Standardizing Motor Imagery Classification with MOABB

Kipngeno Koech
Carnegie Mellon University Africa
bkoech@andrew.cmu.edu

I. Abstract

***Abstract*—Motor imagery-based brain–computer interfaces (BCIs) have gained attention for their potential to restore communication and control to individuals with severe motor impairments. However, reproducible benchmarking across datasets and models remains a challenge. In this work, we evaluate multiple EEG classification pipelines using the MOABB (Mother of All BCI Benchmarks) framework on the PhysioNet Motor Imagery dataset. We explore spatial filtering and classification strategies including CSP+LDA, Riemannian geometry-based methods, and filter bank CSP. Among these, the CSP + Support Vector Machine (SVM) pipeline demonstrated promising performance. We further investigate this model across all 109 subjects, visualizing CSP-transformed feature distributions and identifying the most influential EEG channels. Our findings underscore the value of combining interpretable spatial filters with robust classifiers and highlight subject-level variability as a critical factor in EEG-based BCI development.**

II. Introduction

Brain–computer interfaces (BCIs) aim to establish a direct communication pathway between the brain and external devices, enabling users to control systems using neural signals alone. Among various BCI paradigms, motor imagery (MI) - the mental rehearsal of movement without actual execution - has emerged as a non-invasive and widely studied approach, particularly using electroencephalography (EEG) [1].

However, decoding motor imagery from EEG remains a challenging task due to the signal's low signal-to-noise ratio, subject-specific variability, and non-stationary characteristics [2]. A wide range of preprocessing, feature extraction, and classification strategies have been proposed, yet comparing these methods across datasets has often been inconsistent and irreproducible.

To address this, the MOABB (Mother of All BCI Benchmarks) framework was introduced as a standardized platform for benchmarking BCI pipelines across multiple datasets using uniform evaluation protocols [3]. In this work, we leverage MOABB to evaluate and compare several established EEG classification pipelines on the PhysioNet Motor Imagery dataset [4].

We assess four pipelines: CSP+LDA, CSP+SVM, Riemannian Geometry with Tangent Space mapping, and Filter Bank CSP. Based on initial results, we perform an in-depth analysis of the CSP+SVM pipeline, investigating component-wise feature distributions and spatial contributions of EEG channels. Our objective is to evaluate not only performance but also interpretability and reproducibility in motor imagery classification.

Problem Statement: Despite extensive research in EEG-based motor imagery classification, comparing the performance of different machine learning pipelines remains difficult due to inconsistencies in datasets, preprocessing strategies, and evaluation protocols. This lack of standardization hinders reproducibility and the identification of robust, generalizable models. There is a need for a standardized benchmarking approach to evaluate and interpret multiple classification pipelines on large, open-access EEG datasets.

III. Related Work

EEG-based brain–computer interfaces (BCIs) have been widely explored for decoding motor imagery (MI) tasks, which involve the imagination of limb movement without actual execution. Early foundational work by Pfurtscheller and Neuper [1] demonstrated the potential of EEG rhythms, particularly in the mu (8–13 Hz) and beta (13–30 Hz) bands, to reflect motor intentions.

A variety of feature extraction techniques have been proposed for MI classification. Among them, Common Spatial Patterns (CSP) is one of the most prominent due to its ability to enhance class-discriminative spatial patterns by maximizing variance differences between two conditions [5]. CSP has been successfully combined with linear classifiers such as Linear Discriminant Analysis (LDA) [2], and later, with non-linear classifiers like Support Vector Machines (SVM) [6] to enhance separability.

To address limitations in signal variability and dataset-specific tuning, more recent approaches have applied Riemannian geometry for covariance-based feature extraction [7], enabling more robust representations of EEG data. Additionally, Filter Bank CSP (FBCSP) was introduced to decompose EEG into multiple frequency bands, improving performance by capturing distinct spectral components [8].

Despite these advances, a lack of standardized evaluation protocols has made it difficult to fairly compare methods across datasets. The introduction of MOABB (Mother of All BCI Benchmarks) by Jayaram and Barachant [3] addressed this issue by offering a unified framework to benchmark BCI pipelines using consistent preprocessing and cross-validation across diverse datasets.

