------|
|             | AUC     | $\kappa$   | AUC       | $\kappa$     |
| P01         | 0.73    | 0.45       | 0.74      | 0.50         |
| P02         | 0.65    | 0.35       | 0.67      | 0.25         |
| P03         | 0.65    | 0.30       | 0.61      | 0.20         |
| P04         | 0.59    | 0.25       | 0.57      | 0.20         |
| P05         | 0.61    | 0.40       | 0.62      | 0.25         |
| P06         | 0.62    | 0.35       | 0.60      | 0.30         |
| P07         | 0.60    | 0.25       | 0.59      | 0.25         |
| P08         | 0.56    | 0.30       | 0.55      | 0.20         |
| P09         | 0.55    | 0.25       | 0.57      | 0.20         |
| Average     | 0.62    | 0.32       | 0.61      | 0.26         |
| S.D.        | 0.06    | 0.07       | 0.06      | 0.10         |

Figure 4. Receiver operator characteristics curves reflecting area under the curves for all participants.

Table 2. The 10-fold cross-validation performance in terms of maximum kappa value and the area under ROC Curve (AUC) using CSP and FBCSP with linear discriminant analysis (LDA) and logistic regression (Logreg) models.
between fair and moderate performance, based on the study from [41]. This implies that in the 0.21–0.40 range, the strength of agreement between the predicted and true class is fair, whereas between 0.41–0.60 it is moderate.

Importantly there is not much difference between average scores of CSP-Logreg and FBCSP-Logreg, given as 0.26, and 0.28, respectively. The maximum score was obtained for participant P01, using FBCSP-LDA i.e. 0.55. Similarly, FBCSP-LDA score surpassed remaining models with the maximum average score of 0.41 in discriminating between the two classes.

We therefore can clearly state that FBCSP feature with LDA gave the best 2-class discrimination accuracy than the other feature models exceeding the chance level 60% at \( p < 0.01 \), with highest AUC and \( \kappa \) as shown in figure 5. On average LDA classifier outperformed Logreg with FBCSP, but in case of CSP feature vector, both models resulted in a close average accuracy with a difference of approximately 3%. From figure 6, individual participant performance can be viewed for each model; it can be observed that P01 exhibited minimum mcr comparatively.

### 4. Discussion

In the research reported here, we have analysed \( \mu \) and \( \beta \) EEG features, using the CSP and FBCSP feature extraction methods, following machine learning to classify the left foot and right foot KMI. The proposed models deployed LDA and Logreg algorithms for discrimination of left and right foot KMI tasks. We used the CSP filter patterns for analysis of time invariant EEG source distribution vectors, that elicit upon visual cues i.e. cue-based synchronous BCI (Graz BCI protocol). Overall the CSP patterns implicated the cortical lateralization of ERD/ERS during the left and right foot dorsiflexion KMI. The first pair of CSP pattern, exhibited a centrally focal ERD/SP at the vertex channel Cz. From [5], it is a well-established fact that 'hand motor imagery activates neural networks in the cortical hand representation area which is

![Figure 5](image1.png)

**Figure 5.** Average classification accuracies (in percentage) for each algorithm across participants, red dotted line shows average on and above chance level (\( p < 0.01 \)). The error bars represent standard deviations.

![Figure 6](image2.png)

**Figure 6.** Resulting misclassification rate (in percentage) of CSP-LDA, CSP-Logreg, FBCSP-LDA, and FBCSP-Logreg algorithms for individual participant (\( N = 9 \)). The error bars represent standard deviations.
4.1. FBCSP-LDA model

A previous study [39] underscored that the CSP method could be used in EEG-based classification of left and right foot MI. We therefore experimented with CSP method to improve the performance of our 2-class foot KMI. Initially classical LDA was deployed, that ensued in a classification accuracy of 66.11 ± 3.56 with a kappa score of 0.32, but for the improvement, Logreg algorithm was tested and that resulted in an accuracy of 63.06 ± 4.81 with kappa score of 0.26. This pointed to a difference of approximately 3% in the results of both models, as illustrated in figure 5. The average classification accuracy of both models resulted in above the chance level of 60.0% (p < 0.01) for 80 trials, as described by [40], with 10-fold cross validation. However, compared to the band-power method, the accuracy was low [39]. The study therefore took into account the FBCSP procedure to further improve results obtained by CSP method. FBCSP, in conjunction with LDA and Logreg, resulted in average accuracies of 70.28 ± 4.23 and 64.17 ± 3.31, respectively with average kappa scores of 0.41 and 0.28, respectively. This implied that the maximum 2-class accuracy for left-right foot KMI was with FBCSP-LDA, using 10-fold cross validation as shown in figure 5. For the same model, participant P01 scored the highest kappa of 0.55, following maximum classification accuracy of 77.5% among other participants, as reflected in figure 6.

