

RESEARCH ARTICLE

# Comparing Features for Classification of MEG Responses to Motor Imagery

Hanna-Leena Halme<sup>1,2\*</sup>, Lauri Parkkonen<sup>1,3</sup>

**1** Department of Neuroscience and Biomedical Engineering (NBE), Aalto University School of Science, Espoo, Finland, **2** Radiology Unit, HUS Medical Imaging Center, Helsinki University Hospital, Helsinki, Finland, **3** Aalto Neuroimaging, MEG Core, Aalto University School of Science, Espoo, Finland

\* [hanna-leena.halme@aalto.fi](mailto:hanna-leena.halme@aalto.fi)



## Abstract

### Background

Motor imagery (MI) with real-time neurofeedback could be a viable approach, e.g., in rehabilitation of cerebral stroke. Magnetoencephalography (MEG) noninvasively measures electric brain activity at high temporal resolution and is well-suited for recording oscillatory brain signals. MI is known to modulate 10- and 20-Hz oscillations in the somatomotor system. In order to provide accurate feedback to the subject, the most relevant MI-related features should be extracted from MEG data. In this study, we evaluated several MEG signal features for discriminating between left- and right-hand MI and between MI and rest.

### Methods

MEG was measured from nine healthy participants imagining either left- or right-hand finger tapping according to visual cues. Data preprocessing, feature extraction and classification were performed offline. The evaluated MI-related features were power spectral density (PSD), Morlet wavelets, short-time Fourier transform (STFT), common spatial patterns (CSP), filter-bank common spatial patterns (FBCSP), spatio—spectral decomposition (SSD), and combined SSD+CSP, CSP+PSD, CSP+Morlet, and CSP+STFT. We also compared four classifiers applied to single trials using 5-fold cross-validation for evaluating the classification accuracy and its possible dependence on the classification algorithm. In addition, we estimated the inter-session left-vs-right accuracy for each subject.

### Results

The SSD+CSP combination yielded the best accuracy in both left-vs-right (mean 73.7%) and MI-vs-rest (mean 81.3%) classification. CSP+Morlet yielded the best mean accuracy in inter-session left-vs-right classification (mean 69.1%). There were large inter-subject differences in classification accuracy, and the level of the 20-Hz suppression correlated significantly with the subjective MI-vs-rest accuracy. Selection of the classification algorithm had only a minor effect on the results.

## OPEN ACCESS

**Citation:** Halme H-L, Parkkonen L (2016) Comparing Features for Classification of MEG Responses to Motor Imagery. PLoS ONE 11(12): e0168766. doi:10.1371/journal.pone.0168766

**Editor:** Bin He, University of Minnesota, UNITED STATES

**Received:** July 14, 2016

**Accepted:** December 5, 2016

**Published:** December 16, 2016

**Copyright:** © 2016 Halme, Parkkonen. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Local legislation on medical research, enforced by The Finnish Advisory Board on Research Integrity, prevents open sharing of the data. However, with an explicit prior agreement, the data can be shared with an external party. Data will be available only upon request from the following persons: Hanna-Leena Halme, MSc., [hanna-leena.halme@aalto.fi](mailto:hanna-leena.halme@aalto.fi); Lauri Parkkonen, Prof., [lauri.parkkonen@aalto.fi](mailto:lauri.parkkonen@aalto.fi).

**Funding:** The study was supported by Emil Aaltonen Foundation (<http://emilaaaltonen.fi/>) and Academy of Finland grant #295075 “NeuroFeed” (<http://www.aka.fi/>).

**Competing Interests:** The authors have declared that no competing interests exist.

## Conclusions

We obtained good accuracy in sensor-level decoding of MI from single-trial MEG data. Feature extraction methods utilizing both the spatial and spectral profile of MI-related signals provided the best classification results, suggesting good performance of these methods in an online MEG neurofeedback system.

## 1. Introduction

Motor imagery (MI) augmented with real-time feedback of the modulated brain activity could be an effective method for rehabilitation of motor function in patients who are unable to perform overt limb movements [1–3]. MI, as well as overt motor acts and somatosensory stimulation, is associated with a suppression of 10- and 20-Hz oscillations over the sensorimotor cortex and a rebound of these after the end of MI, motor activity or stimulation [4]. These oscillations are often referred to as the mu rhythm, or more generally as sensorimotor rhythms (SMR). Movement-related suppression and rebound of this rhythm are often referred to as event-related desynchronization (ERD) and event-related synchronization (ERS), respectively. ERD and ERS during overt hand movements were detected in magnetoencephalography (MEG) by Salmelin and Hari [5]. Later, such dynamics has been observed in MEG and electroencephalography (EEG) also during MI [6,7], attempted movements [8], and movement observation [9]. Because the neural activation patterns during MI and executed movements are highly similar [7,10,11], MI-based neurofeedback training could serve as an alternative or ancillary to physical therapy. Regular MI training could facilitate neural plasticity and improve motor skills in e.g. patients suffering from stroke.

The majority of studies on noninvasive MI-based brain–computer interfaces (BCI) have been conducted with EEG. MEG has also been utilized due to its better source localization accuracy compared to EEG [12] and it has already demonstrated its usefulness in neurofeedback [13–15]. Advances in machine-learning and signal-processing algorithms have enabled BCIs which detect users' intentions by classifying brain activity in real time and by giving feedback according to the classification results [16]. In particular, BCIs based on classification of limb MI have been studied and developed, since they have many potential applications in both neurobehavioral experiments and clinical interventions.

In recent years, there has been rapid development of noninvasive MI-based BCI systems (see a review by He a colleagues [17]). SMR are used as a control signal in many rehabilitative and assistive devices, as they offer the possibility for continuous control of direction, velocity or acceleration of the external device. Other types of control signals, e.g. steady-state visual evoked potentials (SSVEP) or P300, do not readily allow continuous control, and are therefore less popular in movement-related BCI systems. Control of two-dimensional [18] and even three-dimensional [19–21] movement has been achieved with EEG-based BCIs utilizing SMR. However, current classification algorithms could be still improved in terms of accuracy, generalizability and learning time; especially the latter is critical in patient studies as the time needed for calibrating the classifier should be as short as possible. As many different methods have been developed for decoding brain activity, it is important to compare their performance to find out the most effective one. Comparison and further development of the existing decoding algorithms is the main motivation for our research.

Extracting the relevant brain activity from MEG or EEG is the main challenge in the classification of neurophysiological measurements. The number of extracted signal features (e.g.

amplitude values, frequency content, spatial patterns) quickly becomes much higher than the available number of single-trial samples corresponding to the task of interest. As most classification algorithms require only a few relevant features for reliable discrimination between classes, it is necessary to reduce the dimensionality of the feature space without losing relevant discriminative information. Dimensionality reduction can be performed with spatial filtering methods, such as Common Spatial Patterns (CSP), which have yielded promising results in discriminating MI patterns in EEG [22–26]. CSP filters project the original signals to components having maximum differences in covariance between two classes. Several extensions, such as Filter Bank CSP (FBCSP, [27]) and Common Spatio—Spectral Pattern (CSSP, [28]), have been suggested for improving the robustness of CSP. Spatio—Spectral Decomposition [29] is a novel spatial filtering method, which maximizes the signal power at a frequency of interest while simultaneously minimizing it at the neighboring frequencies. The method leads to the optimization of signal-to-noise ratio at the selected frequency band and extraction of signal components with a desired spectral profile.

Features extracted by time—frequency analysis are efficient for characterizing changes in oscillatory brain activity, such as those occurring during motor execution and MI [30–32]. With a detailed time—frequency decomposition, one can identify the frequency components that modulate in response to the events of interest and extract those components to be used as discriminative features for classification. To focus on the most informative time—frequency features, both the time window and frequency range of the analysed signal should be chosen appropriately. As multiple neural processes can produce oscillations with similar frequency content, one should also know which brain region to target. In short, sufficient prior knowledge about both the experimental protocol and the neurophysiological phenomena under study is necessary for optimizing feature extraction and classification. Furthermore, in most real-time feedback applications, the total delay of the system should be minimized [33], which rules out computationally expensive signal-processing and classification methods.

Another major problem in designing an MI-BCI is to adapt the classifier to the high inter-subject variability. The most reactive frequency band, the time interval of the most prominent suppression, and the spatial location showing the largest modulation may all differ between subjects and different MI tasks. In addition, the MI-related activation even in the same subject usually varies to some degree between sessions. Despite these sources of variability, the performance of the classifier should ideally remain constant to enable repetitive neurofeedback training over days or weeks. Moreover, an ideal feature set and classifier should generalize across subjects.

In this study, we aimed at selecting optimal feature extraction methods for classifying single-trial sensor-level MEG data of left- and right-hand MI. MI-related feature extraction methods for MEG have been studied previously by Spüler and colleagues [34] who compared functional connectivity features and spectral features for discriminating motor imagery from a mental calculation task. Also, Kang and co-workers [35] compared different CSP methods in the discrimination of left- and right-hand motor imagery in a simultaneous MEG/EEG measurement. In the current study, we compared features extracted by time—frequency decomposition, spatial filtering methods and by the combination of these two. The tested time—frequency methods comprised short-time Fourier transform (STFT), continuous Morlet wavelets (TFR), and multitaper power spectral density (PSD). Spatial feature extraction methods included CSP, FBCSP and SSD, followed by the quantification of band-limited power. In addition, we examined whether using SSD for dimensionality reduction prior to CSP filtering improves the discriminative power of CSP, as suggested in a recent study [36]. Finally, STFT and TFR features were extracted from the CSP-filtered data. Each feature extraction method was evaluated in terms of subsequent classification performance. We also tested four different

classification algorithms to verify that in MI decoding feature extraction plays a more important role than the selection of a particular classifier. The factors underlying the differences in the classification accuracy were investigated by estimating the correlation between the mu modulation amplitude and average classification accuracy across subjects.

## 2. Methods

### 2.1. Subjects

Nine healthy volunteers (5 females, 4 males, age  $25.8 \pm 1.4$  yrs, all right handed by self-report) were recruited for this study. None of the subjects reported any neurological illnesses or motor deficits, and all subjects had normal or corrected-to-normal vision. The study was approved by the Aalto University Research Ethics Committee. The research was carried out in accordance with the guidelines of the declaration of Helsinki, and the subjects gave written informed consent prior to the MEG measurements. None of the subjects had previous experience in MI neurofeedback training.

