MFMSRNet: An Interpretable Multi-frequency and Multi-scale Riemannian Network for Motor Imagery Decoding

Wenhao Rao[1], Xujie Zhao[2], Jianhui Zhao[1*], Bo Du[1], Feixiang Tang[3,4]
[1]School of Computer Science, Wuhan University
[2]School of Computation, Information and Technology, Technical University of Munich
[3]School of Integrated Circuits, Wuhan University
[4]School of Integrated Circuits, Hefei University of Technology
{wenhaorao, jianhuizhao, dubo}@whu.edu.cn, xujie.zhao@tum.de, thomas0209@whu.edu.cn

Abstract

Motor imagery (MI) electroencephalography (EEG) decoding has benefited from deep learning, yet methods operate in Euclidean space and behave as opaque black boxes, neglecting the intrinsic geometry of functional brain connectivity. EEG connectivity descriptors, such as phase synchrony and covariance matrices, naturally reside on the manifold of symmetric positive definite (SPD) matrices, where Euclidean operations are geometrically inconsistent and hinder interpretability. This work proposes the Multi-frequency and Multi-scale Riemannian Network (MFMSRNet), an interpretable end-to-end geometry-aware framework for MI EEG decoding on the SPD manifold. The method constructs kernelized phase-locking value (KPLV) functional connectivity (FC) matrices to capture nonlinear phase synchrony while ensuring positive definiteness. An attention-based Riemannian fusion mechanism adaptively integrates information across multiple frequency bands in the tangent space. Furthermore, a multi-scale Riemannian network extracts global, hemispheric, and local connectivity patterns via manifold-preserving bilinear mappings and smooth eigenvalue rectification. Extensive experiments indicate that MFMSRNet yields more expressive and interpretable representations for robust MI decoding, offering a promising solution for reliable brain–computer interface applications. The code is available at https://github.com/Raeno-Rao/MFMSRNet.

1 Introduction

Brain computer interface (BCI) based on electroencephalogram (EEG) has achieved notable progress in motor imagery (MI) decoding, driven largely by advances in deep learning architectures [Ajmeria *et al.*, 2022; Altaheri *et al.*, 2023]. Convolutional and attention-based neural networks enable automatic feature extraction from raw EEG signals and have demonstrated competitive classification performance across benchmark datasets [Zhang and Li, 2025]. However, the black-box nature of these models poses a significant challenge for practical deployment, as they provide limited transparency into the learned features and underlying decision processes [Rudin and Cynthia, 2019]. This lack of interpretability not only hinders neuroscientific insight but also limits reliability in safety-critical applications, where understanding the mapping between brain dynamics and model outputs is essential for clinical trust [Xu and Yang, 2025].

Beyond interpretability, a fundamental limitation arises from the geometric mismatch between EEG representations and the learning space [Yger *et al.*, 2017]. Multi-channel EEG signals exhibit rich inter-channel dependencies, often represented through covariance or functional connectivity (FC) matrices [Varoquaux *et al.*, 2010; Dai *et al.*, 2019]. Standard neural architectures treat multi-channel EEG as vectors or tensors in Euclidean space, implicitly assuming linearity and disregarding the true geometry of connectivity representations [He and Wu, 2020]. Nevertheless, covariance descriptors, correlation patterns, and FC matrices reside on non-Euclidean manifolds, specifically the manifold of symmetric positive definite (SPD) matrices, where conventional Euclidean operations lack geometric fidelity. Prior studies have shown that Riemannian geometry can be leveraged to better preserve the intrinsic structure of EEG features [Tibermacine *et al.*, 2024; Paillard *et al.*, 2025]. For example, Riemannian geometry–based spatial filtering can enhance discriminative power while maintaining robustness to noise compared to Euclidean counterparts, highlighting the importance of respecting geometric structure in EEG classification tasks [Pan *et al.*, 2025].

