Riemannian fusions of EEG-based features for motor imagery detection under propofol sedation

Valérie Marissens Cueva, Camilla Mannino, Marie-Constance Corsi, Fabien Lotte, Sébastien Rimbert, Laurent Bougrain

RIEMANNIAN FUSIONS OF EEG-BASED FEATURES FOR MOTOR IMAGERY DETECTION UNDER PROPOFOL SEDATION

Valérie Marissens Cueva[⋆†] *Camilla Mannino*[‡] *Marie-Constance Corsi*[‡]
Fabien Lotte[⋆] *Sébastien Rimbert*[⋆] *Laurent Bougrain*[†‡]

[⋆] Inria Center at the University of Bordeaux / LaBRI, Talence, France
[†] Université de Lorraine, CNRS, LORIA, F-54000 Nancy, France
[‡] Sorbonne Université, ICM, CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, Paris, France

ABSTRACT

The brain is a complex system requiring multimodal approaches to better understand cognitive or motor functions. Thus, different and complementary electroencephalographic (EEG) neurophysiological features are available at various spatial, frequency, and temporal scales, e.g., brain connectivity, complexity, or entropy. However, they are usually not investigated all together. In this study, we combine and compare five EEG-based connectivity features with covariance matrices, defining five Riemannian fusion methods and three Euclidean ones as references. We do so for classifying motor imagery EEG signals, both in awake and sedated subjects, with the future goal of detecting accidental awareness during general anesthesia. Covariance matrices alone yielded the best accuracy, with and without sedation. Phase-based connectivity estimators appear to be the most promising fusion with covariances. No significant differences were found between the best fusion of features and that of classifiers.

***Index Terms*—** Fusion, connectivity, covariance, EEG, motor imagery, anesthesia

1. INTRODUCTION

Non-invasive brain-computer interfaces (BCIs) [1] measure brain activity - typically via electroencephalography (EEG) - and translate it into the user's feedback or control signal for applications in health, human-computer interaction or art, among others. One of the most used BCI paradigms is Motor Imagery (MI), i.e., movement imagination, which can be used for device control, motor rehabilitation [1] or, possibly for detecting Accidental Awareness during General Anesthesia (AAGA) [2]. In the latter case, the BCI goal is to detect MIs of patients under an anesthetic, who want to move due to an AAGA, but cannot do so due to the anesthetic cocktail.

Currently, the state-of-the-art method to detect MI in EEG-BCIs is Riemannian geometry [3, 4], which represents and manipulates EEG as covariance matrices. Standard Riemannian classifiers only use EEG covariance matrices capturing the EEG signal's variance/covariance. However, recent works have shown that such matrices could be complemented with functional connectivity matrices, i.e., matrices capturing statistical dependencies between EEG signals from different sensors [5], to improve EEG-based MI classification. However, this approach has only been explored on awake healthy BCI users, but never on patients under anesthetics, whereas anesthetics change EEG dynamics. Thus, any result obtained on awake users may not hold true on anesthetized patients. Moreover, there are many functional connectivity metrics for EEG, and many ways to fuse and combine them in a Riemannian setting with the classical covariance matrices. However, the best way to do so is still unknown.

Thus, in this paper, we design and compare five Riemannian fusion methods combining standard covariance with functional connectivity metrics, to detect MI in patients with and without propofol, an anesthetic agent used in surgery. This enables us to show that certain markers could be more relevant to detect MI according to the level of consciousness in clinical contexts, e.g., to prevent AAGA [2].

2. MATERIAL AND METHODS

2.1. Data description

22 right-handed healthy male subjects (27.38 ± 11.48 years) participated in this study, at the University Hospital of Nancy (France). As this experiment was originally designed to investigate the impact of anesthesia on EEG [2], only males were included to avoid blood pregnancy tests, but no gender differences have been proven in motor imagery EEG activity [6]. All participants signed an informed consent approved by the Inria ethical committee (COERLE 2016-011/01).

Subjects performed right-hand kinesthetic motor imagery (KMI) imagining a maximum of sensations from a clamp between the thumb and index fingers. Prior to the study, they were trained to perform the KMI task progressively [7]. A

The authors thank the French National Research Agency (ANR-22-CE19-0016; project BCI4IA).