### 2.2. MEG measurements

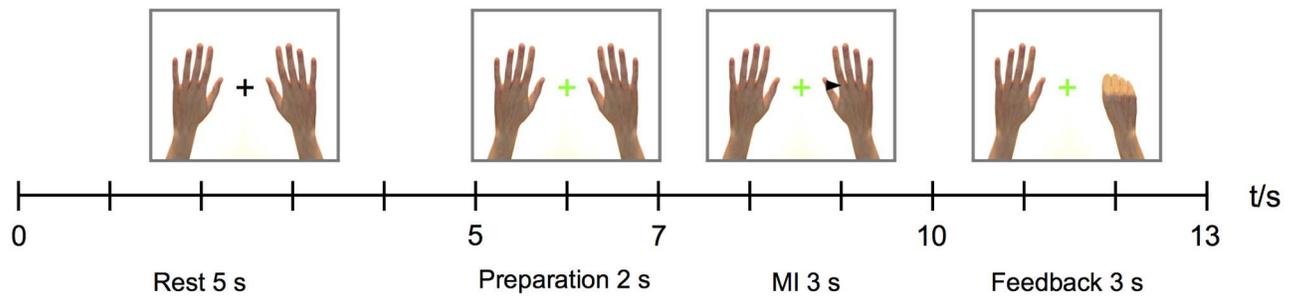
MEG was recorded with a 306-channel Elekta Neuromag™ system (Elekta Oy, Helsinki, Finland) located at the MEG Core of Aalto Neuroimaging, Aalto University. The signals were filtered to 0.1–330 Hz and digitized at the rate of 1 kHz. Four head-position indicator (HPI) coils were attached to the subject's scalp for head position estimation and alignment to a common coordinate frame. The visual cues and feedback (see below) were delivered on a screen located approximately 50 cm in front of the subject's eyes by a projector outside the shielded room. During the recording, the raw MEG data were continuously written in 300-ms segments to a network-transparent ring buffer [33,37] hosted by the MEG acquisition workstation (6-core Intel Xeon CPU at 2.4 GHz, 64-bit CentOS Linux, version 5.3). This buffer was read over a local network connection by another computer (64-bit Ubuntu Linux, version 12.04-LTS), which processed the data in real time using functions implemented in the MNE-Python software [38] and presented the visual stimuli using PsychoPy version 1.83 [39].

### 2.3. Experimental paradigm

The experiment consisted of two sessions, both of which included 40 training trials for calibration of the on-line classifier and 40 trials for testing the classification accuracy. In the beginning of each trial, the subject was presented a realistic image of two hands, pictured from above, on a white background. The subject was instructed to fixate on a small cross at the center of the screen throughout the experiment and not to perform any movements. During the MI task, the subject was instructed by a visual cue (see below for details) to imagine left- or right-hand finger tapping, i.e. touching their thumb with each finger in a sequential manner at a pace of 2–3 touches per second. It was specifically emphasized that the subject should perform kinesthetic instead of visual imagery. Each subject practiced the task for a few minutes outside the magnetically shielded room before the measurements.

Each trial began with a black fixation cross presented in the middle of the screen. After 5 s, the fixation cross turned green to indicate the preparation for the following MI task, and in 2 s a black cue arrow pointing to left or right appeared on the screen for 3 s, prompting for the MI task. Thereafter, the trial structure differed for the training and testing phases as follows:

During the training session, a grasp was shown on the same side as the preceding cue arrow.



**Fig 1. The stimulus paradigm representing the timing and display of the visual stimuli.** The 5-s rest period was followed by 2 s of preparation, 3 s of MI of either the left or the right hand, and 3 s of feedback.

doi:10.1371/journal.pone.0168766.g001

During the test trials, visual feedback of the MI classification was presented for 3 s. The feedback was graded with 20 levels between a resting open hand and a fully closed fist, based on a linear transformation of the classification probability, and it was given separately for the left and right hand. Thus, the movement presented in the feedback was different from the imagery tapping movement, whose strength cannot be visualized as a static image of a hand. Yet, we chose finger tapping as the imagery condition, because in pilot studies it elicited stronger mu-rhythm suppression than imagery of a grasping movement. However, the grasping movement was better for the visual feedback since it could be easily graded according to classification probabilities for each hand.

The timing and event codes of the stimuli were recorded to the MEG stimulus channel to ensure the correct timing of epochs in the offline data analysis. For an illustration of the experimental protocol, see Fig 1. After the training trials, there was a 30-s break, during which the subjects were shown the instructions for the following test trials. The subjects were allowed to rest eyes closed for 2–5 minutes between the two sessions.

## 2.4. Preprocessing for offline analysis

The MEG data were processed with MaxFilter software (version 2.0, Elekta Oy, Helsinki, Finland) [40], including signal space separation (SSS), coordinate frame alignment to the default head position and orientation and automatic bad-channel detection. Only the signals from the 48 gradiometers above the sensorimotor cortices were included in the data analysis, and the number and locations of the sensors were the same for all subjects; we also tested including all gradiometer channels but that yielded poorer classification results. Magnetometers were excluded due to their lower SNR, and sensors outside the parietal area were excluded in order to reduce data dimensionality and diminish the effect of visual activity, as it has been shown that visual imagery can decrease the classification accuracy in MI tasks [41].

## 2.5. Data analysis

The objective of the off-line data analysis was to improve the online classification accuracy and to optimize the analysis pipeline for further real-time measurements. Several different feature extraction algorithms were compared in terms of their ability to detect relevant features for classification.

**2.5.1. Online decoding.** For the online feedback, we applied the same, preselected feature extraction method to all subjects; power spectral densities in 8–30 Hz (bins of 1 Hz) were calculated with the multitaper method over the whole 3-s MI epoch and the selected channels. 50

most relevant features determined by chi-square statistics were retained for classification with LDA.

**2.5.2. Classifier selection.** The most efficient classifier algorithm for offline decoding was selected by comparing the classification accuracies obtained with two commonly used feature extraction methods (CSP and STFT) and four different classifiers implemented in the Scikit-learn software package version 0.17.1 [42]. The efficiency was tested with the left-vs-right classification, since it was considered the most difficult classification task. The classifiers included in this comparison were linear discriminant analysis (LDA), support vector machines with a linear (Linear SVM) and radial basis function (RBF SVM) kernel, and Naïve Bayes (NB). LDA yielded the best accuracy among these classifiers, although the difference was not significant (see Section 3.2). Therefore, LDA was selected for comparing the efficacy of the different features.

**2.5.3. Offline decoding.** The accuracy for left-vs-right MI and MI-vs-rest classification was evaluated with 5-fold cross-validation over each session; the trials were split into five equally-sized partitions, four of which served as training data to calculate feature transformations, while the fifth partition was used to test the classification performance. The procedure was repeated five times and the resulting classification accuracies were averaged.

In addition, inter-session classification over the two sessions of each subject was calculated; one session served as the training data and the other as testing data, and the classification was performed twice such that both measurements were used once for training and once for testing. The purpose of inter-session classification was to investigate how the variability of MEG signals due to e.g. physiological changes over sessions affects the classification accuracy.

When comparing the effectiveness of different features in classification, the goal is to extract only a sufficient subset of relevant features, i.e. the best possible features from each initial feature space. Therefore, in contrast to many machine-learning studies, we did not use a fixed number of features, but instead tried to find the optimal set of features from all different feature spaces derived from the same data. It is also noteworthy that the initial dimensionality of the feature space varies between different methods, and therefore also the number of relevant features is not constant.

**2.5.4. Time—frequency analysis.** Time—frequency features were estimated by multitaper PSD, STFT and Morlet wavelets (TFR). Time—frequency features were calculated for the range 8–30 Hz in 1-Hz bins. Prior to STFT, the data were filtered to this frequency range with a finite impulse response (FIR) filter (filter length equal to the epoch length, transition band 0.5 Hz). In the case of STFT and TFR, temporal resolution and thus the number of features was lowered by averaging over the time windows 0.0–1.0 s, 1.0–2.0 s, 2.0–3.0 s and 3.0–4.0 s with respect to the start of the MI epoch. Furthermore, the features were averaged over frequencies, resulting in four average time—frequency features per measurement channel.

**2.5.5. Spatial filtering methods.** CSP: These transformations were calculated on epochs filtered to 8–30 Hz. The 3 most discriminative spatial filters for each of the two classes, i.e. the 3 first and last columns of the CSP transformation matrix were selected, and the original signals were filtered with them, resulting in 6-component time courses.

FBCSP: The data were filtered to 6 frequency bands of 8–12, 12–16, 16–20, 20–24, 24–28 and 28–32 Hz with FIR filters (filter length equal to the epoch length, transition band 1.0 Hz). CSP transformations and final features were calculated similarly to the basic CSP implementation and separately for all frequency bands. 6 features per each frequency band, i.e. 36 features altogether, were retained. The data dimensionality was further reduced by selecting the 18 best features, according to chi-square statistics, for the final classification.

SSD: This decomposition was performed as described by Nikulin and colleagues [29]. First, the frequency range of maximum mu-rhythm suppression was determined for each subject

and session in sensor space using multitaper PSD. Then, the frequency band of interest was set at the frequency showing the largest suppression  $\pm 2.0$  Hz, and 4-Hz wide side bands below and above the band of interest were set as the flanking frequencies. 20 components in left-vs-right classification and 30 components in MI-vs-rest classification were selected from the covariance matrices for filtering the original data. The number of components was chosen empirically.

**SSD + CSP:** In this case, SSD was used as a preprocessing method, i.e. for extracting the mu-band oscillatory activity, for noise reduction and for linear decomposition of the data before CSP filtering. We used the same frequency of interest and side bands that were used with SSD alone (see above). The data dimensionality was likewise reduced to 20 components in left-vs-right and 30 components in MI-vs-rest classification. The SSD-transformed signal components were further transformed with CSP, and 3 CSP-filtered components per class were retained.

In all aforementioned spatial filtering methods, the filters were calculated only on the training trials, and the transformations were applied to the test trials. In each case the final feature vector was obtained by squaring the extracted component signals, averaging the samples over the selected time window and taking a logarithm of the average. In order to standardize the features and prevent classifier overfitting, all covariance matrices were regularized with Ledoit—Wolf Shrinkage (LWS) before feature extraction and classification.

**2.5.6. CSP + time—frequency analyses.** Finally, the time—frequency decompositions were calculated from the CSP-filtered data. Similarly to the previous analyses, 6 CSP components were retained, and the original signals were projected on those components. PSD, STFT and TFR were calculated on the component signals and averaged over 1-s time windows and 1-Hz frequency bins.

**2.5.7. Statistical analysis.** The statistical differences of feature extraction methods were assessed with Friedman's test, a nonparametric version of two-way analysis of variance. Since the results of the inter-session classification are independent repetitions of the same test, we used the variant of Friedman test that considers multiple observations per each feature—subject combination; in this case the number of observations was two. In the left-vs-right and MI-vs-rest classification the number of observations was one; since the folds of 5-fold cross-validation are not independent of each other, these classification tasks did not meet the criteria of multiple observations.

In addition, we performed a paired-sample t-test for the results from each fold of the 5-fold cross-validation. The null hypothesis was that the differences between classification results  $p(i) = p(a) - p(b)$  were drawn independently from a normal distribution, i.e. method *a* did not yield better results than method *b*. The alternative hypothesis was that better results were obtained with method *a* than with method *b*. The t-test was performed for each pair of methods, separately for left-vs-right and MI-vs-rest classification results.