Despite the advantages of Riemannian geometry in EEG representation learning, existing manifold-based methods for MI decoding have several key limitations [Bronstein *et al.*, 2017; Gerken *et al.*, 2023]. First, most methods focus on covariance or single-band connectivity, ignoring the multi-frequency nature of MI-related neural dynamics [Wen *et al.*, 2024]. Different frequency bands contribute unequally to MI discrimination, especially the mu (8–13 Hz) and beta (13–30 Hz) bands, where event-related desynchronization (ERD) is most prominent and informative. Current multi-frequency fusion strategies often use simple concatenation or Euclidean attention, which can distort the intrinsic geometry of SPD matrices [Suh and Hyung, 2021; Dissanayake *et al.*, 2022]. Second, prior Riemannian networks typically model functional connectivity at a single (whole-brain) spatial scale, failing to capture the hemispheric lateralization and localized sensorimotor interactions critical for MI tasks [Tanamachi *et al.*, 2023]. Although recent geometric deep learning models introduce manifold-aware mappings or attention mechanisms, they rarely integrate multi-scale spatial structure within a unified framework [Shi *et al.*, 2025; Fei *et al.*, 2025]. Finally, while manifold representations are often regarded as more interpretable, most existing works provide limited neurophysiological validation beyond performance improvements, leaving the relationship between learned geometric features and established brain network properties largely unexplored [Dan *et al.*, 2022; Atz *et al.*, 2021].

To address these challenges, this work introduces Multi-frequency and Multi-scale Riemannian Network (MFMSRNet), an end-to-end geometry-aware framework that integrates kernelized FC, Riemannian attention-based fusion, and multi-scale geometric representations for MI decoding. By constructing FC matrices on the SPD manifold and fusing multi-frequency information through attentive geometric operations, the proposed method not only achieves competitive performance across intra-subject and inter-subject benchmarks but also yields representations that are more interpretable in terms of underlying neural interactions. Our contributions can be summarized as follows:

- We propose **MFMSRNet**, an end-to-end Riemannian deep learning framework for MI EEG decoding that jointly models FC, multi-frequency fusion, and multi-scale spatial structure on the manifold.
- We introduce a **kernelized phase-synchrony FC** representation that captures non-linear inter-channel dependencies while ensuring SPD, enabling principled Riemannian analysis of EEG connectivity.
- We develop an **attention-based Riemannian fusion mechanism** that adaptively integrates multi-frequency connectivity representations directly on the manifold, preserving the intrinsic geometric structure to support multi-scale spatial learning.
- Beyond improved classification performance, we conduct **systematic interpretability analyses**, including graph-theoretic small-world topology evaluation and frequency- and region-aware attention profiling, revealing neurophysiologically meaningful patterns underlying the learned representations.

2 Related Work

MI EEG decoding has been extensively studied using deep learning methods due to their strong capacity for automatic feature extraction from raw or minimally processed signals. Convolutional neural networks (CNNs) have been widely adopted to model temporal and spatial EEG patterns, with compact architectures such as EEGNet demonstrating competitive performance across multiple BCI benchmarks while maintaining low computational complexity [Lawhern *et al.*, 2018]. Subsequent works extend this paradigm by introducing multi-scale convolutions, frequency-aware branches, and attention mechanisms to better capture spectral variability and inter-channel dependencies [Mane *et al.*, 2020; Song *et al.*, 2022]. Transformer-based architectures further enhance long-range dependency modeling and provide visualization capabilities through attention maps [Zhang and Li, 2025; Fu *et al.*, 2025; Zhang *et al.*, 2023]. Despite these advances, most deep learning approaches represent multi-channel EEG signals in Euclidean space, implicitly assuming linear relationships between channels [Zhang and Etemad, 2023]. Such representations are insufficient for modeling structured inter-channel interactions and often obscure the underlying FC patterns critical to MI decoding. These limitations motivate the development of connectivity-aware and geometry-preserving frameworks, as adopted in MFMSRNet, which explicitly models EEG interactions through FC on non-Euclidean manifolds.

To address the geometric mismatch between EEG connectivity representations and Euclidean learning spaces, Riemannian geometry–based methods have been proposed to operate directly on SPD matrices. Early studies demonstrated that covariance-based features modeled on SPD manifolds yield improved robustness and discriminability for MI classification [Yger *et al.*, 2017; Kalaganis *et al.*, 2022]. More recent approaches integrate Riemannian geometry into deep architectures via bilinear mappings, eigenvalue rectification, and Log-Euclidean projections, enabling end-to-end learning of manifold-valued representations [Suh and Hyung, 2021; Dissanayake *et al.*, 2022]. Extensions of this paradigm include tensor-based Riemannian networks and SPD-valued graph models that incorporate time–frequency information to capture localized neural dynamics [Ju and Guan, 2022; Ju and Guan, 2024]. However, most existing Riemannian methods focus on covariance descriptors or single-band representations, and multi-frequency information is often fused through handcrafted strategies or Euclidean operations, potentially compromising the intrinsic manifold structure. In contrast, the proposed MFMSRNet performs attention-guided fusion of multi-frequency FC directly on the Riemannian manifold, preserving geometric consistency while enabling adaptive spectral integration.