With the proposed FBCSP model, the selection of time and frequency regions was critical, because there are complex interactions between mu and beta bands which are seemingly rarely observed. Therefore, the frequency regions are defined elaborately in FBCSP method that results in large space dimensionality. Since FBCSP’s feature space dimensionality is larger than CSP’s, there is a tradeoff between more flexibility and the risk of overfitting. We therefore compared the performance of FBCSP with the standard CSP [13] and used the family-wise error rate. For multiple comparison corrections, the Bonferroni correction was deployed. Consequently, adjusted p-values are used in the study. According to [43], larger number of training trials and longer length of the experimental trial could prevent overfitting. Following this, it is overall observed that a tradeoff also exists between the classifier models and feature vector strategy, i.e. if FBCSP-LDA performs high, CSP-LDA performs lower, and similarly if FBCSP-Logreg performs better, CSP-Logreg elicited a lower performance.

Some closely related methods for EEG feature optimization and classification based on MI have recently been reported. Zhichao Jin, et al [44] proposed a sparse Bayesian extreme learning machine (SBELM)-based algorithm to improve the classification performance of MI based BCI. The method ‘automatically controls the model complexity and excludes redundant hidden neurons by combining advantages of both ELM and sparse Bayesian learning’. In another recent review [45], authors compared the traditional classification methods with deep learning techniques. With a comprehensive analysis they concluded that ‘deep learning not only enables to learn high-level features automatically from BCI signals, but also depends less on manual-crafted features and domain knowledge’. For EEG-based BCI studies that deploy MI, discriminative models such as, multi-layer perceptron (MLP), recurrent neural networks (RNN), or convolutional neural networks (CNN), overall elicit highest classification accuracies in BCI applications.

Further useful methods include a sparse group representation model (SGRM) for increasing the efficiency of MI-based BCI, presented lately [46]. Using CSP features, a dictionary matrix is constructed with training samples from both the target and other subjects. The optimal representation of a test sample of the target subject is estimated as a linear combination of columns in the dictionary matrix, by exploiting within-group and group-wise sparse constraints. Consequently, classification is done by calculating the class-specific representation based on the significant training samples corresponding to the nonzero representation coefficients. This effectively reduces the required training samples from target subject because of auxiliary data available from other subjects. Using left versus right hand MI, their study depicted a kappa score of 0.57 and 0.55 for two datasets respectively. Recently, a novel algorithm, temporally constrained sparse group spatial pattern (TSGSP) has been presented [47]. It concurrently optimizes filter bands and time-windows within CSP in order to enhance EEG based MI classification. Their classification results were 88.5%, 83.3%, and 84.3%, for 4-class MI left hand, right hand, feet, tongue, for 2-class MI left versus right hand, and for 4-class MI left hand, right hand, feet, tongue, respectively.

4.2. Band-power feature for classification

The band power or time-frequency method has successfully been used in numerous (offline) BCI studies based on MI [5, 7, 9, 39, 48]. This study is based
on foot KMI. We therefore followed the same experimental paradigm as in the earlier studies [7, 9, 39, 10], i.e. with no prior feedback training. Contradictory to [39], the CSP method did not improve the performance of the left-right foot KMI BCI system, however the FBCSP-LDA model improved the average performance for nine participants. Although both CSP and FBCSP resulted in classification accuracies above the statistical chance level of 60.00% ($p < 0.01$), it was less than the maximum accuracy of 81.6% in single trial analysis [39], i.e. a maximum accuracy of 77.5% was attained in our case. However, the average accuracy of band power method using LDA was 69.3% ± 6.1 [39], whereas our study resulted in improved accuracy of 70.28 ± 4.23. The maximum average kappa statistic for this study is in the 0.41 < 0.60 range i.e. the strength of agreement between classes is moderate [41]. Participant P01 outperformed with a kappa statistic of 0.55 which is also in the moderate range. The strength of agreement between classes needs to be more substantial, which is not in case of CSP-Logreg and FBCSP-Logreg method.

This was followed by no practice (no feedback training) of BCI in advance, that could mark a difference in results [49]. As suggested by [43], larger number of training trials could also prevent overfitting and improve results.

Based on our experimental outcomes, we would suggest an alteration in the experimental protocol, i.e. it could be modified by the inclusion of feedback training, since training without feedback might be inclusive of irrelevant imageries. In future we aim at increasing the practice sessions as well. Furthermore, the suggested methodology procedures from our study could potentially be deployed by BCI systems run by multiple users, as its decoding technique could allow for the selection of optimal feature bands suitable for multiple users. The FBCSP could be exploited in combination with neural networks to investigate for an enhancement in the classification accuracy of foot KMI.