**2.5.8. Level of mu rhythm suppression.** The suppression levels of the 10-Hz (actual frequency range 8–12 Hz) and 20-Hz (16–24 Hz) oscillations were calculated for each subject and session in order to examine whether MI-induced suppression was correlated with the subject's average classification accuracy. Average power was calculated for the baseline (−4.0 –−2.0 s before the start of the MI epoch) and MI (1.0–3.0 s from the start of the MI epoch) for the gradiometers above the sensorimotor cortex. In addition, we calculated the average baseline and MI power for CSP-filtered signals, using the same parameters for CSP as in feature extraction. The power values were averaged over epochs, gradiometers (or CSP components) and frequencies, separately for the 10- and 20-Hz rhythms, resulting in average power values during

baseline and MI periods. The suppression percentage was then calculated as [4]:

$$\text{suppression} = [(power_{MI} - power_{rest}) / power_{rest}] \cdot 100 \%$$

Pearson's correlation between 1) 10-Hz suppression and left-vs-right accuracy, 2) 10-Hz suppression and MI-vs-rest accuracy, 3) 20-Hz suppression and left-vs-right accuracy, and 4) 20-Hz suppression and MI-vs-rest accuracy were calculated. We also estimated the correlation between MI-vs-rest cross-validated accuracy and the 10/20-Hz power at rest. In addition, we computed the suppression percentage and correlations between classification accuracies and suppression individually for each gradiometer, and visualized the single CSP component showing the highest suppression for each subject.

## 2.6. Data availability

The data used in the current analyses are available from the authors upon request.

## 3. Results

### 3.1. MI-induced MEG signals

In order to illustrate the MI-induced mu-rhythm suppression and rebound, we calculated time—frequency representations (TFR) over the time course of the MI epoch and the frequency range 5–45 Hz. The TFRs were averaged over subjects, sessions and epochs, and converted to Z-scores:  $Z = (x - \mu) / \sigma$ , where  $\mu$  is the mean and  $\sigma$  the standard deviation of the TFR values. Mean and standard deviation were calculated separately for each 1 Hz frequency band in 0.5 s time windows. The resulting grand-average TFRs for left- and right-hand MI (averaged over channels) and corresponding spatial patterns (averaged over time) are shown in Fig 2. As MI is cued to begin at 0 s, there is an initial power increase at frequencies below 15 Hz, probably reflecting preparation for the following MI. The suppression (ERD) begins at around 0.5 s and changes to a rebound (ERS) after 3.0 s as MI ends. The 0.5-s delay in the beginning of MI is due to the reaction time.

### 3.2. Online left-vs-right classification accuracy

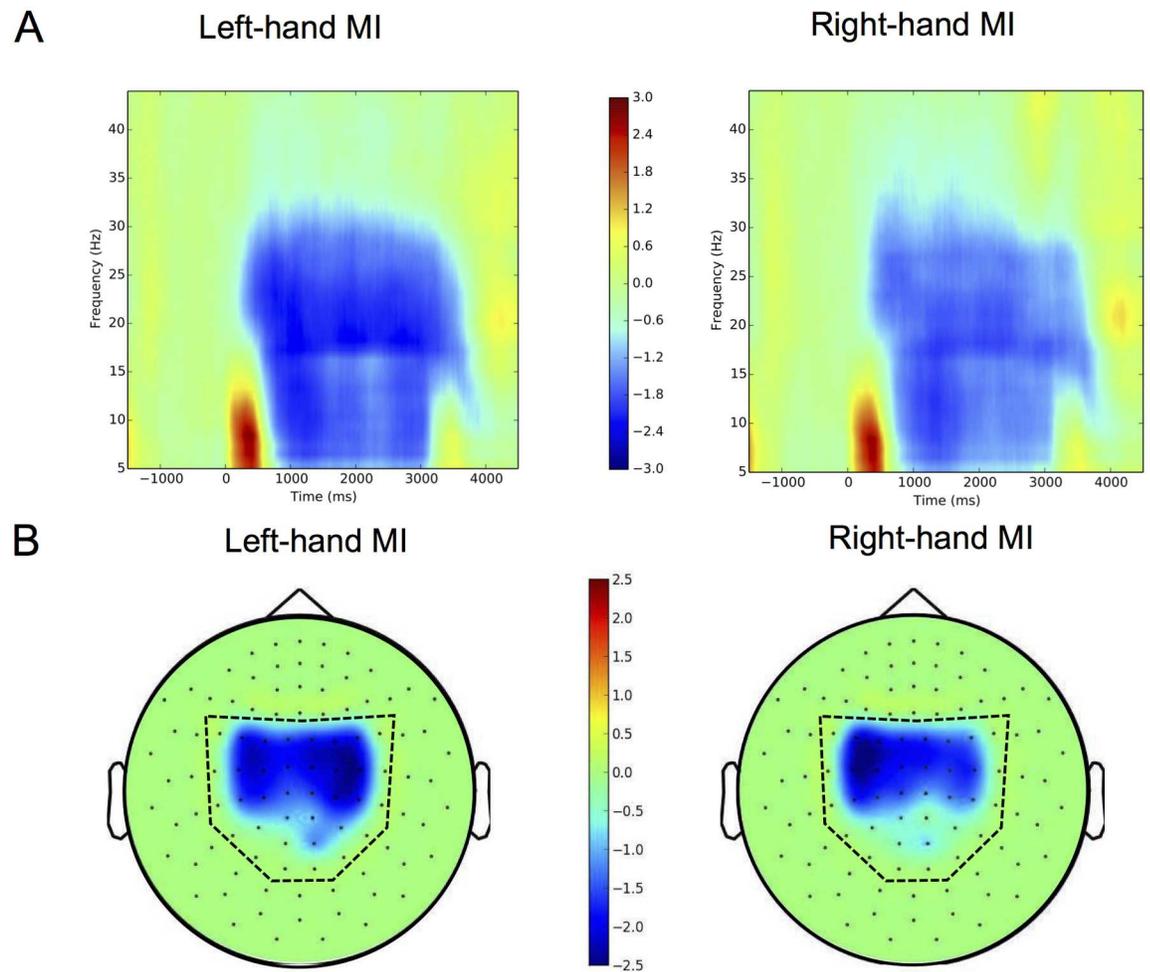
The online classification was calculated for the gradiometer data. Left-vs-right accuracy of the online feedback was fairly low; several subjects remained at the chance level (50%), and the average accuracy over subjects was 62.4%. The results for online classification are summarized in Table 1.

### 3.3. Offline classifier comparison

The accuracies obtained with the four compared classifiers are presented in Table 2. The average classification over subjects and two feature extraction methods (CSP and STFT) was calculated. LDA yielded an average accuracy of 68.60% appearing superior to the other classifiers. However, the difference was not significant.

### 3.4. Offline cross-validated classification accuracies

The average cross-validated single-trial accuracy over all methods and subjects was 68.7% (range 35.7–91.4%) in left-vs-right classification and 77.7% (range 49.3–99.4%) in MI-vs-rest classification. Applying CSP filtering before time—frequency decomposition improved the mean accuracies of all time—frequency feature extraction methods.



**Fig 2. MI-induced MEG signals averaged over subjects.** (A) Time—frequency maps for left- and right-hand MI, representing the Z-score with respect to baseline (−2.0–0.0 s from the cue onset), averaged over subjects, sessions, epochs and parietal gradiometers. (B) Corresponding topographic maps, representing the average Z-score over subjects, sessions, epochs and time window of 1.0–4.0 s from the cue onset. The set of 48 planar gradiometers included in the analysis is indicated by a dashed line.

doi:10.1371/journal.pone.0168766.g002

SSD+CSP combination yielded the best accuracies in both left-vs-right (mean 73.7%, range 59.3–90.2%) and MI-vs-rest (mean 81.3%, range 66.5–99.4%) classification.

Tables 3 and 4 show the mean accuracies of the left-vs-right classification and MI-vs-rest classification, respectively.

According to the Friedman test, there was a statistically significant difference between the classification results obtained using different features, both in left-vs-right ( $\chi^2 = 25.48$ ,  $p = 0.0025$ ) and MI-vs-rest classification ( $\chi^2 = 22.39$ ,  $p = 0.0077$ ). Post-hoc analysis was conducted using Bonferroni correction for multiple comparisons, implemented in MATLAB

**Table 1. Online left-vs-right MI classification accuracies for each subject and session.** Multitaper PSD features and an LDA classifier were used in the online implementation.

1A	1B	2A	2B	3A	4A	4B	5A	5B	6A	6B	7A	7B	8A	8B	9A	9B	MEAN
48,8	58,5	67,7	61,0	64,4	48,8	56,1	92,7	82,9	51,2	78,1	51,2	41,5	70,7	63,4	58,5	65,9	62,4

doi:10.1371/journal.pone.0168766.t001

**Table 2. Mean left-vs-right MI accuracy obtained using CSP and STFT for feature extraction and four different classifiers.**

Classifier	CSP	STFT	MEAN
SVM, RBF kernel	69,56	65,23	<b>67,39</b>
SVM, linear kernel	70,39	65,46	<b>67,92</b>
Naive Bayes	67,42	61,70	<b>64,56</b>
LDA	70,07	67,13	<b>68,60</b>

doi:10.1371/journal.pone.0168766.t002

**Table 3. Left vs. right hand 5-fold cross-validation accuracy for each subject and session.** The best accuracy for each session is in bold.