FC analysis provides a principled way to characterize inter-regional neural interactions and has been increasingly combined with attention mechanisms to enhance MI decoding. Attention-based CNNs and feature selection networks emphasize informative frequency bands or spatial regions, leading to improved classification performance [Altaheri *et al.*, 2022; Qin *et al.*, 2024]. More recent works introduce self-attention mechanisms on correlation or covariance manifolds, enabling adaptive weighting of connectivity patterns under Riemannian constraints [Wang *et al.*, 2024; Hu *et al.*, 2025]. Nevertheless, most existing attention-based FC methods focus on either frequency-level or region-level selection in isolation, and few consider connectivity patterns across multiple spatial scales within a unified framework. Moreover, although manifold-based representations are often claimed to be more interpretable, systematic validation against established neurophysiological network properties remains limited. MFMSRNet addresses these limitations by jointly modeling global, hemispheric, and local connectivity structures on the SPD manifold, while providing quantitative interpretability analyses based on graph-theoretic measures and attention-based frequency and region profiling.

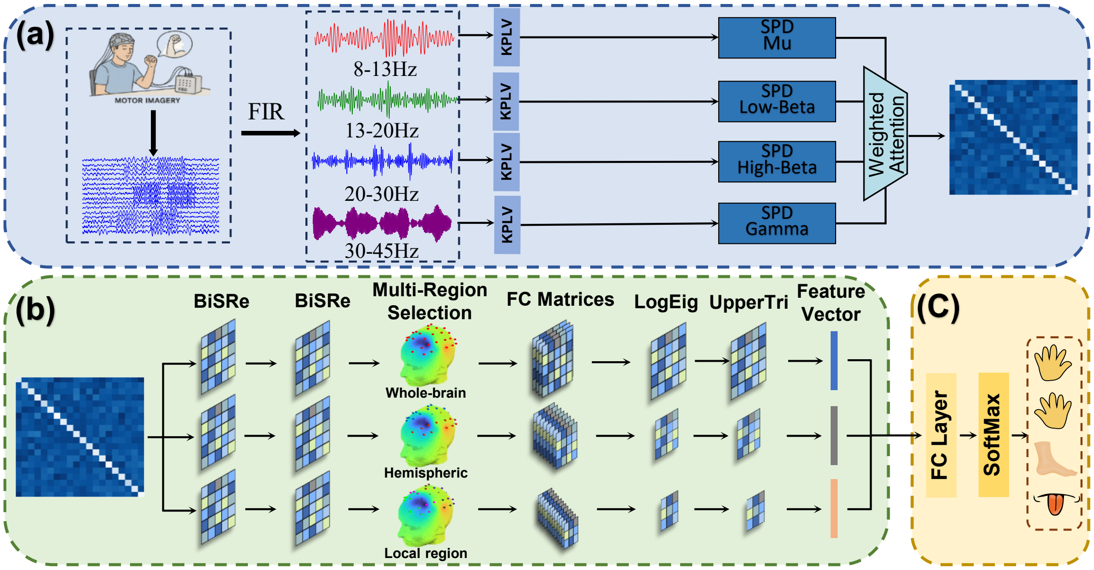

Figure 1: Overview of the proposed end-to-end MI decoding framework MFMSRNet. (a) Multi-frequency brain connectivity construction and fusion, (b) Multi-scale connectivity modeling, and (c) Fully connected layer and softmax layer for final classification.

3 Proposed Method

This work proposes a unified end-to-end framework for decoding MI EEG based on multi-frequency FC matrices construction, Riemannian geometry–aware feature learning, attention-guided frequency fusion, and multi-scale connectivity modeling. The entire pipeline (Figure 1) preserves the intrinsic geometry of FC matrices—modeled as SPD matrices—while enabling adaptive and discriminative representation across frequency bands.

3.1 Kernelized Phase-Synchrony FC Matrices

Given multi-channel EEG signals $\mathbf{X} \in \mathbb{R}^{N\times C\times T}$, where N denotes the number of trials, C the number of channels, and T the number of time points, band-pass filtering is first applied to extract frequency-specific components.

To characterize functional interactions between EEG channels, we adopt phase-based connectivity modeling. Phase synchronization has been shown to be particularly suitable for MI EEG analysis, as it is less sensitive to amplitude variability and volume conduction effects, while effectively capturing oscillatory coupling underlying MI processes [Liu *et al.*, 2016]. Accordingly, for each frequency band X_n, analytic signals are obtained via the Hilbert transform, and the instantaneous phase $\phi_c(t)$ is extracted for each channel.