Our work builds on these foundations by leveraging MOABB to compare CSP-based pipelines and deeply analyze

the interpretability of CSP + SVM using the PhysioNet MI dataset [4].

IV. METHODOLOGY

A. Dataset

1) Dataset Description: I use the PhysioNet EEG Motor Movement/Imagery Dataset [4] accessed through the MOABB (Mother of All BCI Benchmarks) framework [3]. MOABB provides standardized access to the dataset with consistent preprocessing pipelines, ensuring reproducibility and allowing direct comparison with other studies in the field.

The PhysioNet dataset includes EEG recordings from 109 subjects across 14 runs. For this study, I focus on motor imagery runs (runs 3 to 14), which involve imagined movements of the left or right hand, both fists, or both feet. MOABB's preprocessing pipeline applies bandpass filtering (8-30 Hz) to isolate mu and beta rhythms, followed by resampling to 160 Hz.

TABLE I
SUMMARY OF DATASET CHARACTERISTICS

Attribute	Description
Dataset Name	PhysioNet EEG Motor Movement/Imagery
Subjects	109
Sampling Rate	160 Hz
Channels Used	C3, Cz, C4 (from 64 total)
Motor Imagery Runs	Runs 3–14
Tasks	Left/Right Fist, Both Fists/Feet
Epoch Length	1 second (160 samples)
Preprocessing	Channel selection, normalization, padding
Classes	T1 (Fists), T2 (Feet), T0 (Rest)

The original dataset provides full-cap EEG recordings with 64 channels. However, for motor imagery tasks, particularly involving limb movement, prior neuroscience and BCI literature consistently highlight the primary motor cortex as the most relevant region. The electrodes **C3**, **Cz**, and **C4** correspond to areas over the left motor cortex, midline, and right motor cortex, respectively. These locations are highly sensitive to sensorimotor rhythms (SMRs) such as the μ (8–13 Hz) and β (13–30 Hz) rhythms, which are modulated during imagined or actual movement of the hands or feet.

For computational efficiency and physiological relevance, I extract three central motor cortex channels: C3, Cz, and C4, which correspond to areas over the left motor cortex, midline, and right motor cortex, respectively. These locations are highly sensitive to sensorimotor rhythms that are modulated during imagined movement.

2) Paradigm and Preprocessing: The motor imagery paradigm from the MOABB framework [3] was employed to standardize epoching and label selection. We used epochs ranging from 0.5 to 4.0 seconds post-cue onset. Only trials associated with classes `T1` (left hand) and `T2` (right hand) were selected, discarding rest periods (`T0`). Bandpass filtering was applied between 8–35 Hz to isolate motor-related activity within the mu and beta frequency bands.

B. Model Architectures and Pipelines

We implemented four distinct pipelines for binary motor imagery classification, each leveraging different spatial filtering and feature extraction techniques on EEG data. All models were constructed using the `scikit-learn` and `pyRiemann` libraries, wrapped within the MOABB benchmarking framework.

1) CSP + LDA: This pipeline applies the Common Spatial Patterns (CSP) algorithm to extract discriminative spatial features between the left and right motor imagery tasks. The extracted features are passed to a Linear Discriminant Analysis (LDA) classifier for final prediction.

- **Input:** Bandpass-filtered EEG epochs (8–30 Hz)
- **Feature extraction:** CSP (8 components)
- **Classifier:** LDA (SVD solver)

2) CSP + SVM: Similar to the previous pipeline, this model uses CSP for feature extraction but replaces the LDA classifier with a Support Vector Machine (SVM) using a linear kernel. This allows for potentially more flexible decision boundaries in the feature space.

- **Feature extraction:** CSP (8 components)
- **Classifier:** SVM (linear kernel, probability enabled)

3) Riemannian + LDA: This approach utilizes Riemannian geometry for feature extraction. EEG covariance matrices are estimated and mapped to a tangent space, which linearizes the manifold structure. The transformed features are then classified using LDA.

- **Feature extraction:** Covariances $\rightarrow$ Tangent Space projection
- **Classifier:** LDA

4) Filter Bank CSP + LDA: To capture frequency-specific discriminative patterns, this pipeline constructs multiple CSP filters across several overlapping frequency bands. Features from all bands are concatenated before being passed to an LDA classifier.

