



WaveCSP: a robust motor imagery classifier for consumer EEG devices

Mohamed Athif^{1,2} · Hongliang Ren¹

Received: 10 March 2018 / Accepted: 7 January 2019 / Published online: 22 January 2019
© Australasian College of Physical Scientists and Engineers in Medicine 2019

Abstract

There is an increasing demand for reliable motor imagery (MI) classification algorithms for applications in consumer level brain-computer interfacing (BCI). For the practical use, such algorithms must be robust to both device limitations and subject variability, which make MI classification a challenging task. This study proposes methods to study the effect of limitations including a limited number of electrodes, limited spatial distribution of electrodes, lower signal quality, subject variabilities and BCI literacy, on the performance of MI classification. To mitigate these limitations, we propose a machine learning approach, WaveCSP that uses 24 features extracted from EEG signals using wavelet transform and common spatial pattern (CSP) filtering techniques. The algorithm shows better performance in terms of subject variability compared to existing work. The application of WaveCSP to Physionet MI database shows more than 50% of the 109 subjects achieving accuracy higher than 64%. The data obtained from a commercial EEG headset using the same experimental protocol result in up to four out of five subjects who had prior BCI experience (out of a total of 25 subjects) performing with accuracy higher than 64%.

Keywords Brain Computer Interfacing · Electroencephalogram · Motor imagery classification · Wavelet decomposition · Common spatial patterns · Machine learning

Introduction

Brain-computer interfacing (BCI) technology is becoming increasingly popular with the availability of commercial electroencephalogram (EEG) acquiring devices. Compared to medical grade EEG devices, commercial devices are low cost and more portable. They do not require special expertise to use and are more comfortable to wear. Therefore, BCI systems implemented using these devices have a potential use in a variety of entertainment [1–3], aesthetic [4], rehabilitative [5, 6], therapeutic [7], prosthetic [8] and personal health monitoring [9] applications in addition to conventional applications in clinical healthcare and research.

Generally, a BCI system works by extracting information about a user's intention from his or her EEG activity

and translating it into a control signal. Motor imagery (MI) is one of the popular paradigms of extracting information from EEG. Here, imagined or planned movements of one's body parts can be identified using the information extracted from EEG. Classifying different MI is, therefore, a crucial component of such a BCI.

Compared to other EEG paradigms such as steady state visually evoked potentials (SSVEP), MI is less strenuous for the user. Popular alternatives such as SSVEP or oddball paradigms using P300 potential require the user to be attentive throughout complicated and extensive series of external visual or auditory stimuli. Many studies have reported this to cause discomfort to users when used for prolonged periods [10, 11]. However, MI does not depend on visual stimulations to elicit EEG potentials. It may require the subject to respond to cues only during the training phase. In practical BCI applications such as prosthetics, rehabilitation, and gaming, MI is more intuitive and natural as control signals compared to other paradigms.

However, the identification and classification of MI are more challenging compared to SSVEP or P300 [12]. Because of this, even if MI is an ideal paradigm in the user's end, practical application of MI as a BCI paradigm is limited [13]. In addition to this, the lower signal quality

✉ Hongliang Ren
ren@nus.edu.sg; hlren@ieee.org

Mohamed Athif
140052X@uom.lk

¹ Department of Biomedical Engineering, National University of Singapore, Singapore, Singapore

² Department of Electronic and Telecommunication Engineering, University of Moratuwa, Moratuwa, Sri Lanka

and limited distribution density of electrodes in commercial EEG devices, in general, make MI detection more difficult when such devices are used. It is also important for MI classification algorithms to be robust and accurate for a wide range of subjects with varying degrees of BCI literacy [8]. Therefore, it is required to improve the performance of MI identification and classification if this user-friendly paradigm is to be successfully implemented in practically used consumer-level BCI.

Many recent studies have been undertaken in the field of MI classification. Several feature extraction methods have been used including common spatial patterns (CSP) [14], cross-correlation [15] and wavelet decomposition [16]. Proposed MI classification algorithms include machine learning algorithms such as naive Bayes classifiers [14], K nearest neighbors (KNN), support vector machines (SVM), linear discriminant analysis (LDA), decision trees [15, 16], ensemble classifiers [16], voting methods [15], neural networks [17] and convolutional neural networks [18, 19].