979-8-3503-2411-2/25/$31.00 ©2025 IEEE

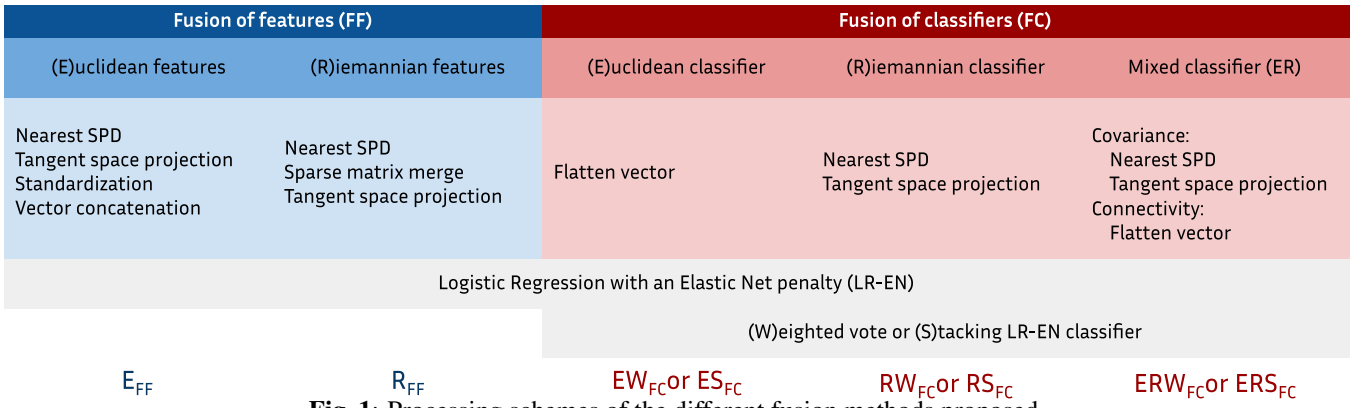

Fig. 1: Processing schemes of the different fusion methods proposed.

low-frequency beep indicated the start of a 2-second motor imagery period, ended by a second beep, and followed by a 6-second resting period with a 0-2 s random delay. All subjects performed 51 trials, while lying on a comfortable bed with eyes closed, and the right forearm on a cushion to prevent movement. This study included two phases: without sedation and under mild sedation with 1 μg/ml propofol, a dose at which subjects remained conscious.

EEG were recorded with OpenViBE with a Biosemi Active Two, from 128 sites (ABC layout) sampled at 2048 Hz. An external electromyogram (EMG) electrode was used to check that there was no movement during the MI task.

2.2. Methods

EEG signals were downsampled to 128 Hz and filtered within 8-30 Hz. Then, 1.5 seconds MI and rest epochs were extracted from each trial, respectively at [-2; -0.5] s and [0.5; 2] s, with 0 denoting the beginning of motor imagery. 40 electrodes were used for classification, covering mainly but not only the motor cortex: C17, B26, A23, D23, C10, B29, B24, B16, B5, A21, A8, D26, D21, D10, C32, C21, C12, C3, B31, B22, B18, B3, A32, A19, A5, A6, D28, D19, D12, D3, C25, C23, C2, B20, B2, A3, D16, D14, D2, A1, following the 128-channels Biosemi ABC layout. Covariance and functional connectivity measures were extracted on each epoch. These features were either combined into a single feature used by a classifier or used independently in an ensemble classification approach to distinguish MI from rest. A 5-fold cross-validation was applied, training on 40 trials and testing on the remaining 11. The following sections detail i) how to compute these features and ii) the strategies for feature-level and classifier-level fusions.

2.2.1. Features

Covariance

Riemannian classifiers employing covariance matrices are currently the state-of-the-art in EEG-based BCIs [4]. Let $\mathbf{X} \in \mathbb{R}^{e \times s}$ be an EEG trial, with e the number of electrodes and s the number of samples. The sample spatial covariance matrix $\Sigma \in \mathbb{R}^{e \times e}$ of $\mathbf{X}$ can be estimated as follows, with ^{T} the matrix transpose:

$$\Sigma = \frac{\mathbf{X}\mathbf{X}^T}{s-1}$$

Functional connectivity

Oscillatory activity plays a crucial part in the tasks performed by the subjects, each estimation of the functional connectivity between signals s_i and s_j relied on the power spectra (PSD). PSD were obtained by using multitapers [8], with time windows of length 1.5s and an overlap of 0.5s for each trial.