The phase-locking value (PLV) P_{ij} between channels i and j is computed as:

$$P_{ij} = \left| \frac{1}{T} \sum_{t=1}^{T} \exp\left(\mathrm{i}\left(\phi_i(t) - \phi_j(t)\right)\right) \right|, \tag{1}$$

where i represents the imaginary unit, $P_{ij} \in [0, 1]$, with larger values indicating stronger phase synchronization.

Although PLV provides a robust measure of phase coupling, it yields a bounded similarity that does not directly conform to the geometric structure required by Riemannian learning. To address this limitation, we propose a Kernelized PLV (KPLV) representation, which maps PLV-based relationships into the space of SPD matrices.

Specifically, PLV is first converted into a distance measure:

$$D_{ij} = 1 - P_{ij}. \tag{2}$$

A Gaussian kernel is then applied to obtain a similarity matrix:

$$K_{ij} = \exp\left(-\frac{D_{ij}^2}{2\sigma^2}\right), \tag{3}$$

where σ controls the decay rate of similarity, which is empirically determined by grid search on the verification set. This kernel operation introduces nonlinearity and ensures that the resulting FC matrix is positive definite, thereby enabling principled Riemannian geometry–based analysis. Compared with raw PLV or linear correlation, KPLV captures higher-order dependencies while explicitly preserving the manifold structure of FC.

To enhance numerical stability, we apply a regularization term:

$$K_{\text{reg}} = K + \lambda\mathbf{I}, \tag{4}$$

where λ is a small positive constant. The matrix is then trace-normalized:

$$K_{\text{norm}} = \frac{K_{\text{reg}}}{\mathrm{tr}(K_{\text{reg}})}, \tag{5}$$

and symmetrized to eliminate numerical asymmetry introduced by floating-point computation:

$$K_{\text{final}} = \frac{1}{2}(K_{\text{norm}} + K_{\text{norm}}^{\top}). \tag{6}$$

Each K_{final} corresponds to the kernelized FC of a specific frequency band. Let $F_n = K_{\text{final}}^{(n)}$ denote the SPD matrix for the n-th band. The set $\{F_n\}_{n=1}^{B}$ then serves as input to the subsequent Riemannian feature learning and attention-based fusion modules, where B is the number of frequency bands.

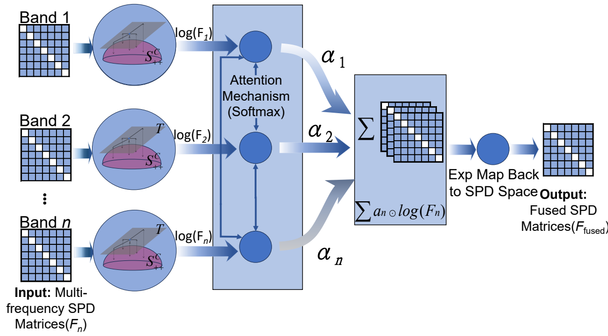

Figure 2: Attention-based multi-frequency FC fusion.

3.2 Attention-Based Multi-frequency FC Matrices Fusion

Multi-frequency fusion is essential for MI decoding because different frequency ranges contribute unequally to motor execution and imagery, particularly in the mu and beta bands. However, directly fusing multi-branch SPD representations is non-trivial, as naive concatenation or Euclidean averaging may violate the underlying manifold geometry and fail to capture frequency-specific relevance. To address this, we design an SPD-constrained attention-based fusion mechanism operating in the Log-Euclidean domain, as illustrated in Figure 2.

To fuse SPD matrices, each encoded matrix is projected to the tangent space via the matrix logarithm:

$$L_n = \log\left(F_n\right). \tag{7}$$

This mapping converts the non-Euclidean SPD manifold into a Euclidean vector space where linear operations are valid. To obtain a compact representation for attention learning, we extract a band-level embedding from each L_n by concatenating its diagonal elements and row-wise means, denoted as z_n.

An attention scoring network $g(\cdot)$, implemented as a lightweight MLP (or alternatively a multi-head self-attention module), is then applied to compute the importance score of each frequency band:

$$s_n = g(z_n). \tag{8}$$

The attention weights are obtained via a softmax normalization:

$$\alpha_n = \frac{\exp(s_n)}{\sum_{j=1}^{B}\exp(s_j)}, \tag{9}$$

which satisfies $\alpha_n \geq 0$ and $\sum_{n=1}^{B}\alpha_n = 1$.