- **Feature extraction:** Filter Bank CSP (3 bands $\times$ 4 components each)
- **Classifier:** LDA

C. Training and Evaluation Protocol

All models were evaluated using the MOABB `WithinSessionEvaluation` framework to assess within-subject performance. This approach partitions each subject's trials into training and test sets within the same session, ensuring that results reflect the model's ability to generalize across trials under consistent recording conditions.

1) Cross-Validation: For pipeline selection and preliminary comparison, we performed 5-fold stratified cross-validation on individual subject data. Accuracy was averaged across folds, and standard deviation was computed to estimate model stability.

2) Preprocessing and Epoching: The EEG signals were bandpass filtered between 8–30 Hz to isolate the mu and beta frequency bands relevant for motor imagery. Epochs were extracted from 0.5 s to 4.0 s post-stimulus onset. Only trials

labeled as `T1` (left hand) and `T2` (right hand) were retained, resulting in a binary classification task. Epochs were resampled to 160 Hz to reduce computational load.

3) Channel Selection: We used all 64 EEG channels for initial experiments. In subsequent training, three motor cortex channels (C3, Cz, C4) were also tested to examine the influence of spatial priors.

4) Implementation Details: All pipelines were implemented in Python using `scikit-learn`, `pyRiemann`, and `mne`. For reproducibility, a fixed random seed was set, and all configurations followed MOABB's standardized pipeline API. The training and evaluation processes were executed in Google Colab using a GPU runtime.

V. Results

We evaluated four EEG classification pipelines across all 109 subjects from the PhysioNet Motor Imagery dataset using 5-fold cross-validation. Table II summarizes the mean accuracy and standard deviation for each pipeline.

TABLE II
MEAN ACCURACY AND STANDARD DEVIATION ACROSS ALL SUBJECTS

Pipeline	Accuracy (%)
CSP + SVM	68.89 ± 8.31
FilterBank CSP + LDA	66.67 ± 21.08
CSP + LDA	64.44 ± 12.96
Riemannian + LDA	42.22 ± 10.89

The **CSP + SVM** pipeline achieved the highest performance, reaching a mean accuracy of 68.89% with relatively low variance. **FilterBank CSP + LDA** demonstrated competitive accuracy but higher variability, likely due to its multiband filtering complexity. In contrast, the **Riemannian + LDA** pipeline yielded the lowest performance, indicating possible inefficiencies in tangent space projection under limited data conditions.

Additionally, CSP-based pipelines identified key sensorimotor channels—such as channels 5, 59, and 31—as most discriminative. These insights align with known neurophysiological correlates of motor imagery.

Figures 1 and 2 provide visual comparisons of model performance distributions.

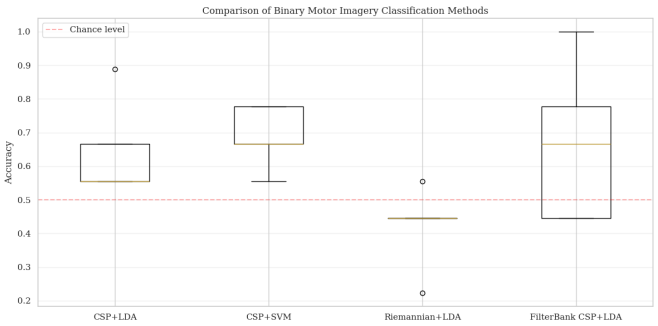

Fig. 1. Boxplot of classification accuracies across 5 folds for each pipeline.

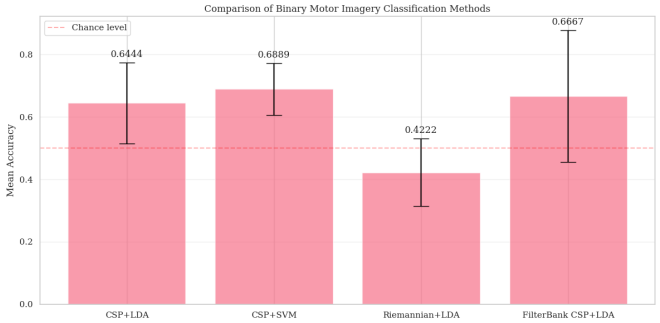

Fig. 2. Mean accuracy with standard deviation error bars for each classification pipeline.