SUBJECT/ SESSION	STFT	Morlet	PSD	FBCSP	SSD	CSP	SSD+CSP	CSP+STFT	CSP+Morlet	CSP+PSD	SUBJECT MEAN
1A	75,3	70,3	69,2	76,6	70,4	82,7	<b>82,8</b>	80,4	81,5	79,0	<b>76,8</b>
1B	75,2	71,5	74,0	<b>76,5</b>	60,3	75,1	65,2	74,0	70,4	70,4	<b>71,3</b>
2A	59,4	63,4	<b>70,2</b>	54,1	66,4	66,3	67,5	59,4	65,0	52,9	<b>62,5</b>
2B	70,4	<b>79,0</b>	69,2	67,7	51,8	65,2	66,5	66,5	70,4	66,6	<b>67,3</b>
3A	69,0	74,1	73,8	68,6	62,9	59,6	<b>77,9</b>	57,4	58,4	59,4	<b>66,1</b>
4A	76,5	75,2	<b>79,0</b>	72,9	60,4	75,1	76,8	70,4	71,6	67,9	<b>72,6</b>
4B	58,1	63,0	56,8	64,1	58,0	79,0	76,5	76,5	<b>79,1</b>	72,8	<b>68,4</b>
5A	80,2	77,7	68,0	77,6	<b>81,5</b>	69,3	80,2	75,3	77,9	74,1	<b>76,2</b>
5B	79,0	82,9	77,9	74,2	82,9	86,6	<b>90,2</b>	89,0	84,1	77,9	<b>82,5</b>
6A	66,6	67,9	71,8	65,7	56,6	67,7	75,2	74,3	71,5	<b>75,3</b>	<b>69,3</b>
6B	86,5	<b>91,4</b>	<b>91,4</b>	86,4	59,4	84,0	90,2	87,7	88,9	86,5	<b>85,2</b>
7A	59,3	65,6	62,9	57,1	65,5	65,4	<b>66,6</b>	63,1	61,9	61,8	<b>62,9</b>
7B	49,2	59,0	56,5	56,5	35,7	61,8	59,3	65,4	65,5	<b>65,6</b>	<b>57,5</b>
8A	55,4	64,0	66,5	57,9	45,5	60,4	65,4	65,4	62,9	<b>67,9</b>	<b>61,1</b>
8B	64,6	61,9	61,8	60,7	61,6	57,0	<b>65,7</b>	<b>65,7</b>	60,7	64,6	<b>62,4</b>
9A	66,8	59,3	61,7	65,4	53,1	69,0	76,5	76,6	73,0	<b>79,0</b>	<b>68,0</b>
9B	49,4	55,4	59,2	44,4	65,4	53,3	<b>70,4</b>	64,3	63,0	55,6	<b>58,0</b>
MEAN	<b>67,1</b>	<b>69,5</b>	<b>68,8</b>	<b>66,3</b>	<b>61,0</b>	<b>69,3</b>	<b>73,7</b>	<b>71,3</b>	<b>70,9</b>	<b>69,3</b>	<b>68,7</b>

doi:10.1371/journal.pone.0168766.t003

**Table 4. MI-vs rest 5-fold cross-validation accuracy for each subject and session.** The best accuracy for each session is in bold.

SUBJECT/ SESSION	STFT	Morlet	PSD	FBCSP	SSD	CSP	SSD+CSP	CSP+STFT	CSP+Morlet	CSP+PSD	SUBJECT MEAN
1A	74,0	78,9	79,5	<b>97,5</b>	82,9	95,1	94,4	95,7	94,4	93,8	<b>88,6</b>
1B	83,8	86,3	80,7	<b>99,4</b>	80,2	98,2	<b>99,4</b>	95,7	96,3	92,6	<b>91,3</b>
2A	71,2	69,2	<b>73,8</b>	70,0	70,4	73,2	73,1	72,5	73,2	<b>73,8</b>	<b>72,0</b>
2B	72,1	71,8	72,3	61,5	64,3	71,5	<b>75,2</b>	67,1	67,1	67,1	<b>69,0</b>
3A	85,5	86,5	<b>87,5</b>	70,2	81,5	76,9	79,3	86,1	86,5	80,8	<b>82,1</b>
4A	78,9	80,2	80,2	83,2	72,9	<b>88,2</b>	85,1	85,7	87,6	87,0	<b>82,9</b>
4B	82,6	80,7	85,7	76,3	53,0	<b>87,5</b>	85,7	82,0	83,2	82,6	<b>79,9</b>
5A	67,7	70,2	71,4	73,3	53,0	70,2	68,3	72,1	<b>76,4</b>	73,3	<b>69,6</b>
5B	63,9	67,0	63,3	<b>69,6</b>	67,8	67,8	66,5	63,4	65,2	66,5	<b>66,1</b>
6A	91,3	91,3	88,8	<b>93,2</b>	87,7	90,7	<b>93,2</b>	88,8	90,7	87,6	<b>90,3</b>
6B	78,2	85,7	80,7	<b>88,2</b>	85,2	83,2	85,1	84,5	86,3	78,2	<b>83,5</b>
7A	71,5	72,7	73,9	78,3	64,1	85,1	79,5	85,1	<b>86,4</b>	85,1	<b>78,2</b>
7B	70,2	77,6	74,5	79,5	70,3	84,4	83,2	86,9	<b>87,5</b>	78,2	<b>79,2</b>
8A	73,9	75,2	73,9	68,9	79,2	65,2	<b>80,8</b>	63,9	67,6	64,5	<b>71,3</b>
8B	85,7	86,3	84,5	83,3	82,7	74,5	<b>88,2</b>	71,4	70,7	65,8	<b>79,3</b>
9A	60,5	<b>70,1</b>	68,3	64,6	49,3	58,4	69,5	57,2	62,1	59,0	<b>61,9</b>
9B	80,0	80,8	<b>83,2</b>	63,3	62,9	67,6	75,1	77,0	77,6	74,5	<b>74,2</b>
MEAN	<b>75,9</b>	<b>78,3</b>	<b>77,8</b>	<b>77,7</b>	<b>71,0</b>	<b>78,7</b>	<b>81,3</b>	<b>78,5</b>	<b>79,9</b>	<b>77,1</b>	<b>77,6</b>

doi:10.1371/journal.pone.0168766.t004

**Table 5. Statistical tests ( $p$ -values) of the differences of the tested methods (first column vs. other columns) in the left-vs-right classification (5 cross-validation folds).** Significant ( $p < 0.05$ ) values are in bold.

	STFT	Morlet	PSD	FBCSP	SSD	CSP	SSD+ CSP	CSP+ STFT	CSP+ Morlet	CSP+ PSD
STFT		0.973	0.916	0.180	<b>0.003</b>	0.896	0.350	0.991	0.990	0.875
Morlet	0.027		0.315	<b>0.016</b>	<b>0.000</b>	0.503	0.085	0.906	0.900	0.499
PSD	0.084	0.685		<b>0.038</b>	<b>0.000</b>	0.613	0.121	0.937	0.921	0.602
FBCSP	0.820	0.984	0.962		<b>0.021</b>	0.988	0.638	1.000	1.000	0.991
SSD	0.997	1.000	1.000	0.979		1.000	0.996	1.000	1.000	1.000
CSP	0.104	0.497	0.387	<b>0.012</b>	<b>0.000</b>		<b>0.041</b>	0.966	0.956	0.495
SSD+CSP	0.650	0.915	0.879	0.362	<b>0.004</b>	0.959		0.997	0.995	0.943
CSP+STFT	<b>0.009</b>	0.094	0.063	<b>0.000</b>	<b>0.000</b>	<b>0.034</b>	<b>0.003</b>		0.403	<b>0.016</b>
CSP+Morlet	<b>0.010</b>	0.100	0.079	<b>0.000</b>	<b>0.000</b>	<b>0.044</b>	<b>0.005</b>	0.597		<b>0.015</b>
CSP+PSD	0.125	0.501	0.398	<b>0.009</b>	<b>0.000</b>	0.505	0.057	0.984	0.985	

doi:10.1371/journal.pone.0168766.t005

(R2015a; Mathworks Inc., MA, USA) function *multcompare*. In the case of left-vs-right classification, there were significant differences between the mean ranks of SSD and SSD+CSP ( $p = 0.0018$ ) and FBCSP and SSD+CSP ( $p = 0.0029$ ). Other pairwise comparisons of the methods did not yield significant differences. In case of MI-vs-rest classification, a significant difference was found between the results of SSD and SSD+CSP ( $p = 0.0113$ ) and between SSD and CSP+Morlet ( $p = 0.0156$ ).

The  $p$ -values from paired-samples t-test for the cross-validation folds are shown in Table 5 (left-vs-right) and Table 6 (MI-vs-rest). The results represent the significance level for the hypothesis that the method on each row is better than the methods on the columns. According to these results, CSP+STFT and CSP+Morlet were significantly better than most of the other methods in left-vs-right classification. SSD+CSP yielded significantly better results compared to SSD, but not to any other methods, probably due to the high variance in the folds of cross-validation. On the other hand, SSD+CSP was significantly better compared to all other methods on MI-vs-rest classification.

### 3.5. Inter-session accuracy

CSP+Morlet yielded the best results in inter-session left-vs-right MI classification (mean 69.1%, range 50.0–85.2%), and the average accuracy over all methods and subjects was 65.4%. Table 7 represents the results of the inter-session classification. The session indicated in the

**Table 6. Statistical test ( $p$ -values) of the differences of the tested methods (first column vs. other columns) in the MI-vs-rest classification (5 cross-validation folds).** Significant ( $p < 0.05$ ) values in bold.

	STFT	Morlet	PSD	FBCSP	SSD	CSP	SSD+ CSP	CSP+ STFT	CSP+ Morlet	CSP+ PSD
STFT		0.993	0.860	0.487	<b>0.000</b>	0.961	1.000	0.910	0.984	0.615
Morlet	<b>0.007</b>		0.099	0.156	<b>0.000</b>	0.789	0.999	0.610	0.889	0.200
PSD	0.140	0.901		0.304	<b>0.000</b>	0.926	1.000	0.824	0.971	0.392
FBCSP	0.513	0.844	0.696		<b>0.001</b>	0.982	1.000	0.923	0.993	0.640
SSD	1.000	1.000	1.000	0.999		1.000	1.000	1.000	1.000	1.000
CSP	<b>0.039</b>	0.211	0.074	<b>0.018</b>	<b>0.000</b>		0.991	0.145	0.765	<b>0.001</b>
SSD+CSP	<b>0.000</b>	<b>0.001</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.009</b>		<b>0.001</b>	<b>0.044</b>	<b>0.000</b>
CSP+STFT	0.090	0.390	0.176	0.077	<b>0.000</b>	0.855	0.999		0.997	<b>0.008</b>
CSP+Morlet	<b>0.016</b>	0.111	<b>0.029</b>	<b>0.007</b>	<b>0.000</b>	0.235	0.956	<b>0.003</b>		<b>0.000</b>
CSP+PSD	0.385	0.800	0.608	0.360	<b>0.000</b>	0.999	1.000	0.992	1.000	

doi:10.1371/journal.pone.0168766.t006

**Table 7. Inter-session accuracy for each subject.** The indicated session was used as the training data. The best accuracy for each session is in bold.