This yields interpretable band relevance, highlighting which frequency bands contribute most to MI classification. We then fuse multi-frequency FC in the Log-Euclidean domain:

$$L_{\text{fused}} = \sum_{n=1}^{B}\alpha_n L_n. \tag{10}$$

Finally, the fused SPD matrix is reconstructed via the matrix exponential:

$$F_{\text{fused}} = \exp\left(L_{\text{fused}}\right). \tag{11}$$

This guarantees that the fused output remains a valid SPD matrix, enabling further processing in Riemannian layers.

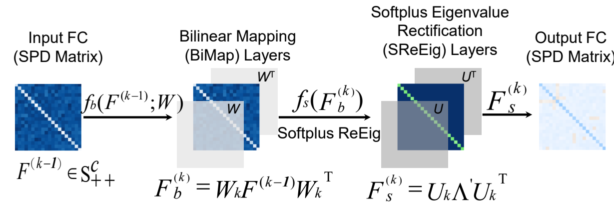

Figure 3: The architecture of the BiSRe module.

3.3 Bilinear and Rectification Matrix Mapping Module

The FC matrices extracted by the proposed KPLV module are SPD matrices, which lie on a connected Riemannian manifold denoted by $\mathcal{S}_{++}^{C}$ [Tuzel *et al.*, 2007]. Directly applying conventional Euclidean operations to such matrices may distort their intrinsic geometry. To enable discriminative learning while preserving manifold structure, we introduce a Bilinear Softplus Rectification (BiSRe) module for SPD-valued feature transformation. The fused FC matrix is treated as an SPD-valued neural network input and processed through two stacked BiSRe layers to perform nonlinear geometric transformations directly on the SPD manifold.

As illustrated in Figure 3, each BiSRe module consists of a Bilinear Mapping (BiMap) layer followed by a SoftPlus Eigenvalue Rectification (ReEig) layer. Let $F^{(k-1)} \in \mathcal{S}_{++}^{C}$ denote the input SPD matrix to the k-th BiSRe layer. The BiMap layer applies a congruence transformation defined as:

$$F_b^{(k)} = f(F^{(k-1)}, W_k) = W_k F^{(k-1)} W_k^{\top}, \tag{12}$$

where $W_k \in \mathbb{R}^{C\times C}$ is a learnable full-rank matrix initialized close to the identity. This operation maps the input SPD matrix to another SPD matrix without vectorization, thereby preserving the intrinsic functional connectivity structure among EEG channels. The positive definiteness of the output is guaranteed by the property of congruence transformations, i.e., $W_k F W_k^{\top} \succ 0$ whenever $F \succ 0$.

To ensure that the mapped matrix remains on the SPD manifold and to introduce nonlinearity into the Riemannian network, an eigenvalue rectification layer is applied after the BiMap transformation.

Specifically, the eigen-decomposition of $F_b^{(k)}$ is computed as:

$$F_b^{(k)} = U_k \Lambda_n U_k^{\top}, \tag{13}$$

where U and Λ_n denote the eigenvectors and eigenvalues, respectively. In classical Riemannian neural networks, the ReEig layer introduces nonlinearity via hard-thresholding of small eigenvalues with a predefined constant ε . Although this guarantees strict positive definiteness of the output matrix, it results in a non-smooth eigenvalue transformation, potentially causing unstable gradients when eigenvalues are near the threshold.

To address this, we replace the standard hard thresholding with a Softplus-based eigenvalue rectification. The rectification function is defined as:

$$E_k(i,i) = \text{Softplus}(\Lambda_n(i,i)) + \varepsilon = \log(1+\exp(\Lambda_n(i,i))), \tag{14}$$

where $\varepsilon > 0$ is a small constant to prevent numerical degeneration. The output of the ReEig layer is reconstructed as:

$$F_s^{(k)} = f_s(F_b^{(k)}) = U_k E_k U_k^\top. \tag{15}$$

This operation guarantees that all eigenvalues remain strictly positive, thereby preserving the SPD property of the matrix while introducing nonlinearity analogous to the ReLU function in conventional convolutional neural networks. Two BiSRe layers per frequency band enable the encoder to learn non-linear SPD-to-SPD transformations, capturing MI task–relevant connectivity structure within each band.