Despite the abundance of work related to general MI classification, there are only a few studies that focus specifically on MI classification with commercial EEG devices. Vamvakousis and Ramirez [20] have studied whether the EPOC headset is suitable to detect the presence or absence of imaginary and actual toe movement, however, the ability to differentiate between two imaginary movements have not been studied here. Fakhruzzaman et al. [21] have also studied the suitability of the commercial headset to identify MI from the lack of it. Their analysis using CSP algorithm available through the OpenViBE software [22] concludes that the low-cost device is not recommendable for BCI applications. The algorithmic component of these work is similar to those used with clinical grade EEG devices and therefore lack robustness required in an application with commercial BCI. Takehara et al. [23] have compared the headset and a medical grade device for the effectiveness of using signal power as a discriminant feature to classify MI and found the low-cost device to be satisfactory for BCI applications. Martinez-Leon et al. [24] has compared the performance of a MI classification based on dFasArt models proposed by Cano-Izquierdo et al. [25] on EEG data acquired from the headset against BCI competition IV dataset and arrived at similar conclusions. A few studies have focused on optimizing MI classification algorithms to suit BCI implemented with commercial EEG devices. Hurtado-Rincon et al. [26] have obtained a set of features used in MI classification from EEG recorded using the same device and have proposed a feature relevance study to select the best discriminating features. Yang et al. [27] proposed a novel subject-specific channel selection method based on a criterion derived from Fisher's discriminant analysis that can be applied with a small number of electrodes typically available with a commercial EEG device. Schiatti et al. [28] have used a feature optimization algorithm to select band power features to classify MI.

The primary objective of this study is twofold. Firstly, we focused on developing a robust classification algorithm WaveCSP that can be used to differentiate MI using EEG acquired by low-cost commercial devices. WaveCSP incorporates wavelet transform and CSP algorithms to extract features from the mu-beta rhythm of EEG. The objective was to increase the number of features to capture intra-band time and frequency domain discrimination between the classes. We evaluated the performance of left hand vs right hand clenching MI classification using the proposed algorithm for EEG acquired using an EPOC EEG headset [29]. Secondly, we focused on the specific challenges associated with MI classification using EEG from a commercial device including a limited number of electrodes, limited spatial distribution of electrodes, lower signal quality and subject variabilities. To independently evaluate the effect of these limitations in isolation, we evaluated the algorithm with a publicly available MI database of EEG acquired using a medical grade device from 109 subjects.

Methods

Dataset

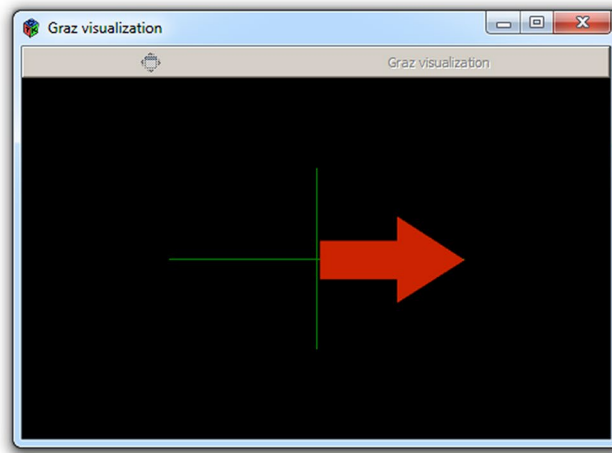
EEG data from Physionet

The publicly available EEG data were sourced from the Physionet MI dataset [30]. We used the left-hand versus right-hand MI paradigm where the subject imagines clenching his right or left fist according to a cue displayed on a screen. When the cue appears directing towards either the left or the right side of the screen, the subject imagines opening and closing the corresponding fist until the cue disappears. Then the subject relaxes a random time of T_R (< 2 s) until another cue is displayed. Figure 1—Top shows the timeline of a single trial of the experiment. The EEG recorded simultaneously consists of signals from 64 EEG channels, each sampled at 160 Hz and saved along with annotation markers for the display of right and left cues. Data were available for 109 volunteering subjects who performed a total of 45 left and right MI tasks in three experiments that took 2 min each.

Data from the commercial EEG headset

The EEG data were collected using the aforementioned EEG headset. We adopted the same experimental protocols reported under which the publicly available data were collected. We used the OpenViBE EEG platform for acquiring EEG, generating stimulations for cues, displaying cues and for recording signals and annotation markers as GDF files. The target cue for left and right MI were chosen to be arrows pointing to the corresponding direction (Fig. 1

Fig. 1 Top-The timeline of a single trial of MI experiment. **Bottom**-The cue displayed to instruct the subject to imagine right-hand clenching MI



-Bottom). The EEG signals consisted of 14 EEG channels, each sampled at 128 Hz. Each of the 25 volunteering subjects performed a total of 46 left and right MI tasks. Five out of the 25 subjects had previously participated in at least one BCI experiment. The rest had no previous experience with any type of BCI.

Analysis

The data collected from the two sources were used for five studies as described in Table 1. The experimental procedure and the parameters used for training and testing the classifiers were kept constant across all the studies. Based on these studies, five different analyses were undertaken to comparatively evaluate the performance of the proposed algorithm on a medical grade and consumer grade EEG with inherent limitations as shown in Table 2. Stages of the WaveCSP algorithm are shown in Fig. 2.