5. Conclusions

In this study, we proposed the novel approach to incorporate CSP and FBCSP in conjunction with LDA and Logreg model for the selection of significant filter bands, to improve the left-right foot KMI classification accuracy. FBCSP feature with LDA resulted in highest discrimination accuracy than the other feature models exceeding the chance level 60% at $p < 0.01$ with 10-fold cross validation and the highest $\kappa$ statistic. These results encourage the classification of left-right foot KMI and can be exploited as control commands in a bionic foot-BCI operation or a foot neuroprosthesis. The left-right foot KMI discrimination results are encouraging in view of the covert anatomical representation area of foot in the human sensorimotor cortex compared to that of the hand. We next aim to monitor the repetitive use of neurofeedback training and its effects on classification accuracy.

Acknowledgments

None.

Conflict of interest

None of the authors have potential conflicts of interest to be disclosed.

Competing interest

Authors declare no competing interests.

ORCID iDs

Pavel M Trivailo @ https://orcid.org/0000-0002-3129-8044

Milan Simic @ https://orcid.org/0000-0002-3403-4765

References

[1] Tariq M, Trivailo P and Simic M 2018 EEG-based BCI control schemes for lower-limb assistive-robots Frontiers in Human Neuroscience 12 312

[2] He Y et al 2018 Brain–machine interfaces for controlling lower-limb powered robotic systems J. Neural Eng. 15 021004

[3] Lebedev M A and Nicolelis M A 2017 Brain–machine interfaces: from basic science to neuroprostheses and neurorehabilitation Physiol. Res. 97 767–837

[4] Cervera MA et al 2018 Brain–computer interfaces for post-stroke motor rehabilitation: a meta-analysis Annals of Clinical and Translational Neurology 5 651–63

[5] Pfurtscheller G et al 2006 Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks Neuro. Image 31 153–9

[6] Tariq M, Trivailo P and Simic M 2018 Event-related changes detection in sensorimotor rhythm Int. Rob. Autom. J. 4 119–20

[7] Tariq M et al 2017 Mu-beta rhythm ERD/ERS quantification for foot motor execution and imagery tasks in BCI applications 2017 8th IEEE Int. Conf. on Cognitive Infocommunications (CogInfoCom) (Debrecen, Hungary) (IEEE) pp 000091–96

[8] Tariq M, Trivailo P M and Simic M 2017 Detection of knee motor imagery by Mu ERD/ERS quantification for BCI based neurorehabilitation applications 2017 11th Asian Control Conf. (ASCC) (Gold Coast, QLD, Australia) (IEEE) pp 2215–19

[9] Tariq M et al 2019 Comparison of event-related changes in oscillatory activity during different cognitive imaginary movements within same lower-limb Acta Politechnica Hungarica. 16 77–92

[10] Neuper C and Pfurtscheller G 1996 Post-movement synchronization of beta rhythms in the EEG over the cortical foot area in man Neurosci. Lett. 216 17–20

[11] McFarland D J et al 1997 Spatial filter selection for EEG-based communication Electroencephalogr. Clin. Neurophysiol. 103 386–94

[12] Ramoser H, Muller-Gerking J and Pfurtscheller G 2000 Optimal spatial filtering of single trial EEG during imagined hand movement IEEE Trans. Rehabil. Eng. 8 441–6