3.4 Multi-Scale Riemannian Network

After refining F_{fused} with stacked BiSRe modules to obtain the discriminant SPD representation on the Riemannian manifold, it is further processed by a designed multi-scale Riemannian network to capture the functional interactions related to MI at different spatial resolutions. This design is motivated by neurophysiological evidence that MI modulates brain connectivity at global, hemispheric, and local sensorimotor scales. To facilitate scale-aware modeling, EEG channels are reordered as $[\mathbf{C}_{\mathbf{left}}, \mathbf{C}_{\mathbf{right}}]$, with channels within each hemisphere grouped by anatomical regions, enabling structured extraction of multi-scale connectivity patterns on the SPD manifold.

The network consists of three complementary branches.

Whole-brain branch. The full-channel FC matrix is directly processed by two cascaded BiSRe layers to encode global brain network organization.

Hemispheric branch. To capture lateralized MI patterns, the FC matrix is partitioned into intra- and inter-hemispheric principal submatrices corresponding to left–left, left–right, right–left, and right–right connectivity. According to Theorem 1, each submatrix, which remains SPD as a principal submatrix of F_{fused}, is embedded using shared BiSRe layers, and the resulting features are concatenated to model hemispheric interactions relevant to left/right MI.

Theorem 1. *Let $M \in \mathcal{S}_{++}^{C}$. For any index set $I \subset \{1, \ldots, C\}$ with $|I| = w$, the principal submatrix $M[I]$ is in $\mathcal{S}_{++}^{w}$. (The proof is provided in Appendix A.)*

Local branch. Fine-grained regional connectivity is extracted using a sliding-window mechanism on the FC matrix, producing a set of local SPD submatrices that capture micro-scale sensorimotor interactions. These local representations are mapped via shared BiSRe layers and aggregated to characterize localized connectivity patterns.

After scale-specific Riemannian encoding, the resulting SPD matrices are projected onto the tangent space using the Log-Euclidean mapping and vectorized by extracting upper triangular elements. Feature vectors from the whole-brain, hemispheric, and local branches are concatenated and fed into a fully connected classifier for MI decoding.

By jointly aggregating connectivity representations across multiple spatial scales while preserving manifold structure, the proposed network captures complementary global, lateralized, and local brain network characteristics, leading to more discriminative and interpretable MI representations.

4 Experiments

4.1 Datasets and Evaluation Protocols

Experiments are conducted on two complementary MI EEG datasets: the standardized BCI Competition IV 2a (22×22 FC matrix) and a clinical lower-limb MI (LL-MI) dataset from stroke patients (40×40 FC matrix). The two complement each other and are used to evaluate methods in standard benchmarks and real clinical scenarios. Performance is evaluated using accuracy, F1 score, and Cohen's kappa. All tests run on an RTX 5090 D GPU (32GB) and an i7 CPU @ 3.4GHz. The average end-to-end processing time per trial is approximately 0.22s, with core MI inference around 0.1s.

BCI-IV 2a. The dataset [Tangermann *et al.*, 2012] comprises EEG recordings from 9 healthy subjects performing four-class MI tasks: left hand (LH), right hand (RH), both feet (BF), and tongue (Tg). Owing to its standardized experimental design and well-established evaluation protocols, this dataset serves as a canonical benchmark for MI decoding. Both intra-subject and inter-subject evaluation protocols are adopted to assess the robustness and generalization ability of the proposed method in a multi-class setting.

LL-MI. The dataset [Liu *et al.*, 2025] contains EEG recordings from 27 post-stroke patients performing lower-limb MI tasks, representing a challenging and clinically relevant scenario. In this work, we focus on the gait-phase-encoded sequential electrical stimulation paradigm, which includes gait MI and idle-state tasks. Evaluation is conducted across subjects within this paradigm to examine the robustness and interpretability of the proposed framework in rehabilitation-oriented BCI applications.

4.2 Comparison with Prior Art

The results in Table 1 show that our MFMSRNet consistently surpasses representative baselines in both intra- and inter-subject evaluations on BCI-IV 2a and LL-MI datasets. In both settings, Euclidean-based models (e.g., EEGNet, EEG-Inception) perform less effectively. This limitation suggests that shallow spatial-temporal convolutions alone are inadequate for capturing discriminative patterns in multi-class MI. Transformer-hybrid architectures such as EEG-Conformer achieve moderate improvements but lag behind geometry-aware and attention-enhanced methods, particularly in the more challenging inter-subject evaluation, where cross-individual variability dominates.