Algorithm

Pre-processing

The EEG acquired from the Physionet database [31] were clean and noise-free due to the superior quality of the acquisition system. Therefore, no filtering was applied to these signals. However, the signals recorded using the headset were noisy and therefore a filter was used to smoothen the signal as shown in Eq. 1. A moving average filter of width $w = 10$ was used to optimize classification accuracy.

$$y[n] = \frac{1}{w} \sum_{i=0}^{w-1} x[n - i] \tag{1}$$

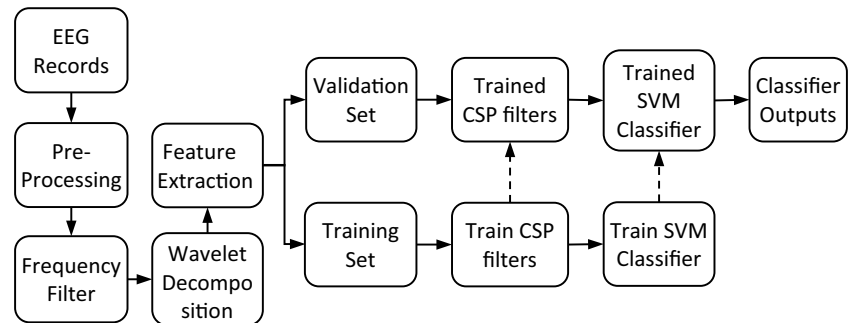
The signals from both the data sets were epoched into blocks of time series signals using the event markers for each left-hand and right-hand MI trials using the EEGLAB toolbox [32] available in MATLAB.

Table 1 Description of the studies performed

Study	Description
1	Physionet data with all 64 electrodes
2	Physionet data with only the 14 electrodes available in the commercial EEG headset
3a	Headset data with all 14 electrodes in the regular position (all 25 subjects)
3b	Headset data with all 14 electrodes in the regular position (only the five subjects with prior BCI experience)
4	Headset data with all 14 electrodes and headset pulled backwards (only the five subjects with prior BCI experience)

Table 2 Description of the analysis of the studies performed

Analysis	Studies analyzed	Purpose
1	Study 3a, study 3b	Studying the effect of BCI literacy on the performance of the algorithm
2	Study 1, study 2	Studying the effect of the limited number of electrodes independent of other factors
3	Study 3b, study 4	Studying the effect of the spatial arrangement of electrodes over the scalp
4	Study 2, study 3a	Studying the effect of signal quality independent of other factors
5	Study 1, study 3a	Studying the overall performance difference between a commercial EEG device and a professional EEG device

Fig. 2 The schematic representation of the algorithm

Frequency filter

Mu and beta rhythms, defined to be the EEG signal bandwidth confined in the ranges of 7.5–12.5 Hz and 16–31 Hz respectively have been well documented as corresponding to the motor activity of the brain. Therefore, we extracted the signal components $y[n]$ belonging to the frequency range 7.5–31 Hz from the raw EEG signals $x[n]$ using a Hamming windowed sinc filter $f[n]$ as shown in Eq. (2) where the $*$ indicates convolution operation. The filter order was optimally chosen by EEGLAB implementation of the filter.

$$y[n] = f[n] * x[n] \quad (2)$$

Wavelet decomposition

The combined mu-beta component of each signal was decomposed into three wavelet components using discrete wavelet transform (DWT) using Harr mother wavelet. The objective of applying wavelet transform was to capture the intra-band discriminators for the two classes that span both in time and frequency domains. Each of the 64 (or 14) channels were decomposed into three components such that there were 192 (or 42) signals available for the next stage.

CSP filter training

Each of the three sets of components (with 64 or 14 channels each) resulting from wavelet decomposition were used

to train a CSP filter. CSP is a spatial filtering technique that maximizes the ratio of variance between two sets of time series signals.

Consider X_L and X_R as the two sets of signals belonging to the two classes (left hand vs right hand) of MI. Each of X_L and X_R consist of $n=64$ or 14 signals of length T samples each in the form of a matrix of dimension $n \times T$. Here T was determined by the sampling rate of the recorded EEG and the time interval for which the cue arrow was displayed. The goal of CSP filter training is to obtain a spatial filter W such that the ratio of variance between the resulting signals after multiplying the original signal with the filter is maximized.

$$W = \arg \max_w \frac{\|WX_L\|^2}{\|WX_R\|^2} \quad (3)$$

The solution for this optimization problem can be obtained by simultaneous diagonalization of the averaged normalized spatial covariance of the two matrices X_L and X_R . The normalized spatial covariance of the two sets of signals R_L and R_R can be calculated as follows:

$$R_L = \frac{X_L X_L^T}{\text{trace}(X_L X_L^T)} \quad (4)$$

$$R_R = \frac{X_R X_R^T}{\text{trace}(X_R X_R^T)}$$

Average of R_L and R_R is computed over the 23 (or 22) trials of left and right imagery for each subject to produce the

average normalized spatial covariance matrices $\overline{R_L}$ and $\overline{R_R}$. Simultaneous diagonalization of $\overline{R_L}$ and $\overline{R_R}$ is equivalent to Eigen decomposition of $\overline{R_L}^{-1} \overline{R_R}$:

$$\overline{R_L}^{-1} \overline{R_R} = PDP^{-1} \tag{5}$$

Here D is the diagonal matrix of eigenvalues and columns of P represent the corresponding eigenvector. When the columns of P are sorted in descending order of the corresponding eigenvalue and transposed we obtain n spatial filters in the n rows of the resulting matrix V :