SUBJECT/ SESSION	STFT	Morlet	PSD	FBCSP	SSD	CSP	SSD+CSP	CSP+STFT	CSP+Morlet	CSP+PSD	SUBJECT MEAN
1A	60,5	56,8	58,0	54,3	64,2	<b>71,6</b>	69,1	69,1	<b>71,6</b>	69,1	<b>62,2</b>
1B	69,1	71,6	65,4	61,7	69,1	74,1	72,8	71,6	69,1	<b>79,0</b>	<b>64,7</b>
2A	52,6	<b>55,1</b>	52,6	53,9	44,9	50,0	50,0	50,0	50,0	50,0	<b>52,7</b>
2B	54,6	58,4	54,6	<b>59,7</b>	35,1	55,8	55,8	55,8	55,8	55,8	<b>55,4</b>
4A	67,9	74,1	77,8	72,8	58,0	<b>81,5</b>	75,3	79,0	<b>81,5</b>	<b>81,5</b>	<b>73,3</b>
4B	70,4	<b>75,3</b>	72,8	59,3	56,8	69,1	69,1	72,8	<b>75,3</b>	66,7	<b>66,8</b>
5A	81,5	80,3	82,7	71,6	82,7	82,7	<b>84,0</b>	<b>84,0</b>	81,5	77,8	<b>79,5</b>
5B	81,5	84,0	85,2	<b>91,4</b>	82,7	82,7	80,2	82,7	85,2	81,5	<b>81,8</b>
6A	85,2	<b>86,4</b>	84,0	84,0	59,3	70,4	82,7	80,3	81,5	72,8	<b>71,0</b>
6B	66,7	66,7	71,6	<b>81,5</b>	69,1	70,4	66,7	65,4	67,9	64,2	<b>70,1</b>
7A	56,8	<b>67,9</b>	61,7	58,0	56,8	58,0	64,2	63,0	56,8	58,0	<b>56,9</b>
7B	58,0	64,2	58,0	63,0	44,4	61,7	56,8	63,0	<b>66,7</b>	55,6	<b>59,6</b>
8A	63,0	<b>64,2</b>	63,0	54,3	55,6	58,0	<b>64,2</b>	54,3	55,6	51,9	<b>61,3</b>
8B	55,6	60,5	60,5	60,5	55,6	<b>64,2</b>	<b>64,2</b>	61,7	63,0	59,3	<b>59,3</b>
9A	66,7	69,1	65,4	71,6	63,0	60,5	<b>76,5</b>	64,2	65,4	63,0	<b>63,4</b>
9B	65,4	64,2	69,1	69,1	75,3	75,3	72,8	70,4	<b>79,0</b>	75,3	<b>68,7</b>
MEAN	<b>66,0</b>	<b>68,7</b>	<b>67,7</b>	<b>66,7</b>	<b>60,8</b>	<b>67,9</b>	<b>69,0</b>	<b>68,0</b>	<b>69,1</b>	<b>66,3</b>	<b>65,4</b>

doi:10.1371/journal.pone.0168766.t007

first column of the table was used as the training data and the other session for the same subject as testing data. Subject 3 had only one session recorded due to technical problems and was thus excluded from this analysis.

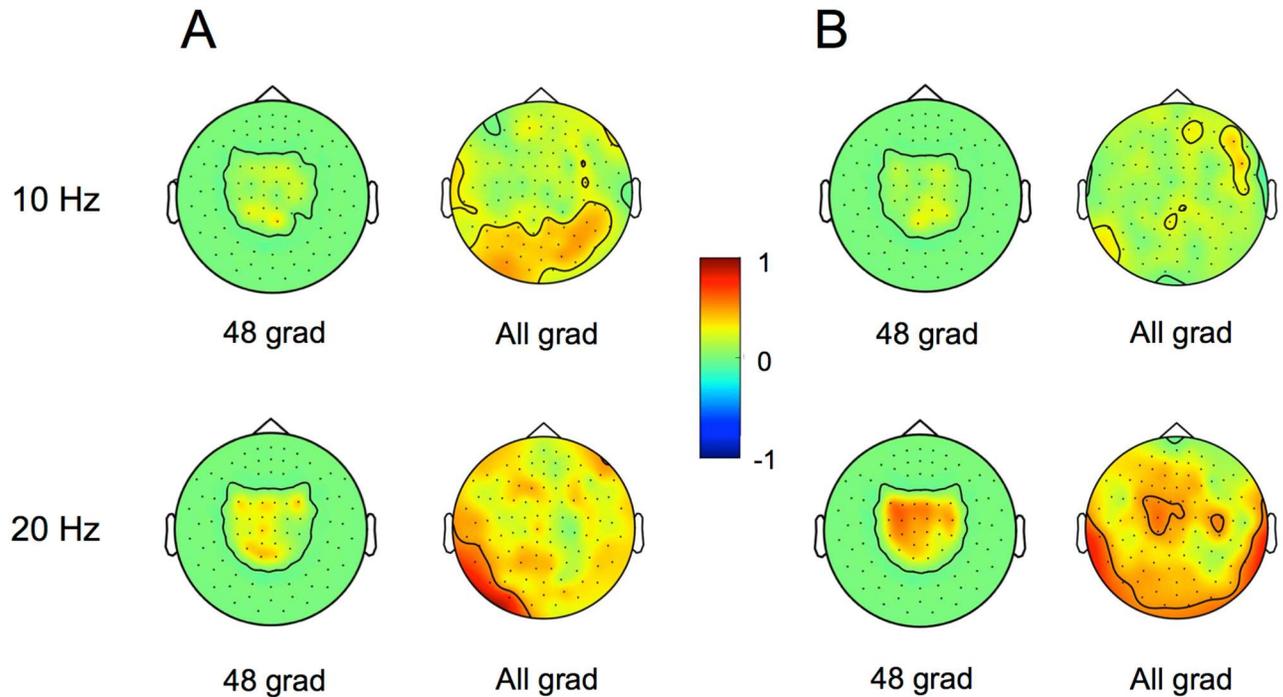
According to the Friedman test, there was a statistically significant difference between the results obtained using different features ( $\chi^2 = 20.79, p = 0.014$ ). However, the post-hoc test conducted with Bonferroni correction did not reveal any significant differences between the methods.

### 3.6. The effect of individual mu-rhythm characteristics on classification accuracy

We investigated how 1) 10- and 20-Hz amplitudes at rest, 2) the suppression of 10- and 20-Hz rhythms during MI were related to the subject's average MI classification accuracy. The correlation was calculated both on gradiometer and CSP-filtered data.

For the data averaged over parietal gradiometers, no correlation was found between left-vs-right mean accuracy and suppression of sensorimotor rhythm during MI (for 10 Hz,  $r = 0.143, p = 0.584$ ; for 20 Hz,  $r = 0.324, p = 0.204$ ). MI-vs-rest mean accuracy was significantly correlated with the 20-Hz suppression amplitude ( $r = 0.510, p = 0.037$ ) but not with the 10-Hz suppression ( $r = -0.088, p = 0.737$ ). The level of sensorimotor rhythms during baseline was not correlated with MI-vs-rest mean accuracy (for 10 Hz,  $r = -0.260, p = 0.314$ ; for 20 Hz,  $r = 0.123, p = 0.639$ ). The results were very similar for the CSP-filtered data averaged over components. Left-vs-right mean accuracy did not correlate with suppression of SMR (for 10 Hz,  $r = 0.182, p = 0.484$ ; for 20 Hz,  $r = 0.326, p = 0.202$ ), nor did SMR power at rest (for 10 Hz,  $r = -0.351, p = 0.168$ ; for 20 Hz,  $r = -0.200, p = 0.441$ ). Also in this case MI-vs-rest mean accuracy was significantly correlated with the 20-Hz suppression amplitude ( $r = 0.582, p = 0.014$ ) but not with the 10-Hz suppression ( $r = 0.008, p = 0.977$ ).

Correlations were also calculated separately for each gradiometer included in the analyses as well as for each CSP-filtered component. Fig 3 illustrates the correlations between mu-rhythm suppression at each gradiometer and average classification accuracy in left-vs-right



**Fig 3. Correlations of classification performance and measures of 10/20-Hz activity at each gradiometer.** Correlations between (A) left-vs-right and (B) MI-vs-rest cross-validated accuracy and 10/20-Hz suppression level at 48 parietal gradiometers (left) and all gradiometers (right). In the case of 48 gradiometers, the black contour lines show the outlines of non-zero correlation values.

doi:10.1371/journal.pone.0168766.g003

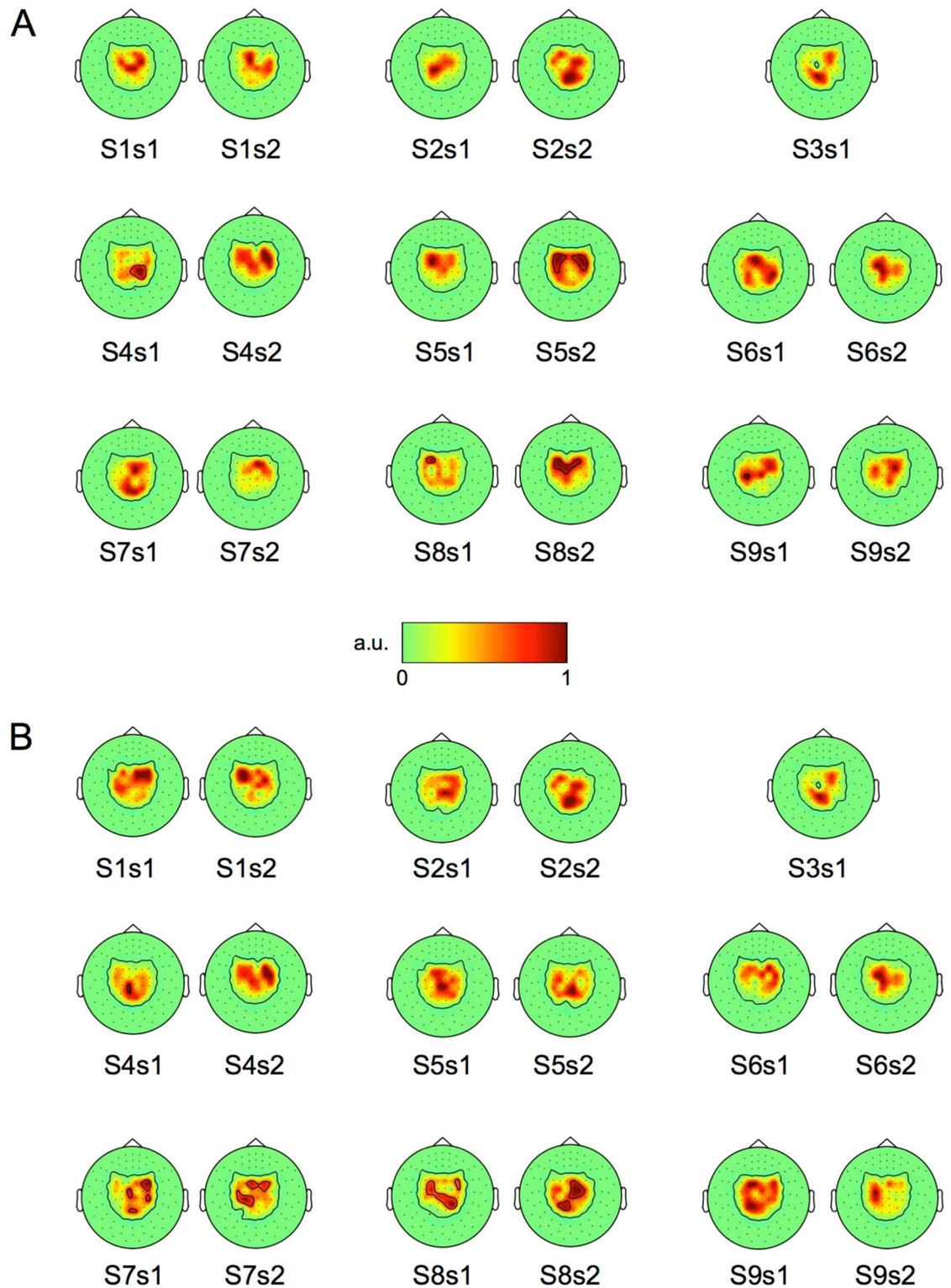
and MI-vs-rest classification tasks. The level of beta-band suppression over the sensorimotor cortex shows high correlation with both left-vs-right and MI-vs-rest accuracy. Alpha-band suppression in occipital areas has a fairly high correlation with left-vs-right accuracy. Fig 4 shows, for each subject and session, the spatial pattern of the CSP component showing the highest mu-rhythm suppression in alpha and beta bands.