Tensor-CSPNet and Graph-CSPNet, which incorporate geometric insights into SPD manifold structures, deliver stronger performance than purely Euclidean networks, yet they remain constrained by their focus on second-order statistics without adaptive frequency fusion. Attention-driven models such as M-FANet and CorAtt-MIX further narrow this gap by leveraging multi-feature attention or correlation manifold self-attention, demonstrating the value of adaptive weighting mechanisms. However, these approaches still do not fully exploit both FC geometry and adaptive multi-frequency integration.

Notably, our MFMSRNet achieves the highest accuracy, F1 scores, and Cohen's kappa values across all benchmarks, substantially surpassing CorAtt-MIX and other manifold baselines. On the LL-MI dataset, MFMSRNet attains 94.69%

Baseline	BCI-IV 2a (Intra-subject)			BCI-IV 2a (Inter-subject)			LL-MI		
	Accuracy(%)	F1 Score(%)	Kappa	Accuracy(%)	F1 Score(%)	Kappa	Accuracy(%)	F1 Score(%)	Kappa
EEGNet (2018)	70.59 ± 7.07	69.48 ± 4.14	0.6079	59.03 ± 7.36	57.92 ± 4.26	0.4538	91.24 ± 1.39	90.27 ± 1.01	0.8248
FBCNet (2020)	72.95 ± 6.08	69.88 ± 5.66	0.6343	61.38 ± 6.1	59.21 ± 6.21	0.4851	88.97 ± 2.42	87.35 ± 1.5	0.7814
EEG-Inception (2021)	69.37 ± 8.37	67.38 ± 7.02	0.5750	57.81 ± 8.17	55.82 ± 7.52	0.4458	87.19 ± 1.36	87.44 ± 1.73	0.7688
EEG-Conformer (2022)	73.10 ± 8.50	71.05 ± 6.43	0.6633	60.94 ± 8.43	58.33 ± 7.94	0.4692	89.07 ± 2.16	88.24 ± 2.03	0.7952
Tensor-CSPNet (2022)	75.53 ± 7.95	71.57 ± 6.94	0.6738	63.62 ± 8.21	59.65 ± 7.04	0.5148	92.24 ± 1.19	90.93 ± 1.66	0.8448
Graph-CSPNet (2023)	76.4 ± 7.4	75.09 ± 6.99	0.6803	64.84 ± 7.41	64.41 ± 6.35	0.5312	92.62 ± 1.86	92.15 ± 1.73	0.8524
M-FANet (2024)	78.05 ± 9.18	76.12 ± 9.22	0.7023	66.49 ± 9.19	64.33 ± 10.58	0.5532	91.09 ± 1.53	91.21 ± 1.75	0.8218
CorAtt-MIX (2025)	77.2 ± 7.6	76.07 ± 8.12	0.6911	65.64 ± 7.75	65.39 ± 7.77	0.5418	92.82 ± 1.06	92.93 ± 1.46	0.8564
MFMSRNet (Ours)	**81.72** ± 6.07	**79.41** ± 5.78	**0.7364**	**67.38** ± 6.33	**67.29** ± 6.11	**0.5851**	**94.69** ± 1.2	**94.81** ± 1.39	**0.8938**

Table 1: Performance comparison with baselines in terms of accuracy (in %), F1 score (in %) and kappa on BCI-IV 2a and LL-MI datasets.

accuracy, 94.81% F1 score, and 0.8938 kappa, indicating not only strong discriminative capability but also robust generalization in a binary classification paradigm. These consistent improvements underscore that the proposed MFMSRNet yields a more expressive and resilient representation of EEG dynamics compared to existing state-of-the-art methods.

4.3 Explainability Analysis

The proposed framework produces fused FC representations with inherent interpretability by explicitly modeling EEG interactions on the Riemannian manifold of SPD matrices, enabling direct neurophysiological analysis [Wang *et al.*, 2025].

Graph-theoretic Interpretability Analysis

To assess whether the fused FC representations preserve meaningful brain organization, we conduct a graph-theoretic analysis based on network topology metrics (see Appendix B for metric calculations), with a focus on small-world properties [Gallego-Molina *et al.*, 2022]. As shown in Figure 4, all fused FC networks exhibit significant small-world properties (> 1), indicating a balance between local segregation and global integration [Bassett and Bullmore, 2017].

For the BCI-IV 2a dataset (Figure 4 (a)), hand MI shows significantly higher small-world properties than Tg and BF imagery ($p < 0.05$), linked to greater local clustering without compromising path length — consistent with the denser cortical representation of hand movements. No significant differences were observed between Tg and BF imagery or between LH and RH.