$$V = \text{sort}(P)^T \tag{6}$$

We define W of dimension $2m \times n$ as the spatial filter matrix by selecting the first m and last m ($< n$) rows of V . Using W we obtain $2m$ spatially filtered signals that maximize the variance between the two classes represented by each row of F_L and F_R where:

$$\begin{aligned} F_L &= WX_L \\ F_R &= WX_R \end{aligned} \tag{7}$$

Feature extraction

Three spatial filters W_1 , W_2 , and W_3 were obtained for each of the three wavelet decomposition components from the previous step. We selected $m = 2$ after evaluating the performance of the algorithm for $m = 1, 2$ and 3 . The resulting filtered signals F_1 , F_2 , and F_3 contained four filtered signals each. Two sets of features f_A and f_B were extracted from each of the signals to produce a total of 24 features for each EEG epoch (Table 3). Here $F_{i,j}$ denotes the j^{th} row of F_i :

$$\begin{aligned} f_A(i,j) &= \log \left(\frac{\text{var}(F_{i,j})}{\sum_{j=1}^4 \text{var}(F_{i,j})} \right) \\ f_B(i,j) &= \sum_{k=1}^l F_{i,j}[k] \text{ where } l = \text{length of } F_{i,j} \\ \text{for } i &= 1, 2, 3 \text{ and } j = 1, 2, 3, 4 \end{aligned} \tag{8}$$

Table 3 Feature vector used for machine learning

	Wavelet component (i)			
	1	2	3	
CSP Filter (j)	1	$f_A(1, 1)$	$f_A(2, 1)$	Variance feature (A)
	2	$f_A(1, 2)$	$f_A(2, 2)$	
	3	$f_A(1, 3)$	$f_A(2, 3)$	
	4	$f_A(1, 4)$	$f_A(2, 4)$	
	1	$f_B(1, 1)$	$f_B(2, 1)$	Power feature (B)
	2	$f_B(1, 2)$	$f_B(2, 2)$	
	3	$f_B(1, 3)$	$f_B(2, 3)$	
	4	$f_B(1, 4)$	$f_B(2, 4)$	

Classifier

Three commonly used machine learning algorithms, a linear SVM classifier, an LDA classifier and a KNN classifier were tested as the machine learning framework. The three algorithms have been widely used in the binary classification of MI. A linear SVM classifier separates two classes of training data points by maximizing the gap between the decision boundaries. New predictions are then made by assigning a test data point to a category based on side of the boundary that it belongs. LDA assigns a class to a test data point based on whether it satisfies a criterion derived as a function of a linear combination of the training data points. In KNN a test data point is assigned to the class most common among its K nearest neighbors where Euclidean distance is considered in the multidimensional feature space.

The model parameters for the classifiers were heuristically chosen to optimize the performance using preliminary experiments. Based on the results of the preliminary experiments, we selected a Matlab implementation of linear SVM which automatically selects a suitable scale factor by subsampling the input data. The soft margin parameter was set to 1. LDA was implemented in Matlab using uniform prior probabilities and same diagonal covariance matrix for both the classes. KNN was implemented with $K = 1$ neighbors and uniformly weighted Euclidean distance. All three algorithms standardize the predictor variables by the corresponding mean and standard deviation.

Performance evaluation

For each of the subjects, 45 trials of either left or right MI were available. Fivefold cross-validation was performed by dividing the 45 trials into five groups of 9 trials each and using a combination of four of them as the training set and the remaining set of records as the testing set in each of the fivefolds of validation. The classifiers were trained using the feature vector of 24 features obtained from each trial.

The classifier outputs were compared with the class labels of the testing trials and accuracy of prediction was evaluated as the percentage of correctly classified trials out of the total number of trials conducted for the subject. In contrast to other physiological signals such as ECG, EEG signals depend on the subject’s thought process. Termed BCI literacy, the degree to which one can successfully communicate through BCI differs from person to person [33]. Particularly, BCI illiteracy in MI tasks is known as “Motor Imagery Inability” that signifies the inability of some people to elicit imaginary movements [34]. These subjects are unable to perform imaginary motor activity even if they perfectly understood the instructions and fully intend to engage in the experiment. Therefore, the mean accuracy of all the subjects may not accurately describe the true performance of the classification algorithm. To address this issue, we adopted a measure

proposed by Muller-Putz et al. [35] to identify subjects for whom the algorithm performed significantly well as opposed to “random guessing”. The random process of classifying a signal into either left or right MI can be modeled as a binomial process with expected value 50%. The upper bound of the 95% confidence interval of this expected value indicates a percentage above which a classifier can be considered to have performed significantly better than randomly guessing. According to the method proposed by Muller-Putz et al. [35], for 45 trials the upper bound of the confidence interval is 64%. We developed the metrics described in Table 4 to evaluate the performance of the algorithm.