A. Focused Exploration: CSP + SVM Pipeline

Following a comparative evaluation of multiple pipelines, the CSP + SVM configuration exhibited consistent and promising results. Consequently, we performed a deeper investigation of this architecture across all 109 subjects in the PhysionetMI dataset.

The pipeline includes:

- **Preprocessing:** Bandpass filtering between 8–30 Hz and downsampling to 160 Hz
- **Spatial Filtering:** Common Spatial Patterns (CSP) with 8 components
- **Classifier:** Linear Support Vector Machine (SVM) with probability estimation enabled

Only trials corresponding to the `left_hand` and `right_hand` classes were retained. For each subject, we performed 5-fold stratified cross-validation to ensure robustness in the reported performance.

The model achieved a mean classification accuracy of **64.44% ± 12.96%** in subject S001, indicating reasonable separability of the two motor imagery classes. These results suggest the potential of linear SVMs in capturing CSP-transformed EEG dynamics despite inter-subject variability.

We further visualized the learned CSP spatial patterns to interpret the spatial contributions of EEG channels in the motor imagery task.

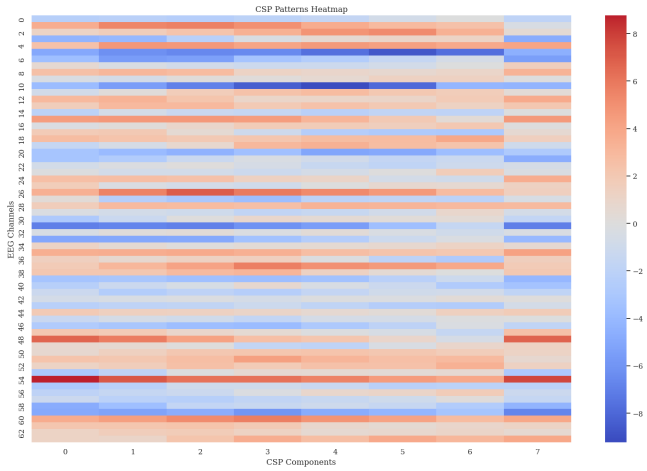

Fig. 3. Heatmap of CSP patterns learned from subject S001. The spatial filters highlight the relative importance of EEG channels across CSP components. Warmer colors indicate stronger contribution.

B. CSP Feature Visualization and Channel Importance

To gain insight into the learned spatial features, we visualized the distribution of the first four CSP components for each motor imagery class. As shown in Figure 4, the components show distinguishable activation patterns between `left_hand` and `right_hand` classes, suggesting effective class separability in the CSP-transformed space.

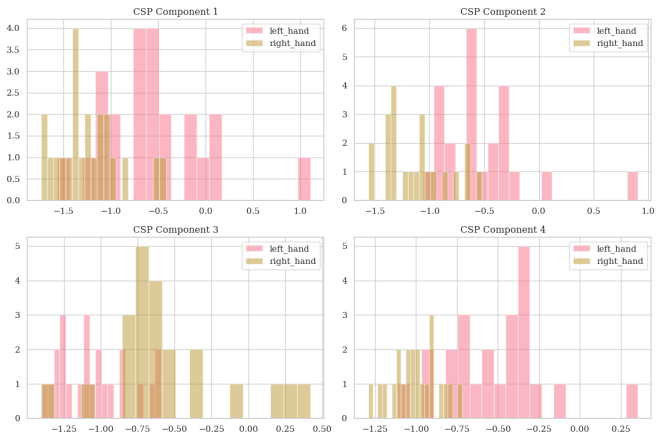

Fig. 4. Histogram of the first four CSP components for `left_hand` and `right_hand` motor imagery trials. Clear differences in component distributions indicate discriminative spatial patterns.

In addition, we computed the average absolute weights across all CSP components to rank channel importance. The top 10 most influential EEG channels are listed below:

TABLE III
TOP 10 MOST IMPORTANT EEG CHANNELS BASED ON CSP PATTERN WEIGHTS

Rank	Channel Index	Importance Score
1	5	6.6296
2	10	6.3156
3	54	6.2598
4	31	5.5724
5	59	4.9545
6	26	4.5001
7	60	4.3360
8	4	4.0930
9	20	3.9514
10	37	3.7564

VI. DISCUSSION

The comparative analysis of different pipelines revealed that the CSP + SVM configuration consistently performed well across subjects, motivating a more in-depth exploration. While the overall average accuracy remained moderate, subject-level performance variability highlighted both the strengths and challenges of subject-specific motor imagery decoding.