#### 4. Discussion

The main objective of this study was to validate optimal feature extraction methods for characterizing single-trial hand MI in the context of an MEG-based neurofeedback system. Multiple spatial filtering and time—frequency analysis methods, commonly used for feature extraction from EEG/MEG signals, were evaluated in terms of their ability to capture relevant features for subsequent classification.

##### 4.1. Evaluation of feature extraction methods

According to our results, spatial filtering methods in general outperformed time—frequency methods in all single-trial classification tasks. In addition, time—frequency features extracted from CSP-filtered signals yielded better classification than the corresponding time—frequency features extracted from gradiometer signals. Among the evaluated methods, the best was the combination of SSD and CSP, which outperformed other decomposition methods (SSD, CSP and Filter-Bank CSP) in both left-vs-right and MI-vs-rest classification. CSP has been found efficient for discriminating single-trial MI patterns [22–24], and our results agreed with these findings, since already the basic CSP yielded good accuracies in the current study. However, adding SSD filtering prior to CSP further improved the classification. As has been shown by



**Fig 4. Spatial patterns of the CSP components showing the highest 10/20-Hz suppression for each subject (S) and session (s).** Components showing the highest (A) alpha-band and (B) beta-band suppression during MI. The scale is normalized to 0–1 (arbitrary units) for clarity. The outermost black contour lines show the outline of non-zero values.

doi:10.1371/journal.pone.0168766.g004

Blankertz and colleagues [24], CSP is the optimal method for the extraction of variance differences between two classes, and a prior linear decomposition does not add any new information to that. The increase in classifier performance was most likely due to decreased overfitting compared to CSP alone [36]. The superior performance of SSD+CSP on both left-vs-right and MI-vs-rest indicates the flexibility of the method in discriminating the relevant oscillatory activity from noise. Another advantage of spatial filtering with linear decomposition is low computational cost: the computation time of the proposed methods was negligible (below 0.5 s per epoch for all methods) and thus they are suitable for a real-time interface. In the real-time setup, additional time could be saved if the filters were fitted to the training data only in the beginning of the experiment.

## 4.2. Accuracies in different classification tasks

As expected, the discrimination between baseline and MI yielded better accuracy than left-vs-right-hand MI classification. Regarding the possible application of these methods in neurofeedback rehabilitation, the classification between baseline and MI might be as important as left-vs-right classification. Many stroke patients with motor disabilities have difficulties in mu rhythm modulation in the affected hemisphere. Instead, the unaffected hemisphere shows increased motor-related activity due to reduced intracortical inhibition, as shown during voluntary movements [43] and using paired-pulse transcranial magnetic stimulation [44]. Thus, in the early phase of rehabilitation it might be more useful to target discriminating rest and MI, and once the patients are capable of producing a detectable mu-rhythm suppression during MI of the affected limb, the training could focus on differentiating right- and left-side MI.

Despite the good results obtained with healthy subjects, it is yet impossible to predict the performance of the presented methods in patients. Sensorimotor brain activity is often disrupted and variable in patients with a stroke in the territory of middle cerebral artery. Therefore, training the classifier with individual MEG/EEG data might not be the optimal approach. Tangwiriyasakul and colleagues [45] reported that using motor execution -related EEG signals from the unaffected hemisphere of acute hemiparetic stroke patients was sufficient for training a classifier that differentiated rest from MI. One could also train the classifier with a large database of healthy subjects' MI-related signals. In this case, the neurofeedback-assisted rehabilitation could be initiated without first collecting the individual (likely low-quality) training data from each patient. However, decoding one subject's brain activity using data transferred from another subject is challenging and requires additional regularization of the spatial filters [46]. Yet another plausible solution for training the classifier might be to use data collected during passive movements, as suggested by Kaiser and co-workers [47], or during functional electrical stimulation of the target muscles [48].

The within-session left-vs-right accuracies were higher than the inter-session classification accuracies, indicating that there was some variation in the measured brain responses between the two sessions of the same subject. Spatial filtering methods involving CSP performed well also in this situation, yielding above-chance decoding accuracies for most subjects. These findings suggest that spatial filtering increases the robustness of MI classification also over sessions. It should be noted that LWS regularization of the covariance matrix was applied to all spatial filters, as such regularization is known to reduce classifier overfitting.

## 4.3. Inter-subject differences

Classification accuracies varied substantially across the subjects, even though all of them were naïve to neurofeedback and BCI training. The best subjects achieved > 90% left-vs-right accuracy with several different feature extraction methods, whereas the least successful subjects

hardly performed over chance level (50%). A likely reason for the low accuracies is the relatively low number of training trials. Previous studies reporting excellent MI classification accuracies have used both highly trained subjects and a large number of trials for both training and testing. For example, the MEG/EEG datasets of Berlin Brain—Computer Interface Competition [49] include hundreds of MI trials recorded from trained subjects. Foldes and colleagues [50] reported that only five minutes of training data are sufficient for training an MEG-based BCI, but they used overt movements instead of MI, and the discrimination was done only between rest and movement. However, it is noteworthy that in our study up to 91% left-vs-right MI classification accuracies were achieved despite the small amount of data and untrained subjects.

We assumed that the inter-subject differences in classification performance could be explained by the differences in the suppression of mu rhythm during MI. As expected, we found that the level of the 20-Hz suppression correlated significantly with the subject's average accuracy of MI-vs-rest classification. This correlation was slightly stronger on the left hemisphere (Fig 3), probably due to the fact that all subjects were right-handed. In contrast, the left-vs-right classification accuracy could not be explained by the level of mu rhythm suppression. A possible reason for this finding is that the spatial patterns could be more relevant in discriminating right- and left-hand MI, and these patterns varied significantly between subjects. This variation is illustrated in Fig 4, which represents the CSP components with the highest alpha- and beta-band suppression: the spatial patterns tend to be more similar between sessions than between subjects.

In contrast to the results of Blankertz and colleagues [51], we did not find a significant correlation between mu rhythm baseline amplitude and classification accuracy. This finding is probably due to the fact that in our study the baseline amplitude was measured during the short 5-s rest periods between the MI epochs, whereas Blankertz and colleagues collected the baseline data in a separate longer measurement during which the subjects were resting eyes closed. Although in our study the subjects were instructed to perform MI only during the cued periods, it is possible that movement-related activity was present also during the rest periods.

It has been reported that some subjects are BCI illiterate, i.e. not able to operate an MI-BCI despite excessive training [51,52], and our results suggest that insufficient mu rhythm suppression can partially explain this finding. Guger et al. [53] reported that 93% of untrained subjects achieved over 60% left-vs-right MI accuracy after two sessions of training, and 7% of subjects remained below 60% performance. Our off-line cross-validated accuracies are not directly comparable to the aforementioned results, but it is noteworthy that all nine subjects achieved higher than 59% cross-validated left-vs-right accuracy with SSD+CSP features. It is also likely that the subjects can improve mu-rhythm modulation, and thus the classification accuracy, with repetitive BCI training. It would be interesting to examine whether the least successful subjects are able to achieve sufficient control over the MI-based BCI after several training sessions. On the other hand, unsuccessful subjects could be excluded from further studies by measuring suppression of sensorimotor rhythms in an initial screening session.

#### 4.4. Limitations and future studies

There were certain limitations in the current study that should be addressed before we can design a real-time MEG neurofeedback system feasible for e.g. rehabilitation. First, we only used visual stimuli and feedback in our experimental paradigm. However, proprioceptive feedback could enhance the subjective feeling of successful motor imagery and thus be a more natural feedback modality for neurofeedback learning [54,55], especially during rehabilitation of motor functions [56]. In a recent study, Piitulainen and co-workers [57] introduced an MEG-

compatible pneumatic hand/foot stimulator, which would be suitable for delivering the proprioceptive feedback. In further studies, we will investigate the effect of simultaneous MI and proprioceptive stimulation to mu rhythm modulation and decoding accuracy.

Second, in the current study we did not apply inverse modeling prior to feature extraction since our main goal was to examine whether sufficient decoding can be achieved with features extracted in sensor space and linearly-transformed subspaces. Operating in sensor space and SSD/CSP-transformed subspaces requires less computation and subject preparation than source-space analysis; for example, MRI scans prior to MEG measurements are not needed. It has been argued by Blankertz and colleagues [24] that CSP is an optimal method for finding variance differences in a given dataset, and inverse modeling does not bring any additional information to that problem. However, several studies have shown that source-level information increases the separability of various MI tasks [17,58,59] and would thus be beneficial in BCI systems. In addition, source-level analysis would have the spatial specificity required for targeting specific brain regions. This issue might be especially critical when dealing with classification tasks with more than two classes, as CSP is often suboptimal for such classification. Source imaging has been successfully applied to several real-time neurofeedback experiments using EEG [58,60] and MEG [61–64], and used in an offline classification analysis [65]. The added value of source-space feature extraction and classification will be evaluated in our further studies.

Third, the experimental protocol of this study was synchronous, i.e. the timing of epochs was pre-determined and the subjects were instructed to follow the visual cues indicating MI and rest. In recent studies, also asynchronous approaches for MI-BCI feedback have been proposed [66,67], in which the subject is allowed to perform MI at a self-determined pace. In this case, the data analysis should be done in sliding time windows over the course of the measurement, which requires more computation than processing one epoch at a time. In the current study, feature extraction and classification were done only during the discrete MI epochs and the feedback was shown immediately after the epoch. In the asynchronous feedback paradigm, the output of the classifier should be updated continuously, preferably with a minimal delay. Although this approach might be computationally more challenging, the asynchronous paradigm would be more realistic and thus give the user a stronger feeling of control over the neurofeedback system.

## 5. Conclusions

We compared different feature extraction methods for classifying MI-related MEG signals measured from healthy, untrained subjects. Spatial filtering with combined SSD and CSP outperformed other methods in discriminating left- vs. right-hand MI and rest from MI. Other spatial filtering methods yielded good accuracies as well, and CSP filtering also improved the discriminability of all time—frequency features. The level of 20-Hz suppression during MI correlated with the subjective MI-vs-rest classification accuracy, implying the potential use of 20-Hz suppression as a pre-screening tool before BCI training.