We compared the performance of the proposed algorithm against previously reported algorithms for data obtained using a medical grade EEG device (in study 1) and for data obtained using a commercial EEG device (in study 3). In study 1, we compared the performance of the proposed algorithm against two recent works based on SUTCCSP algorithm [36, 37] that are evaluated on the Physionet MI dataset which is acquired using a medical grade EEG device. In study 3a, we compared the performance of the proposed algorithm against two previously reported works where the same commercial headset was used to classify left vs right MI using CSP features [21] and band power features [28] respectively. We limited the comparison for only the studies that used the same device (commercial or medical grade) and the same MI classification (left vs. right hand) because of the performance variability across devices and EEG paradigms.

Results

The study 1 for the three classifiers was done using data from all the 109 subjects of the Physionet dataset. All three classifiers performed nearly equal, with KNN slightly leading the others in terms of SPP (Table 5). To evaluate the performance of the algorithm independent of the limitations imposed by the headset, we compared the results of our algorithm with two recently published work that used the same medical grade EEG dataset. Both the algorithms use EEG features extracted with strong-uncorrelating Transform Complex CSP (SUTCCSP) which is an extension of CSP

method for the analysis of pair-wise MI data. Kim et al. [37] show a superior SA but a low SPP resulting in an overall loss in mean accuracy. Park et al. [36] have reported slightly lower SA and slightly higher SPP than the proposed method. However, they have not included data from four subjects citing poor signal quality. Moreover, they have neither reported the overall mean accuracy or the mean accuracy of subjects who performed below the significant accuracy.

In study 2, when only the 14 channels of EEG signals acquired from the medical grade data using electrodes available in the commercial headset were used, a clear reduction of performance was observed. Interestingly, unlike in study 1, SVM outperformed the other two algorithms.

In study 3a, data collected from 25 subjects using the headset was used to evaluate the performance of the algorithm, all three classifiers performed poorly. SVM showed the best SA with significant accuracy level for 4 subjects out of 25. In comparison, Fakhruzzaman et al. [21] used online processing with an OpenViBE implementation of CSP algorithm. Following CSP filtering, the signal is filtered in the alpha/beta [8–30] Hz range, split into blocks of 1 s every 16th second and the logarithmic band power is computed as features for an LDA classifier. The performance was evaluated on a single subject. Schiatti et al. [28] used band power features extracted from all 14 electrodes available in the commercial headset. These features included five frequency bands [$\mu = (8–13)$ Hz, $B_1 = (13–18)$ Hz, $B_2 = (18–23)$ Hz, $B_3 = (23–28)$ Hz, and $B = (13–30)$ Hz] extracted from three time windows [TW1 = (0–2) s, TW2 = (1–3) s, and TW3 = (2–4) s] after the stimulus presentation. Mutual information-based feature selection was used to select optimal features and subsequently, an SVM classifier was trained. While their results are slightly better than the proposed algorithm, they evaluated the performance only on three subjects.

In study 3b, headset data with all 14 electrodes in regular position from only the five subjects with prior BCI experience were used. A significant improvement in all three performance metrics was observed. The SVM classifier showed significant performance for four out of the five subjects who had prior BCI experience while LDA and KNN both showed significant performance for three out of the five subjects.

Study 4 was performed to investigate the effect of electrode distribution of the commercial headset. As shown in Fig. 3, the device does not have electrodes over the motor cortex in its usual position. Therefore, this study was performed with the headset positioned backwards from its usual position such that the frontal sensors align with the motor and premotor cortices. We used the data from the subject who had prior BCI experience in this study because of the very poor performance of the other subjects. Better signal acquisition from the motor cortex was expected and hence better performance. However, the mean accuracy and SPP have reduced considerably while SA has considerably increased.

Table 4 Definition of performance metrics

Metric	Definition
Mean accuracy (MA)	MA of all the subjects evaluated
Significant performance percentage (SPP)	Percentage of subjects for whom the algorithm performed with accuracy higher than 64%
Significant accuracy (SA)	MA for the subjects for whom the algorithm performed with accuracy higher than 64%

Fig. 3 The recommended positioning of electrodes in the EPOC headset and the region of motor cortex mapped over the scalp