Spectral estimators - The coherence coefficient is the frequency equivalent to the time domain cross-correlation. It is defined as:

$$Coh_{ij}(f) = \frac{|E[S_{ij}(f)]|}{\sqrt{E[S_{ii}(f)] \cdot E[S_{jj}(f)]}}$$

where $S_{ij}(f)$ the cross-spectral density and $S_{ii}(f)$ the auto-spectral density. The imaginary part of coherence (ImCoh) is less sensitive to volume conduction effects [9].

Phase estimators - The Phase-Locking Value (PLV) and the Phase-Lag Index (PLI) assess phase synchrony between signals in a specific frequency band. The PLV is obtained from the absolute value of the mean phase between s_i and s_j, is defined as [10]:

$$PLV = |\mathbb{E}[e^{j\Delta\phi(t)}]|$$

where $\Delta\phi$ represents the phase difference between the signals.

The PLI evaluates the distribution of phase differences across observations by weighting them as -1 or 1 depending on their sign [11] and is defined as:

$$PLI = \mathbb{E}[\mathrm{sgn}(\Im(S_{ij}))]$$

where: $\Im(S_{ij})$ denotes the imaginary part of S_{ij}; sgn is the sign function. Given its sensitivity to field spread and noise, we considered a variant of the PLI: the weighted PLI (wPLI) [12]. Here, the different phase relationships are weighted based on their values and not on their sign.

2.2.2. Fusion

As covariance and connectivity matrices contain complementary information, we considered fusing them at different levels, namely at the features and classifier levels (see Fig. 1).

Fusion of features (FF)

1) E_{FF}: Euclidean fusion. In order to merge covariance and connectivity information, connectivity matrices were first approximated by their nearest semi-positive definite (SPD) matrices [13]. Each feature was then projected into its tangent space, standardized independently and concatenated into a single vector, following [14]. MI vs. rest classification was performed using a Logistic Regression with an Elastic-Net penalty (LR-EN), with the optimal L1 ratio selected via 5-fold cross-validation from $\{0.0, 0.25, 0.5, 0.75, 1.0\}$.

2) R_{FF}: Riemannian fusion. Features were merged into a block-diagonal sparse SPD matrix, with each block corresponding to a feature type. The resulting matrix was projected into the tangent space at the geometric mean of the training SPD matrices, according to [3], and LR-EN was used for classification as above.

Fusion of classifiers (FC)

The fusion of classifier approach integrates covariance and/or functional connectivity estimators within a two-level classification framework. All three variants use LR-EN as the first-level classifier. The first-level classifier outputs are then fused in different ways. The EN combines ℓ_1 and ℓ_2 penalties and optimizes the weight vector $\mathbf{w}$ by minimizing the following objective function:

$$\min_{\mathbf{w}} \|\text{vec}(X)\mathbf{w} - \mathbf{y}\|^2 + \alpha\rho\|\mathbf{w}\|_1 + \alpha(1-\rho)\|\mathbf{w}\|_2^2 \quad (1)$$

where the regularization parameters were set to $\alpha = 1$ and $\rho = 0.15$ based on preliminary experiments.

First-level classifiers:

1) E_{FC}: Euclidean Classifiers. Covariance and functional connectivity features are vectorized by flattening the upper triangular part of each matrix. A first-level classifier is trained for each estimator.

2) R_{FC}: Riemannian Classifiers. The nearest SPD matrices of the covariance and functional connectivity features is computed. Then, they are projected in their own tangent space. A specific first-level classifier is trained on each projected estimator. It is inspired by the method proposed in [5].

3) ER_{FC}: Fusion of Riemannian and Euclidean Classifiers. This variant merges features from two domains: covariance matrices processed in Riemannian space and functional connectivity estimators computed in Euclidean space. A first-level classifier is trained for each estimator. It is inspired by the method proposed in [5].

Second-level classifier:

A) W: Weighted vote. The predictions of the first-level classifiers are combined using weights derived from confidence intervals. These weights are estimated on the validation set by calculating each estimator's average accuracy across multiple trials.

B) S: Stacking Ensemble. Predictions of first-level classifiers are combined using a second-level stacking classifier, also based on LR-EN.

Six possible fusions of classifiers can be obtained by combining the three kinds of classifiers and the two ensemble methods: EW_{FC}, RW_{FC}, ERW_{FC}, ES_{FC}, RS_{FC}, and ERS_{FC} (see Fig. 1).