## Author Contributions

**Conceptualization:** HLH LP.

**Data curation:** HLH.

**Formal analysis:** HLH.

**Funding acquisition:** HLH LP.

**Investigation:** HLH.

**Methodology:** HLH LP.

**Project administration:** LP.

**Resources:** LP.

**Software:** HLH LP.

**Supervision:** LP.

**Validation:** HLH.

**Visualization:** HLH.

**Writing – original draft:** HLH.

**Writing – review & editing:** HLH LP.

## References

1. Birbaumer N, Cohen LG. Brain-computer interfaces: communication and restoration of movement in paralysis. *J Physiol*. 2007 Mar 15; 579(3):621–36.
2. Daly JJ, Wolpaw JR. Brain—computer interfaces in neurological rehabilitation. *Lancet Neurol*. 2008; 7(11):1032–43. doi: [10.1016/S1474-4422\(08\)70223-0](https://doi.org/10.1016/S1474-4422(08)70223-0) PMID: [18835541](https://pubmed.ncbi.nlm.nih.gov/18835541/)
3. Ang KK, Guan C, Chua KSG, Ang BT, Kuah C, Wang C, et al. Clinical study of neurorehabilitation in stroke using EEG-based motor imagery brain-computer interface with robotic feedback. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. 2010. p. 5549–52.
4. Pfurtscheller G, Lopes da Silva FH. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiol*. 1999; 110(11):1842–57. PMID: [10576479](https://pubmed.ncbi.nlm.nih.gov/10576479/)
5. Salmelin R, Hari R. Spatiotemporal characteristics of sensorimotor neuromagnetic rhythms related to thumb movement. *Neuroscience*. 1994; 60(2):537–50. PMID: [8072694](https://pubmed.ncbi.nlm.nih.gov/8072694/)
6. Pfurtscheller G, Neuper C. Motor imagery activates primary sensorimotor area in humans. *Neurosci Lett*. 1997; 239(2–3):65–8. PMID: [9469657](https://pubmed.ncbi.nlm.nih.gov/9469657/)
7. Schnitzler A, Salenius S, Salmelin R, Jousmäki V, Hari R. Involvement of Primary Motor Cortex in Motor Imagery: A Neuromagnetic Study. *NeuroImage*. 1997; 6(3):201–8. doi: [10.1006/nimg.1997.0286](https://doi.org/10.1006/nimg.1997.0286) PMID: [9344824](https://pubmed.ncbi.nlm.nih.gov/9344824/)
8. Kauhanen L, Rantanen P, Lehtonen J, Tarnanen I, Alaranta H, Sams M. Sensorimotor cortical activity of tetraplegics during attempted finger movements. *Biomed Tech*. 2004; 49(1):59–60.
9. Cochin S, Barthelemy C, Roux S, Martineau J. Observation and execution of movement: similarities demonstrated by quantified electroencephalography. *Eur J Neurosci*. 1999 May; 11(5):1839–42. PMID: [10215938](https://pubmed.ncbi.nlm.nih.gov/10215938/)
10. McFarland DJ, Miner L a, Vaughan TM, Wolpaw JR. Mu and beta rhythm topographies during motor imagery and actual movements. *Brain Topogr*. 2000; 12(3):177–86. PMID: [10791681](https://pubmed.ncbi.nlm.nih.gov/10791681/)
11. Miller KJ, Schalk G, Fetz EE, den Nijs M, Ojemann JG, Rao RPN. Cortical activity during motor execution, motor imagery, and imagery-based online feedback. *Proc Natl Acad Sci U S A*. 2010; 107(9):4430–5. doi: [10.1073/pnas.0913697107](https://doi.org/10.1073/pnas.0913697107) PMID: [20160084](https://pubmed.ncbi.nlm.nih.gov/20160084/)
12. Cohen D, Cuffin BN. EEG versus MEG localization accuracy: Theory and experiment. *Brain Topogr*. 1991; 4(2):95–103. PMID: [1793693](https://pubmed.ncbi.nlm.nih.gov/1793693/)
13. Mellinger J, Schalk G, Braun C, Preissl H, Rosenstiel W, Birbaumer N, et al. An MEG-based Brain-Computer Interface (BCI). *NeuroImage*. 2007; 36(3):581–93. doi: [10.1016/j.neuroimage.2007.03.019](https://doi.org/10.1016/j.neuroimage.2007.03.019) PMID: [17475511](https://pubmed.ncbi.nlm.nih.gov/17475511/)
14. Buch E, Weber C, Cohen LG, Braun C, Dimyan MA, Ard T, et al. Think to Move: a Neuromagnetic Brain-Computer Interface (BCI) System for Chronic Stroke. *Stroke*. 2008 Mar 1; 39(3):910–7. doi: [10.1161/STROKEAHA.107.505313](https://doi.org/10.1161/STROKEAHA.107.505313) PMID: [18258825](https://pubmed.ncbi.nlm.nih.gov/18258825/)
15. Foldes ST, Weber DJ, Collinger JL. MEG-based neurofeedback for hand rehabilitation. *J Neuroeng Rehabil*. 2015 Dec 22; 12(1):85.
16. Pfurtscheller G, Neuper C. Motor imagery and direct brain-computer communication. *Proc IEEE*. 2001 Jul; 89(7):1123–34.

17. He B, Baxter B, Edelman BJ, Cline CC, Wenjing WY. Noninvasive brain-computer interfaces based on sensorimotor rhythms. *Proc IEEE*. 2015; 103(6):907–25.
18. Wolpaw JR, McFarland DJ. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proc Natl Acad Sci U S A*. 2004 Dec 21; 101(51):17849–54. doi: [10.1073/pnas.0403504101](https://doi.org/10.1073/pnas.0403504101) PMID: [15585584](https://pubmed.ncbi.nlm.nih.gov/15585584/)
19. McFarland DJ, Sarnacki WA, Wolpaw JR. Electroencephalographic (EEG) control of three-dimensional movement. *J Neural Eng*. 2010 Jun 1; 7(3):36007.
20. Doud A, Lucas J, Pisansky M, He B. Continuous three-dimensional control of a virtual helicopter using a motor imagery based brain-computer interface. *PLoS One*. 2011;
21. LaFleur K, Cassady K, Doud A, Shades K, Rogin E, He B. Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain—computer interface. *J Neural Eng*. 2013 Aug 1; 10(4):46003.
22. Koles ZJ, Lazar MS, Zhou SZ. Spatial patterns underlying population differences in the background EEG. *Brain Topogr*. 1990; 2(4):275–84. PMID: [2223384](https://pubmed.ncbi.nlm.nih.gov/2223384/)
23. Blankertz B, Muller K-R, Curio G, Vaughan TM, Schalk G, Wolpaw JR, et al. The BCI Competition 2003: Progress and Perspectives in Detection and Discrimination of EEG Single Trials. *IEEE Trans Biomed Eng*. 2004 Jun; 51(6):1044–51. doi: [10.1109/TBME.2004.826692](https://doi.org/10.1109/TBME.2004.826692) PMID: [15188876](https://pubmed.ncbi.nlm.nih.gov/15188876/)
24. Blankertz B, Tomioka R, Lemm S, Kawanabe M, Muller K. Optimizing Spatial filters for Robust EEG Single-Trial Analysis. *IEEE Signal Process Mag*. 2008; 25(1):41–56.
25. Guger C, Ramoser H, Pfurtscheller G. Real-time EEG analysis with subject-specific spatial patterns for a brain-computer interface (BCI). *IEEE Trans Rehabil Eng*. 2000 Dec; 8(4):447–56. PMID: [11204035](https://pubmed.ncbi.nlm.nih.gov/11204035/)
26. Wang Y, Gao S, Gao X. Common Spatial Pattern Method for Channel Selection in Motor Imagery Based Brain-computer Interface. In: 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference. IEEE; 2005. p. 5392–5.
27. Ang KK, Chin ZY, Zhang H, Guan C. 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). IEEE; 2008. p. 2390–7.
28. Lemm S, Blankertz B, Curio G, Müller K-R. Spatio-spectral filters for improving the classification of single trial EEG. *IEEE Trans Biomed Eng*. 2005 Sep; 52(9):1541–8. doi: [10.1109/TBME.2005.851521](https://doi.org/10.1109/TBME.2005.851521) PMID: [16189967](https://pubmed.ncbi.nlm.nih.gov/16189967/)
29. Nikulin V V., Nolte G, Curio G. A novel method for reliable and fast extraction of neuronal EEG/MEG oscillations on the basis of spatio-spectral decomposition. *NeuroImage*. 2011; 55(4):1528–35. doi: [10.1016/j.neuroimage.2011.01.057](https://doi.org/10.1016/j.neuroimage.2011.01.057) PMID: [21276858](https://pubmed.ncbi.nlm.nih.gov/21276858/)
30. Wang T, Deng J, He B. Classifying EEG-based motor imagery tasks by means of time—frequency synthesized spatial patterns. *Clin Neurophysiol*. 2004; 115(12):2744–53. doi: [10.1016/j.clinph.2004.06.022](https://doi.org/10.1016/j.clinph.2004.06.022) PMID: [15546783](https://pubmed.ncbi.nlm.nih.gov/15546783/)
31. Pfurtscheller G, Brunner C, Schlögl A, Lopes da Silva FH. Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. *NeuroImage*. 2006; 31(1):153–9. doi: [10.1016/j.neuroimage.2005.12.003](https://doi.org/10.1016/j.neuroimage.2005.12.003) PMID: [16443377](https://pubmed.ncbi.nlm.nih.gov/16443377/)
32. Herman P, Prasad G, McGinnity TM, Coyle D. Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification. *IEEE Trans Neural Syst Rehabil Eng*. 2008 Aug; 16(4):317–26. doi: [10.1109/TNSRE.2008.926694](https://doi.org/10.1109/TNSRE.2008.926694) PMID: [18701380](https://pubmed.ncbi.nlm.nih.gov/18701380/)
33. Sudre G, Parkkonen L, Bock E, Baillet S, Wang W, Weber DJ. rtMEG: A Real-Time Software Interface for Magnetoencephalography. *Comput Intell Neurosci*. 2011; 2011:1–7.
34. Spüler M, Rosenstiel W, Bogdan M. Using Coherence for Robust Online Brain-Computer Interface (BCI) Control. *Nonlinear Dyn Electron Syst*. 2014;363–70.
35. Kang S, Ahn M, Jun SC. Performances among various common spatial pattern methods for simultaneous MEG/EEG data. In: Ryu J, Chong KT, Ikeura R, Han Q, editors. *Proc SPIE 7500, ICMIT 2009: Mechatronics and Information Technology, 75000X*. International Society for Optics and Photonics; 2010.
36. Haufe S, Dähne S, Nikulin V V. Dimensionality reduction for the analysis of brain oscillations. *NeuroImage*. 2014; 101:583–97. doi: [10.1016/j.neuroimage.2014.06.073](https://doi.org/10.1016/j.neuroimage.2014.06.073) PMID: [25003816](https://pubmed.ncbi.nlm.nih.gov/25003816/)
37. Oostenveld R, Fries P, Maris E, Schoffelen J-M. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput Intell Neurosci*. 2011; 2011:1–9.
38. Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, et al. MEG and EEG data analysis with MNE-Python. *Front Neurosci*. 2013; 7:267. doi: [10.3389/fnins.2013.00267](https://doi.org/10.3389/fnins.2013.00267) PMID: [24431986](https://pubmed.ncbi.nlm.nih.gov/24431986/)