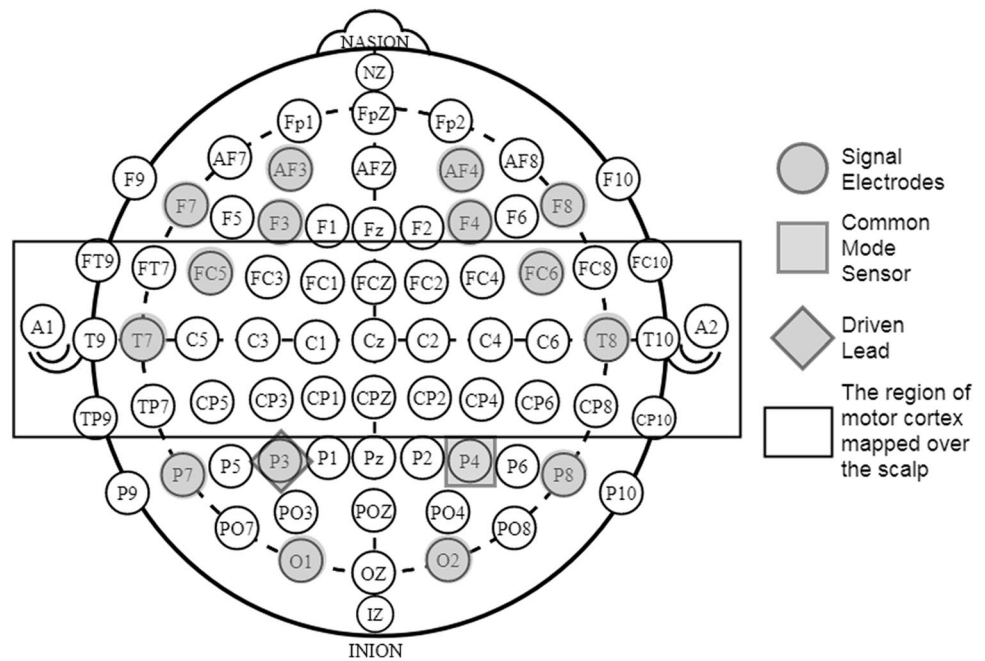


Table 5 Performance evaluation in the studies conducted

Study	Classifier	MA (%)	SA (%)	SPP (%)
1	WaveCSP–LDA	63.4	74.0	46.8
	WaveCSP–SVM	63.4	74.8	46.8
	WaveCSP–KNN	63.5	74.0	50.5
	SUTCCSP with random forest [37]	57.9 ^a	80.1	22.9 ^a
	SUTCCSP with SVM Gaussian Kernel [36]	NA	72.4 ^b	51.4 ^b
2	WaveCSP–LDA	60.3	72.9	35.8
	WaveCSP–SVM	61.3	72.6	43.1
	WaveCSP–KNN	59.2	53.9	30.3
3a	WaveCSP–LDA	55.0	68.1	16
	WaveCSP–SVM	54.1	68.6	16
	WaveCSP–KNN	52.8	67.8	16
	Fakhruzzaman et al. [21]	28.3	NA	NA
3b	Schiatti et al. [28]	62.0 ^c	62.0 ^{c,d}	100.0 ^{c,d}
	WaveCSP–LDA	64.1	68.7	60
4	WaveCSP–SVM	67.0	68.6	80
	WaveCSP–KNN	61.2	68.8	60
	WaveCSP–LDA	61.1	78.8	40
	WaveCSP–SVM	61.6	89.0	20
	WaveCSP–KNN	61.8	72.8	40

^aCalculated using reported values

^bFour subjects were not evaluated citing poor signal quality

^cCalculated from the accuracy reported for three subjects

^dSignificant accuracy margin of 57% is used as 200 trials were used

Discussion

The performance of an EEG classifier heavily depends on the degree to which the subject was able to accurately perform the mental activity studied, in addition to the inherent

differences between subjects in the electrical activity of the brain. Therefore, having a high SPP is important for a classifier to be implemented in consumer level BCI. It has been widely accepted that MI BCI paradigm requires more training for successful use. Through analysis 1, the results from

studies 3a and 3b clearly show that subjects who had prior experience with BCI (four had performed similar MI experiments, one had experienced an SSVEP experiment) show better performance. With four out of five subjects having significant performance, WaveCSP-SVM shows the best performance for subjects who have some prior BCI experience.

We performed analysis 2 by comparing the results of study 1 and study 2. A clear reduction in performance has been observed in all the performance metrics. Therefore, it can be concluded that independent of the signal quality, the limited number of electrodes available in the headset reduces the performance of MI classification. Since increasing the number of electrodes is not a feasible option for an easy-to-wear headset, increasing the number of classifier features in a non-redundant fashion is demonstrated in our study.

When the EEG headset is positioned such that the frontal sensors align with the mid-region of the scalp, motor cortex activity should be more pronounced in the recorded EEG resulting in better classification accuracy. In analysis 3 we compared the performance of the classifier with data obtained with the headset in the usual position vs the modified position. However, due to the physical design of the headset, the reference electrodes lose contact from the scalp when altered from the ordinary position. This effect varies according to the shape and size of the subject's head as the device has been designed as a one-size-fits-all headset. In study 4, WaveCSP-SVM performance with 89% SA and 20% SPP implies that there was one subject out of the five for whom the classifier worked with very high accuracy. This may be a result of the fact that the modified position of the headset fits well on the subject's scalp.

EEG acquired from the commercial headset is more susceptible to noise and artifacts since the device is typically worn over hair and is not tightly fixed. In analysis 4, comparing study 2 and study 3a, we see a general decrease in all performance metrics when the commercial EEG headset is used instead of the medical grade EEG device. There is a significant drop in SPP when the commercial headset is used, which may be attributed to its lower signal quality. This provides evidence that signal quality of the commercial device has an effect on classification accuracy. It can be observed that, when the medical grade device with better signal quality is used, subjects for whom performance is higher than 64% the classification accuracy is considerably high and for those lower than 64%, performance is considerably low.