3. RESULTS

This section presents MI detection accuracies and statistical differences between the proposed fusion methods using connectivity estimators and covariance. Figure 2 shows the results of both feature fusion strategies, with and without sedation. Due to space limitations, Figure 4 presents only the most relevant classifier fusion accuracies, while all six strategies are described in the text.

3.1. Accuracy

To rank the discriminative power of each feature set and thus identify the most useful ones, we applied the Bradley-Terry model [15] using pairwise comparisons across the E_{FF}, R_{FF}, RS_{FC}, RW_{FC}, ERS_{FC}, and ERW_{FC} fusion methods (Fig. 3). ES_{FC} and EW_{FC} were excluded, as Euclidean-space covariance matrices yielded accuracies below chance level, indicating no discriminative value.

3.1.1. Single features

According to the Bradley-Terry, covariance matrices exhibit the highest discriminative power, outperforming coherence and phase-based features, both with and without sedation. Indeed, covariance yields an accuracy of 69.72% (no sedation) and 66.16% (sedation) using a Riemannian classifier, compared to 62.17% and 61.23% for vectorized coherence features. All covariance-based models significantly outperform coherence-based ones ($p < 0.01$).

3.1.2. Fusion of features

In both sedated and non-sedated conditions, fusing features in a large sparse matrix (R_{FF}) yields significantly lower accuracy than fusing them at the tangent space (E_{FF}). For example, without sedation, the best R_{FF} accuracy of 58.60% is achieved by fusing covariances with coherence, while the worst E_{FF} result of 65.83% comes from covariances fused with both wPLI and PLI. A similar gap is observed with sedation (Fig. 2).

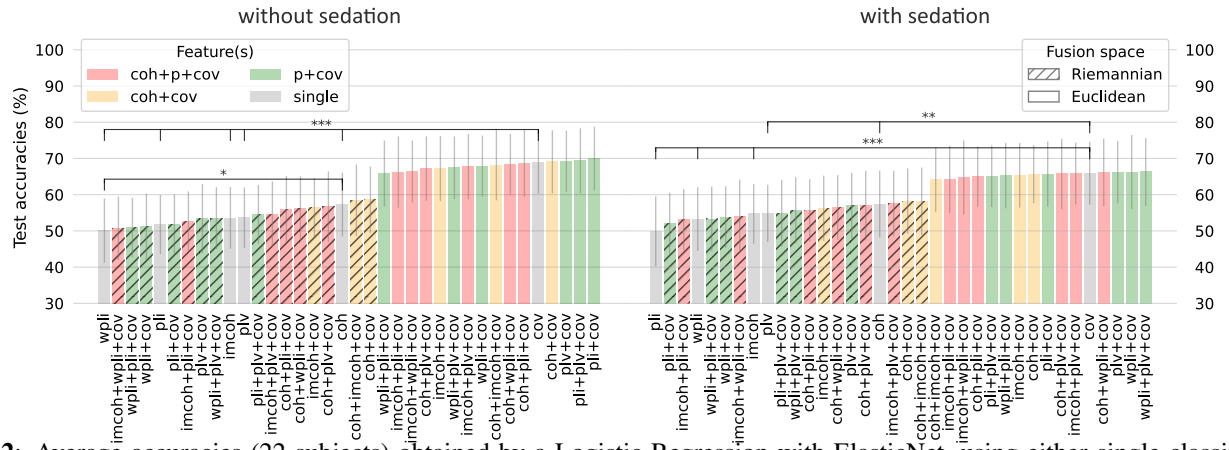

Fig. 2: Average accuracies (22 subjects) obtained by a Logistic Regression with ElasticNet, using either single classifier or fusion classifiers based on covariance (*cov*), coherence (*coh*), or phase (*p*). Riemannian fusion bars are hatched; Euclidean fusion bars are solid. Paired t-tests with Bonferroni correction (6 comparisons) evaluated differences between single features and the best fusion set (e.g., $pli + cov$ without sedation and $wpli + plv + cov$ with sedation) (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$).

3.1.3. *Fusion of classifiers*

Projecting estimator matrices onto the tangent space (Riemannian approaches) outperforms simple vectorization (Euclidean approaches) in both stacking and weighted ensemble methods. However, as illustrated in Figure 4, when both approaches are combined, performance improves across all feature sets and ensemble methods. For example, the best mixed result is achieved in the weighted ERW_{FC} method using classifiers trained on covariance and PLV (69.74%), while the best purely Riemannian result is obtained with RS_{FC} using covariance and imaginary coherence (68.41%).