[13] Zhang Y et al 2015 Optimizing spatial patterns with sparse filter bands for motor-imagery based brain–computer interface J. Neurosci. Methods 235 85–91
[14] Rozado D, Duenser A and Howell B 2015 Improving the performance of an EEG-based motor imagery brain computer interface using task evoked changes in pupil diameter PLoS One 10 e0121262
[15] Kyogonakis I et al 2018 Decoding motor imagery through common spatial pattern filters at the EEG source space Computational Intelligence and Neuroscience 2018 1–10
[16] Blankertz B et al 2008 Optimizing spatial filters for robust EEG single-trial analysis IEEE Signal Process Mag. 25 41–56
[17] Kothe C A and Makeig S 2013 BCILAB: a platform for brain–computer interface development J. Neural Eng. 10 056014
[18] Higashi H and Tanaka T 2013 Simultaneous design of FIR filter banks and spatial patterns for EEG signal classification IEEE Trans. Biomed. Eng. 60 1100–10
[19] Dornhege G et al 2006 Combined optimization of spatial and temporal filters for improving brain–computer interfacing IEEE Trans. Biomed. Eng. 53 2274–81
[20] Lemm S et al 2005 Spatio-spectral filters for improving the classification of single trial EEG IEEE Trans. Biomed. Eng. 52 1541–8
[21] Thomas K P et al 2009 A new discriminative common spatial pattern method for motor imagery brain–computer interfaces IEEE Trans. Biomed. Eng. 56 2730–3
[22] Ang K K et al 2008 Filter bank common spatial pattern (FB CSP) in brain–computer interface 2008 IEEE Int. Joint Conf. on Neural Networks (IEEE World Congress on Computational Intelligence) (Hong Kong, China) (IEEE) (https://doi.org/10.1109/IJCNN.2008.4634130)
[23] Tariq M, Koreshi Z and Trivailo P 2017 Optimal control of an active prosthetic ankle 3rd Int. Conf. on Mechatronics and Robotics Engineering (ICMRE) (Paris, France) pp 113–118
[24] Klem G H et al 1999 The ten–twenty electrode system of the international federation Electroencephalogr. Clin. Neurophysiol. 52 3–6 (https://pdfs.semanticscholar.org/33a7/cfe8b8568c6e6204c080125e5836d5079b.pdf)
[25] Renard Y et al 2010 Openvibe: an open-source software platform to design, test, and use brain–computer interfaces in real and virtual environments Presence: Teleoperators and Virtual Environments 19 35–53
[26] Tariq M, Trivailo P M and Simic M 2018 Motor imagery based EEG features visualization for BCI applications Procedia Computer Science 126 1936–44
[27] Ang K K et al 2012 Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b Frontiers in Neuroscience 6 39
[28] Blankertz B et al 2008 The Berlin Brain–Computer Interface: Accurate performance from first–session in BCI-naive subjects IEEE Trans. Biomed. Eng. 55 2452–62
[29] Lotto F et al 2018 A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update J. Neural Eng. 15 031005
[30] Tharwat A 2018 Classification assessment methods Applied Computing and Informatics Article in press (https://doi.org/10.1016/j.aci.2018.08.003)
[31] Mitchell T M 1997 Machine Learning, 1 (United States of America: McGraw-hill Science) 432 0070428077
[32] Yong X and Menon C 2015 EEG classification of different imaginary movements within the same limb PLoS One 10 e0121896
[33] Kraemer H C 2014 Kappa coefficient Wiley StatsRef: Statistics Reference Online 14 1–4
[34] Lotto F et al 2007 A review of classification algorithms for EEG-based brain–computer interfaces J. Neural Eng. 4 R1
[35] Cohen J 1960 A coefficient of agreement for nominal scales Educational and Psychological Measurement 20 37–46
[36] Tamhane A and Dunlop D 2000 Statistics and data analysis: from elementary to intermediate (New York: Prentice-Hall) p 722
[37] DeFelipe J 2011 The evolution of the brain, the human nature of cortical circuits, and intellectual creativity Frontiers in neuroanatomy 5 29
[38] Mueller S et al 2013 Individual variability in functional connectivity architecture of the human brain Neuron 77 586–95
[39] Hashimoto Y and Ushiba J 2013 EEG-based classification of imaginary left and right foot movements using beta rebound Clinical Neurophysiology 124 2153–60
[40] Muller–Putz G et al 2008 Better than random: a closer look on BCI results International Journal of Bioelectromagnetism. 10 52–55 (http://www.ijbem.org)
[41] Landis J R and Koch G G 1977 The measurement of observer agreement for categorical data Biometrics 159–74
[42] Neuper C, Wörtz M and Pfurtscheller G 2006 EDR/ERS patterns reflecting sensorimotor activation and deactivation Progress in brain research. 159 211–22
[43] Huang G et al 2010 Model based generalization analysis of common spatial pattern in brain computer interfaces Cognitive neurodynamics. 4 217–23
[44] Jin Z et al 2018 EEG classification using sparse Bayesian extreme learning machine for brain–computer interface Neural Computing and Applications 1–9
[45] Zhang X et al 2019 A survey on deep learning based brain computer interface: recent advances and new frontiers 1 1–35
[46] Jiao Y et al 2018 Sparse group representation model for motor imagery EEG classification IEEE Journal of Biomedical and Health Informatics 23 631–41
[47] Zhang Y et al 2018 Temporally constrained sparse group spatial patterns for motor imagery BCI IEEE Transactions on Cybernetics 49 3522–32
[48] Liu Y et al 2019 Analysis of electroencephalography event-related desynchronisation and synchronisation induced by lower-limb stepping motor imagery Journal of Medical and Biological Engineering 39 54–69
[49] Orand A et al 2012 The comparison of motor learning performance with and without feedback Somatosensory & Motor Research 29 103–10