An overall drop in performance is observed when the commercial headset is compared with the medical grade device in analysis 5. In comparison with previous work on the commercial device, WaveCSP-SVM has better performance with respect to mean accuracy and significant accuracy. Meaningful interpretation of SA and SPP requires a large set of test subjects.

MI classification problem is solved in the twofold process of feature extraction and classification. There are many classification algorithms that have been shown to work well with EEG signals. However, the major bottleneck is in feature extraction methods. While the use of deep learning instead of handcrafted features could be a solution to this problem, its suitability for MI classification should be further investigated. Existing work using neural networks and convolutional neural networks do not produce significantly better results than the classical methods. Sparse nature of MI data could be a reason for this observation.

Conclusion

This study proposes a robust MI classifier that attempts to address the limitations posed by a commercial EEG acquisition device. Comparative analysis of the performance of the proposed algorithm for different configurations of data acquired from a commercial device and data acquired from a medical grade device is undertaken. Subjects who had prior experience in BCI performed significantly better compared to those who did not have any BCI experience. The use of a smaller number of electrodes has reduced the performance irrespective of the signal quality and other factors. Conclusive evidence was not found for the hypothesis that positioning the EEG electrodes closer to the motor cortex results in better classification performance. Practical limitations for such modified electrode placement were identified. The quality of the acquired EEG signal was found to affect the performance of MI classification. Further, EEG from medical grade device better discriminates between subjects with significant and poor performance. Variability of subjects and their prior experience and training in using BCI was identified as a significant aspect of MI classification in this work. We collected data from 25 subjects for this study due to the lack of a large EEG dataset acquired with the headset. Future work should focus on collecting a data set with more subjects using a commercial EEG device. The relationship between the levels of training and experience in BCI and MI performance could be explored in detail.

Acknowledgements This work was supported by the Singapore Academic Research Fund under Grant R-397-000-227-112. (corresponding author: Hongliang Ren). The authors would like to acknowledge Cai Xin Chen, National University of Singapore for his assistance in organizing and conducting experiment trials.

Compliance with ethical standards

Conflict of interest The author declares that they have no conflict of interest to declare.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the

National University of Singapore and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study. No identifiable personal information is disclosed.