It is important to note that there is no statistical difference between the top-performing models, preventing the identification of a single best fusion strategy. These results also hold under sedation, with the best performance obtained using ERW_{FC} fusion and classifiers trained on covariance, PLV, and wPLI (66.36%).

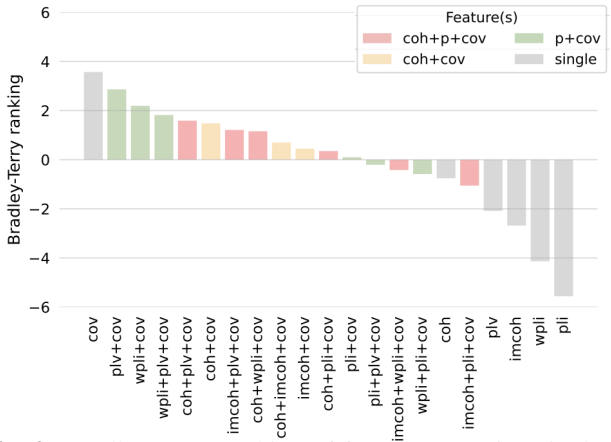

Fig. 3: Bradley-Terry ranking of feature sets using single or fused features based on covariance (*cov*), coherence (*coh*), or phase (*p*), across the fusion methods: E_{FF}, R_{FF}, RS_{FC}, RW_{FC}, ERS_{FC}, and ERW_{FC}, with sedation. Results without sedation are discussed, but not shown.

3.1.4. *Fusion of features vs. fusion of classifiers*

Taking into account both fusion of features and classifiers (Fig. 3), the fusion of PLV and covariance achieves the second-best MI vs. rest discrimination behind covariance, in both sedated and non-sedated conditions. Notably, without sedation, fusions involving covariance and phase estimators fail to discriminate, but their performance improves with sedation. With no significant difference, the best fusion of features E_{FF} yields a better accuracy than the best fusion of classifiers ERW_{FC}, respectively 70.01% against 69.82% (Fig. 4 and 2).

3.2. Effect of sedation on performance

Under sedation, the performance gap between top models and single-feature models narrows in both feature and classifier fusion, as overall accuracies decline. Specifically, top performance drops from 69.24% to 66.28% on E_{FF}, and from 69.82% to 66.36% on ERW_{FC}. All accuracies based on single features, but covariance, are stable with or without sedation.

4. DISCUSSION

4.1. Limited contribution of connectivity features in motor imagery classification

Our results indicate that connectivity-based features do not significantly improve classification accuracy in our MI vs.