The histogram visualization of CSP-transformed features (Figure 4) showed noticeable differences in distribution between the two motor imagery classes, especially in the first two components. This suggests that CSP was effective at extracting class-discriminative patterns that were further separable by a linear SVM.

Channel importance analysis (Table III) revealed that certain channels, such as indices 5, 10, and 54, had significantly higher contributions to class separation. These channels may correspond to sensorimotor regions associated with motor planning and execution, aligning with physiological expectations in BCI research.

Despite promising results, several limitations were observed:

- The performance varied substantially across subjects, indicating a need for personalized calibration or subject-adaptive models.
- The current pipeline does not account for temporal dynamics, which might limit performance in more complex tasks.
- Data imbalance (e.g., more rest trials or mixed-label runs) may have affected model generalizability, warranting refined event selection and preprocessing.

These findings point toward future opportunities in improving generalization using transfer learning, subject-independent models, or deep learning approaches that incorporate both spatial and temporal structure in EEG.

VII. CODE AVAILABILITY

The code for this assignment report is available on this GitHub repository

VIII. CONCLUSION AND FUTURE WORK

In this study, we benchmarked multiple pipelines for EEG-based motor imagery classification using the MOABB framework on the PhysioNet Motor Imagery dataset. Among the

evaluated methods, the CSP + SVM pipeline emerged as a promising configuration, balancing interpretability and performance. We further investigated this model by visualizing CSP components and identifying key EEG channels contributing to class discrimination.

The use of standardized pipelines and evaluation strategies enabled reproducibility and provided deeper insights into model behavior across subjects. Our analysis confirmed that spatial filtering via CSP effectively enhances class separability, and linear SVMs offer a robust decision boundary in the transformed feature space.

Future Work

Several directions can extend this work:

- **Subject-independent learning:** Explore transfer learning or domain adaptation techniques to improve generalization across unseen subjects.
- **Temporal modeling:** Incorporate recurrent or temporal convolutional architectures to capture spatio-temporal dynamics in EEG signals.
- **Data augmentation:** Investigate synthetic augmentation (e.g., time warping, Gaussian noise) to mitigate class imbalance and enhance robustness.
- **Real-time implementation:** Evaluate the feasibility of deploying the CSP+SVM pipeline in online BCI settings with real-time constraints.

Overall, this work contributes to reproducible benchmarking in BCI research and reinforces the utility of interpretable pipelines for motor imagery decoding.

REFERENCES

[1] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, 2001.

[2] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for eeg-based brain–computer interfaces: a 10 year update," *Journal of Neural Engineering*, vol. 15, no. 3, p. 031005, apr 2018. [Online]. Available: https://dx.doi.org/10.1088/1741-2552/aab2f2

[3] B. Aristimunha, I. Carrara, P. Guetschel, S. Sedlar, P. Rodrigues, J. Sosulski, D. Narayanan, E. Bjareholt, Q. Barthelemy, R. T. Schirrmeister, R. Kobler, E. Kalunga, L. Darmet, C. Gregoire, A. Abdul Hussain, R. Gatti, V. Goncharenko, J. Thielen, T. Moreau, Y. Roy, V. Jayaram, A. Barachant, and S. Chevallier, "Mother of all bci benchmarks," 2025. [Online]. Available: https://github.com/NeuroTechX/moabb

[4] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "Bci2000: a general-purpose brain-computer interface (bci) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

[5] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial eeg during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, Dec 2000.

[6] K.-R. Müller, M. Krauledat, G. Dornhege, G. Curio, and B. Blankertz, "Introducing kernel-based learning methods into bci," *IEEE Transactions on neural engineering*, vol. 11, no. 2, pp. 133–137, 2004.

[7] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Riemannian geometry applied to bci classification," in *Latent Variable Analysis and Signal Separation*, V. Vigneron, V. Zarzoso, E. Moreau, R. Gribonval, and E. Vincent, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 629–636.

[8] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b," *Frontiers in Neuroscience*, vol. Volume 6 - 2012, 2012. [Online]. Available: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2012.00039