References

- Liao LD, Chen CY, Wang Ij, Chen SF, Li SY, Chen BW, Chang JY, Lin CT, (2012) Gaming control using a wearable and wireless EEG-based brain-computer interface device with novel dry foam-based sensors. *J NeuroEng Rehabil*. <https://doi.org/10.1186/1743-0003-9-5>
- Khong A, Jiangnan L, Thomas KP, Vinod AP, (2014) BCI based multi-player 3-D game control using EEG for enhancing attention and memory. In: *Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference on*, San Diego, CA, USA,
- Ahn M, Lee M, Choi J, Jun SC, (2014) A review of brain-computer interface games and an opinion survey from researchers, developers and users. *Sensors* 14(8):14601–14633
- Grandchamp R, Delorme A, (2016) The brainarium: an interactive immersive tool for brain education, art, and neurotherapy. *Comput Intell Neurosci* 8:100–111
- Daly JJ, Huggins JE, (2015) Brain-computer interface: current and emerging rehabilitation applications. *Arch Phys Med Rehabil* 96(3):S1–S7,
- Dokkum Lv, Ward T, Laffont I (2015) Brain computer interfaces for neurorehabilitation—its current status as a rehabilitation strategy post-stroke. *Ann Phys Rehabil Med* 58(1):3–8
- McFarland DJ, Daly J, Boulay C, Parvaz MA (2017) Therapeutic applications of BCI technologies. *Brain-Comput Interfaces* 4(1–2):37–52
- Guger C, Daban S, Sellers E, Holzner C, Krausz G, Caraballona R, Gramatica F, Edlinger G (2009) How many people are able to control a P300-based brain-computer interface (BCI)?. *Neurosci Lett* 462(1):94–98
- Annese VF, Mezzina G, Venuto DD, (2016) Towards mobile health care: neurocognitive impairment monitoring by BCI-based game. In: *IEEE SENSORS*, Orlando, FL, 2016
- Dreyer A, Herrmann C, Rieger J, (2017) Tradeoff between user experience and BCI classification accuracy with frequency modulated steady-state visual evoked potentials. *Front Hum Neurosci* 11:391
- Ekandem JI, Davis TA, Alvarez I, James MT, Gilbert JE, Evaluating the ergonomics of BCI devices for research and experimentation. *Ergonomics* 55(5):592–598, 2012
- Mao X, Li M, Li W, Niu L, Xian B, Zeng M, Chen G, (2017) Progress in EEG-based brain robot interaction systems. *Comput Intell Neurosci*. <https://doi.org/10.1155/2017/1742862>
- Zhao J, Li W, Li M, (2015) Comparative study of SSVEP- and P300-based models for the telepresence control of humanoid robots. *PLoS ONE* 10(11):e0142168
- R SS, Rabha J, Nagarjuna KY, Samanta D, Mitra P, Sarma M, Motor imagery EEG Signal processing and classification using machine learning approach. In: *International Conference on New Trends in Computing Sciences (ICTCS) 2017*, 2017
- Krishna DH, Pasha I, Savithri TS (2016) Classification of EEG motor imagery multi class signals based on cross correlation. *Procedia Comput Sci* 85:490–495
- Ramos AC, Hernandez RG, Vellasco M, Vellasco P, Ensemble of classifiers applied to motor imagery task classification for BCI applications. In: *International Joint Conference on Neural Networks (IJCNN)*, 2017
- Hamed M, Salleh S-H, Noor AM, Mohammad-Rezazadeh I, Neural network-based three-class motor imagery classification using time-domain features for BCI applications. In: *IEEE REGION 10 SYMPOSIUM*, 2014
- Tang Z, Li C, Sun S, 2016 Single-trial EEG classification of motor imagery using deep convolutional neural networks. *Opt—Int J Light Electron Opt*. <https://doi.org/10.1016/j.ijleo.2016.10.117>
- Sakhavi S, Guan C, Yan S, Parallel convolutional-linear neural network for motor imagery classification. In: *23rd European Signal Processing Conference (EUSIPCO)*, 2015
- Vamvakousis Z, Ramirez R, (2013) Towards a low cost MuRhythm based BCI. *Proc Fifth Int Brain-Comput Interface Meet 2013*. <https://doi.org/10.3217/978-3-85125-260-6-13>
- Fakhruzzaman MN, Riksakomara E, Suryotrisongko H (2015) EEG wave identification in human brain with emotiv EPOC for motor Imagery. *Procedia Comput Sci* 72:269–276
- Renard Y, Lotte F, Gibert G, Congedo M, Maby E, Delannoy V, Bertrand O, Lecuyer A, (2010) OpenViBE: an open-source software platform to design, test and use brain-computer interfaces in real and virtual environments. *Presence: Teleoper Virtual Environ*. <https://doi.org/10.1162/pres.19.1.35>
- Takehara D, Kayanuma H, Matsubara M, Seki N, Wada K, Kurata M, Ono Y, (2017) Development of an ERD measurement system using emotiv EPOC. *Trans Jpn Soc Med Biol Eng*. 55:555–559
- Martinez-Leon J-A, Cano-Izquierdo J-M, Ibarrola J, (2016) Are low cost brain computer interface headsets ready for motor. *Expert Syst Appl* 49:136–144,
- Cano-Izquierdo J, Almonacid M, Pinzolas M, Ibarrola J (2009) dFasArt: dynamic neural processing in FasArt model. *Neural Networks* 22(4):479–487
- Hurtado-Rincon J, Rojas-Jaramillo S, Ricardo-Cespedes Y, Alvarez-Meza AM, Castellanos-Domínguez G, Motor imagery classification using feature relevance analysis: an emotiv-based BCI system. In: *Image, Signal Processing and Artificial Vision (STSIVA), 2014 XIX Symposium on*, Armenia, Colombia, 2014
- Yang Y, Kyrgyzov O, Wiart J, Bloch I, Subject-specific channel selection for classification of motor imagery electroencephalographic data. In: *International Conference on Acoustics, Speech, and Signal Processing*, Vancouver, 2013
- Schiatti L, Faes L, Tessadori J, Barresi G, Mattos L, Mutual information-based feature selection for low-cost BCIs based on motor imagery. In: *IEEE EMBS*, Orlando, 2016
- Emotiv Inc., San Francisco, USA, Emotiv EPOC, (2017) [Online]. Available: <https://www.emotiv.com>. [Accessed 01 01 2017]
- Goldberger A, Amaral L, Glass L, Hausdorff J, Ivanov P, Mark R, Mietus J, Moody G, Peng C, Stanley H, (2000) Physiobank, physio-toolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation* 101(23):215–220
- Schalk G, McFarland DJ, Hinterberger T, Birbaumer N, Wolpaw JR, (2000) A general-purpose brain-computer interface (BCI) system. *IEEE Trans Biomed Eng* 51(6):1034–1043
- Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J Neurosci Methods* 21:134–139
- Mcavinue LP, Robertson IH (2008) Measuring motor imagery ability: a review. *Eur J Cognit Psychol* 20(2):232–251
- Ahn M, Cho H, Ahn S, Jun SC (2013) High theta and low alpha powers may be indicative of BCI-illiteracy in motor imagery. *PLoS ONE* 8(11):e80886
- Muller-Putz GR, Schrer R, Brunner C, (2008) Better than random? a closer look on BCI results. *Bioelectromagnetism* 10(1):52–55

36. Park C, Took CC, Mandic DP (2014) Augmented complex common spatial patterns for classification of noncircular EEG from motor imagery tasks. *IEEE Trans Neural Syst Rehabil Eng* 22(1):1–10
37. Kim Y, Ryu J, Keun KK, Took CC, Mandic DP, Park C, (2016) Motor imagery classification using Mu and beta rhythms of EEG with strong uncorrelating transform based complex

common spatial patterns. *Comput Intell Neurosci*. <https://doi.org/10.1155/2016/1489692>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.