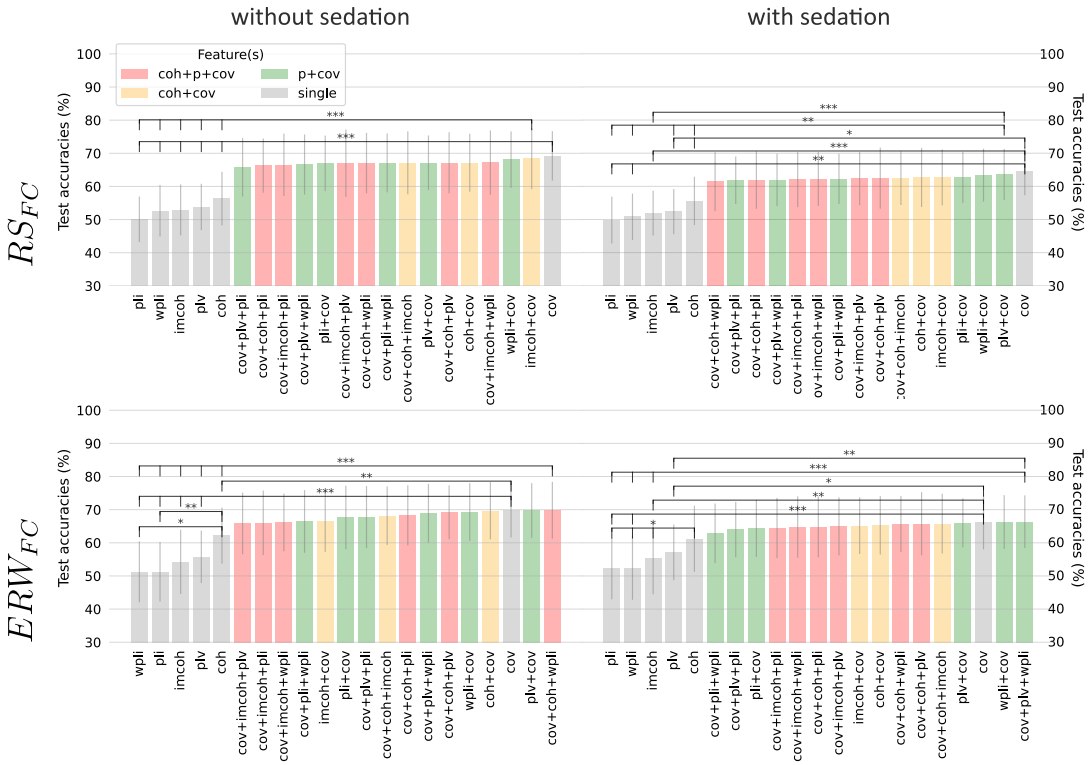

Fig. 4: Mean accuracies (22 subjects) obtained by the stacking ensemble, based on either single or fused classifiers derived from covariance (cov), coherence (coh), or phase (p), with and without sedation. Top: first-level classifiers in Riemannian space (RS_{FC}), bottom: weighted voting with mixed classifiers (ERW_{FC}). Paired t-tests with Bonferroni correction were used to compare accuracies of single features vs. best fusion (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$).

Rest paradigm. This finding contrasts with previous studies, such as [5], which reported enhanced MI decoding using connectivity measures in a right vs. left hand MI classification task. Several factors may account for this divergence. First, the nature of the MI task differs: our study involves a fine-grained kinesthetic motor imagery task—pinching between thumb and index finger whereas other studies used whole-hand imagery with opposing laterality. Connectivity features may be more sensitive to inter-hemispheric differences, which are prominent in right vs. left hand comparisons, but less informative when contrasting unilateral MI with rest. Additionally, resting-state activity may introduce patterns of connectivity that obscure the subtle network-level changes associated with fine-motor MI. This suggests that the discriminative power of connectivity features is highly dependent on the MI paradigm and the complexity of the motor task. One hypothesis to explain why phase estimators do not strongly contribute to the prediction could be that they are difficult to estimate i) in a short 1.5 s window size, ii) when amplitudes are small. Additional studies not presented here for lack of space have shown that increasing the window size to 3 s to compute phase estimators increases the accuracies.

4.2. Features fusion under propofol sedation

Our results show that phase-based features, when fused with covariance matrices, have greater discriminative power for distinguishing MI from rest under light sedation than in the awake state. In [16] it was shown that, during light propofol sedation, wPLI exhibits globally reduced connectivity except in occipital regions, compared to full consciousness, where connectivity is stronger in occipital, parietal, and temporal areas. Future work will assess if it could impact MI detection.

Duclos et al. found that amplitude envelope correlation (AEC) outperformed wPLI and their combination in tracking functional connectivity changes across levels of consciousness [16]. They concluded that AEC is more robust and reproducible than phase-based measures such as wPLI, PLI, or the imaginary part of coherence [17]. It would therefore be valuable in future studies to assess whether AEC can provide complementary or superior information for MI detection under sedation.

Röhr et al. integrated propofol-specific EEG features, such as burst suppression ratio, with covariances in a Riemannian framework to detect postoperative delirium [18]. While burst suppression reflects a state of deep anesthesia incompatible with cognitive activity, including motor imagery or movement intention, its detection could allow the exclu-

sion of EEG periods where MI is physiologically implausible, thereby preventing false positives during deep anesthesia.

4.3. Individual-level analysis

This study evaluated the ability to detect motor imagery from rest using functional connectivity, covariance features, or their fusion. However, analyses were conducted at the group level rather than individually. Future work should assess whether specific features or fusion strategies offer greater advantages for certain subjects. Additionally, separating participants by BCI performances could help identify features more beneficial for particular groups, e.g., the least responsive subjects.