39. Peirce JW. PsychoPy-Psychophysics software in Python. *J Neurosci Methods*. 2007; 162(1–2):8–13. doi: [10.1016/j.jneumeth.2006.11.017](https://doi.org/10.1016/j.jneumeth.2006.11.017) PMID: [17254636](https://pubmed.ncbi.nlm.nih.gov/17254636/)
40. Taulu S, Kajola M, Simola J. Suppression of interference and artifacts by the signal space separation method. *Brain Topogr*. 2004 Jan; 16(4):269–75. PMID: [15379226](https://pubmed.ncbi.nlm.nih.gov/15379226/)
41. Neuper C, Scherer R, Reiner M, Pfurtscheller G. Imagery of motor actions: Differential effects of kinesthetic and visual-motor mode of imagery in single-trial EEG. *Cogn Brain Res*. 2005; 25(3):668–77.
42. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. *J Mach Learn Res*. 2011; 12(Oct):2825–30.
43. Murase N, Duque J, Mazzocchio R, Cohen LG. Influence of Interhemispheric Interactions on Motor Function in Chronic Stroke. *Ann Neurol*. 2004 Mar; 55(3):400–9. doi: [10.1002/ana.10848](https://doi.org/10.1002/ana.10848) PMID: [14991818](https://pubmed.ncbi.nlm.nih.gov/14991818/)
44. Manganotti P, Acler M, Zanette GP, Smania N, Fiaschi A. Motor Cortical Disinhibition During Early and Late Recovery After Stroke. *Neurorehabil Neural Repair*. 2007; 22(4):396–403.
45. Tangwiriyasakul C, Mocioiu V, van Putten MJ a M, Rutten WLC. Classification of motor imagery performance in acute stroke. *J Neural Eng*. 2014 Jun 1; 11(3):36001.
46. Lotte F, Guan C. Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms. *IEEE Trans Biomed Eng*. 2011 Feb; 58(2):355–62. doi: [10.1109/TBME.2010.2082539](https://doi.org/10.1109/TBME.2010.2082539) PMID: [20889426](https://pubmed.ncbi.nlm.nih.gov/20889426/)
47. Kaiser V, Kreiling A, Müller-Putz GR, Neuper C. First Steps Toward a Motor Imagery Based Stroke BCI: New Strategy to Set up a Classifier. *Front Neurosci*. 2011; 5:86. doi: [10.3389/fnins.2011.00086](https://doi.org/10.3389/fnins.2011.00086) PMID: [21779234](https://pubmed.ncbi.nlm.nih.gov/21779234/)
48. Vidaurre C, Pascual J, Ramos-Murguialday A, Lorenz R, Blankertz B, Birbaumer N, et al. Neuromuscular electrical stimulation induced brain patterns to decode motor imagery. *Clin Neurophysiol*. 2013; 124(9):1824–34. doi: [10.1016/j.clinph.2013.03.009](https://doi.org/10.1016/j.clinph.2013.03.009) PMID: [23642833](https://pubmed.ncbi.nlm.nih.gov/23642833/)
49. Blankertz B, Müller K-R, Krusienski DJ, Schalk G, Wolpaw JR, Schlögl A, et al. The BCI competition. III: Validating alternative approaches to actual BCI problems. *IEEE Trans Neural Syst Rehabil Eng*. 2006 Jun; 14(2):153–9. doi: [10.1109/TNSRE.2006.875642](https://doi.org/10.1109/TNSRE.2006.875642) PMID: [16792282](https://pubmed.ncbi.nlm.nih.gov/16792282/)
50. Foldes ST, Vinjamuri RK, Wang W, Weber DJ, Collinger JL. Stability of MEG for real-time neurofeedback. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2011. p. 5778–81.
51. Blankertz B, Sannelli C, Halder S, Hammer EM, Kübler A, Müller K-R, et al. Neurophysiological predictor of SMR-based BCI performance. *NeuroImage*. 2010 Jul 15; 51(4):1303–9. doi: [10.1016/j.neuroimage.2010.03.022](https://doi.org/10.1016/j.neuroimage.2010.03.022) PMID: [20303409](https://pubmed.ncbi.nlm.nih.gov/20303409/)
52. Ahn M, Jun SC. Performance variation in motor imagery brain—computer interface: A brief review. *J Neurosci Methods*. 2015; 243:103–10. doi: [10.1016/j.jneumeth.2015.01.033](https://doi.org/10.1016/j.jneumeth.2015.01.033) PMID: [25668430](https://pubmed.ncbi.nlm.nih.gov/25668430/)
53. Guger C, Edlinger G, Harkam W, Niedermayer I, Pfurtscheller G. How many people are able to operate an EEG-based brain-computer interface (BCI)? *IEEE Trans Neural Syst Rehabil Eng*. 2003 Jun; 11(2):145–7. doi: [10.1109/TNSRE.2003.814481](https://doi.org/10.1109/TNSRE.2003.814481) PMID: [12899258](https://pubmed.ncbi.nlm.nih.gov/12899258/)
54. Gomez-Rodriguez M, Peters J, Hill J, Schölkopf B, Gharabaghi A, Grosse-Wentrup M. Closing the sensorimotor loop: haptic feedback facilitates decoding of motor imagery. *J Neural Eng*. 2011 Jun; 8(3):36005.
55. Ramos-Murguialday A, Schürholz M, Caggiano V, Wildgruber M, Caria A, Hammer EM, et al. Proprioceptive feedback and brain computer interface (BCI) based neuroprostheses. *PLoS One*. 2012; 7(10):e47048. doi: [10.1371/journal.pone.0047048](https://doi.org/10.1371/journal.pone.0047048) PMID: [23071707](https://pubmed.ncbi.nlm.nih.gov/23071707/)
56. Ang KK, Guan C, Phua KS, Wang C, Zhou L, Tang KY, et al. Brain-computer interface-based robotic end effector system for wrist and hand rehabilitation: results of a three-armed randomized controlled trial for chronic stroke. *Front Neuroeng*. 2014; 7(July):30.
57. Piitulainen H, Bourguignon M, Hari R, Jousmäki V. MEG-compatible pneumatic stimulator to elicit passive finger and toe movements. *NeuroImage*. 2015; 112:310–7. doi: [10.1016/j.neuroimage.2015.03.006](https://doi.org/10.1016/j.neuroimage.2015.03.006) PMID: [25770989](https://pubmed.ncbi.nlm.nih.gov/25770989/)
58. Edelman BJ, Baxter B, He B. EEG Source Imaging Enhances the Decoding of Complex Right-Hand Motor Imagery Tasks. *IEEE Trans Biomed Eng*. 2016 Jan; 63(1):4–14. doi: [10.1109/TBME.2015.2467312](https://doi.org/10.1109/TBME.2015.2467312) PMID: [26276986](https://pubmed.ncbi.nlm.nih.gov/26276986/)
59. Yuan H, Doud A, Gururajan A, Bin He. Cortical Imaging of Event-Related (de)Synchronization During Online Control of Brain-Computer Interface Using Minimum-Norm Estimates in Frequency Domain. *IEEE Trans Neural Syst Rehabil Eng*. 2008 Oct; 16(5):425–31. doi: [10.1109/TNSRE.2008.2003384](https://doi.org/10.1109/TNSRE.2008.2003384) PMID: [18990646](https://pubmed.ncbi.nlm.nih.gov/18990646/)
60. Qin L, Ding L, He B. Motor Imagery Classification by Means of Source Analysis for Brain Computer Interface Applications. *J Neural Eng*. 2005; 2(4):65–72. PMID: [16317229](https://pubmed.ncbi.nlm.nih.gov/16317229/)

61. Ahn M, Hong JH, Jun SC. Source space based brain computer interface. In: 17th International Conference on Biomagnetism Advances in Biomagnetism—Biomag2010. In: IFMBE Proceedings. 2010. p. 366–9.
62. Lin PT, Sharma K, Holroyd T, Battapady H, Fei D, Bai O. A High Performance MEG Based BCI Using Single Trial Detection of Human Movement Intention. *Funct Brain Mapp Endeavor to Understand Work Brain*, Signorelli Fr Dr (Ed), InTech. 2013;
63. Boe S, Gionfriddo A, Kraeutner S, Tremblay A, Little G, Bardouille T. Laterality of brain activity during motor imagery is modulated by the provision of source level neurofeedback. *NeuroImage*. 2014; 101:159–67. doi: [10.1016/j.neuroimage.2014.06.066](https://doi.org/10.1016/j.neuroimage.2014.06.066) PMID: [24999037](https://pubmed.ncbi.nlm.nih.gov/24999037/)
64. Florin E, Bock E, Baillet S. Targeted reinforcement of neural oscillatory activity with real-time neuroimaging feedback. *NeuroImage*. 2014; 88:54–60. doi: [10.1016/j.neuroimage.2013.10.028](https://doi.org/10.1016/j.neuroimage.2013.10.028) PMID: [24211817](https://pubmed.ncbi.nlm.nih.gov/24211817/)
65. Battapady H, Lin P, Fei D-Y, Huang D, Bai O. Single trial detection of human movement intentions from SAM-filtered MEG signals for a high performance two-dimensional BCI. In: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2009. p. 524–7.
66. Townsend G, Graimann B, Pfurtscheller G. Continuous EEG classification during motor imagery—Simulation of an asynchronous BCI. *IEEE Trans Neural Syst Rehabil Eng*. 2004; 12(2):258–65. doi: [10.1109/TNSRE.2004.827220](https://doi.org/10.1109/TNSRE.2004.827220) PMID: [15218939](https://pubmed.ncbi.nlm.nih.gov/15218939/)
67. Kus R, Valbuena D, Zygiereicz J, Malechka T, Graeser A. Asynchronous BCI Based on Motor Imagery With Automated Calibration and Neurofeedback Training. *IEEE Trans Neural Syst Rehabil Eng*. 2012; 20(6):823–35. doi: [10.1109/TNSRE.2012.2214789](https://doi.org/10.1109/TNSRE.2012.2214789) PMID: [23033330](https://pubmed.ncbi.nlm.nih.gov/23033330/)