5. CONCLUSION

Our study proposed and evaluated eight techniques for combining six EEG features types at both feature and classifier levels to decode motor imagery under propofol sedation. Covariance-based features alone enabled reliable MI detection in both sedated and non-sedated states, while adding connectivity features did not yield significant improvements. Future work will explore longer time windows and higher sampling resolution to improve the reliability of connectivity-based features.

6. REFERENCES

[1] M Clerc, L Bougrain, and F Lotte, *Brain-computer interfaces 1: Methods and perspectives*, John Wiley & Son, 2016.

[2] S. Rimbert, D. Schmartz, L. Bougrain, C. Meistelman, C. Baumann, and P. Guerci, "Motana: study protocol to investigate motor cerebral activity during a propofol sedation," *Trial*, vol. 20, no. 534, pp. 9, 2019.

[3] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Multiclass Brain-Computer Interface Classification by Riemannian Geometry," *IEEE Trans. on Biomed. Eng.*, vol. 59, no. 4, pp. 920–928, 2012.

[4] S Chevallier, I Carrara, B Aristimunha, P Guetschel, S Sedlar, B Junqueira Lopes, et al., "The largest EEG-based BCI reproducibility study for open science: the MOABB benchmark," preprint, 2024.

[5] M.-C. Corsi, S. Chevallier, F. De Vico Fallani, and F. Yger, "Functional Connectivity Ensemble Method to Enhance BCI Performance (FUCONE)," *IEEE Trans Biomed Eng*, vol. 69, no. 9, pp. 2674–2683, 2022.

[6] V. Gamboa von Groll, N. Leeuwis, S. Rimbert, A. Roc, L. Pillette, F. Lotte, et al., "Large scale investigation of the effect of gender on mu rhythm suppression in motor imagery brain-computer interfaces," *Brain-Computer Interfaces*, vol. 11, no. 3, pp. 87–97, 2024.

[7] S Rimbert, L Bougrain, and S Fleck, "Learning how to generate kinesthetic motor imagery using a BCI-based learning environment: a comparative study based on guided or trial-and-error approaches," in *IEEE SMC*, 2020, pp. 2483–2498.

[8] D. Slepian, "Prolate spheroidal wave functions, fourier analysis, and uncertainty #x2014; V: the discrete case," *Bell Syst. Tech. J.*, vol. 57, no. 5, pp. 1371–1430, 1978.

[9] G. Nolte, O. Bai, L. Wheaton, Z. Mari, S. Vorbach, and M. Hallett, "Identifying true brain interaction from EEG data using the imaginary part of coherency," *Clin Neurophysiol*, vol. 115, no. 10, pp. 2292–2307, 2004.

[10] P. Celka, "Statistical Analysis of the Phase-Locking Value," *IEEE Sig Proc Lett.*, vol. 14, no. 9, 2007.

[11] C. J. Stam, G. Nolte, and A. Daffertshofer, "Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources," *Hum. Brain Mapp.*, vol. 28, no. 11, 2007.

[12] M. Vinck, R. Oostenveld, M. van Wingerden, F. Battaglia, and C. M. A. Pennartz, "An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias," *NeuroImage*, vol. 55, no. 4, 2011.

[13] "Computing a nearest symmetric positive semidefinite matrix," *Lin Alg App*, vol. 103.

[14] M. S. Yamamoto, A. Mellot, S. Chevallier, and F. Lotte, "Novel SPD matrix representations considering cross-frequency coupling for EEG classification using Riemannian geometry," in *EUSIPCO*, 2023.

[15] R. A. Bradley and M. E. Terry, ," *Biometrika*, , no. 3/4.

[16] Catherine Duclos, Charlotte Maschke, Yacine Mahdid, Kathleen Berkun, Jason da Silva Castanheira, Vijay Tarnal, et al., "Differential classification of states of consciousness using envelope- and phase-based functional connectivity," *NeuroImage*, vol. 237, pp. 118171, 2021.

[17] G.L. Colclough, M.W. Woolrich, P.K. Tewarie, M.J. Brookes, A.J. Quinn, and S.M. Smith, "How reliable are meg resting-state connectivity metrics?," *NeuroImage*, vol. 138, pp. 284–293, 2016.

[18] Vera Röhr, Benjamin Blankertz, Finn M. Radtke, Claudia Spies, and Susanne Koch, "Machine-learning model predicting postoperative delirium in older patients using intraoperative frontal electroencephalographic signatures," *Front Aging Neurosci*, vol. 14, 2022.