For the LL-MI dataset (Figure 4 (b)), the Idle state consistently shows a significantly higher small-world property than Gait MI ($p < 0.01$). This reflects higher clustering and shorter path lengths at rest, while Gait MI reduces clustering and slightly increases path length, indicating task-related network reconfiguration [Neyland *et al.*, 2021; Su *et al.*, 2025].

Frequency- and Region-aware Interpretability Analysis

We further investigate the interpretability of the proposed attention-based Riemannian fusion from spectral and regional perspectives. In the LL-MI dataset (Figure 5 (a)), Idle is dominated by mu and low-beta bands, while Gait MI assigns higher attention to high-beta and gamma bands, reflecting task-relevant spectral adaptation.

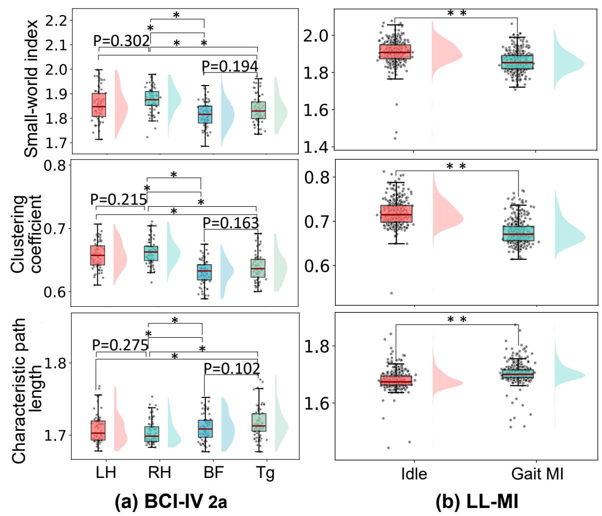

Figure 4: Small-world properties and topology of fused FC matrices are shown. Scatter plots of clustering coefficient and path length reflect local segregation vs. global integration. Paired tests with FDR correction were used for significance ($*p < 0.05$, $**\,p < 0.01$).

Region-wise FC analysis (Figure 5 (b)), after trace normalization and significance filtering, reveals distinct spatial patterns. Gait MI shows stronger connectivity within the central sensorimotor region, reflecting enhanced sensorimotor integration during gait imagery. In contrast, Idle shows more distributed connectivity across bilateral motor and frontal–parietal regions, associated with internally oriented processing without explicit motor demands.

Similar trends appear in the BCI-IV 2a dataset (Figure 5 (c, d)). All MI classes are dominated by mu and beta rhythms, with BF and Tg imagery showing higher contributions from higher-frequency bands. Attention distributions are symmetric between LH and RH MI. Region-wise FC reveals contralateral motor cortex dominance during hand imagery, while foot and tongue imagery involve stronger parietal and parieto-occipital engagement.

Unlike prior methods that provide mainly post-hoc or feature-level interpretability, MFMSRNet offers intrinsic connectivity-level explanations grounded in Riemannian geometry, enabling frequency- and region-specific neurophysiological interpretation.

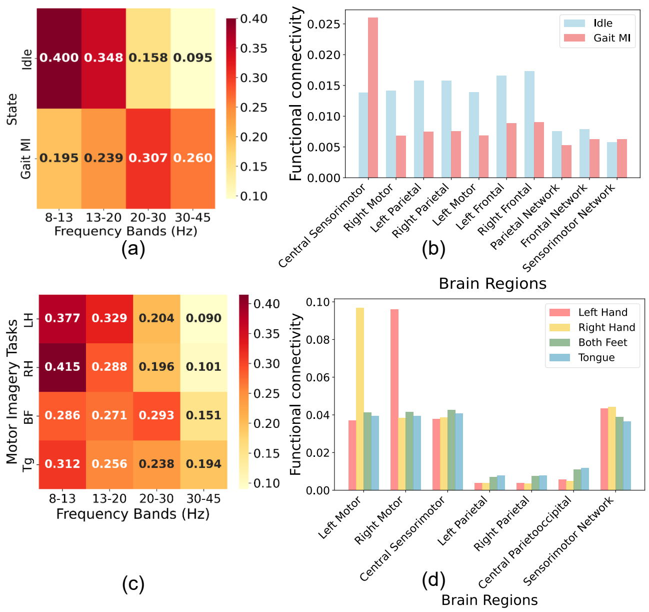

Figure 5: Interpretability analysis of attention-based Riemannian fusion at frequency and region levels. (a, c) Attention weight distribution across frequency bands. (b, d) Region-wise FC strength comparison after trace normalization; only significantly different regions are displayed.

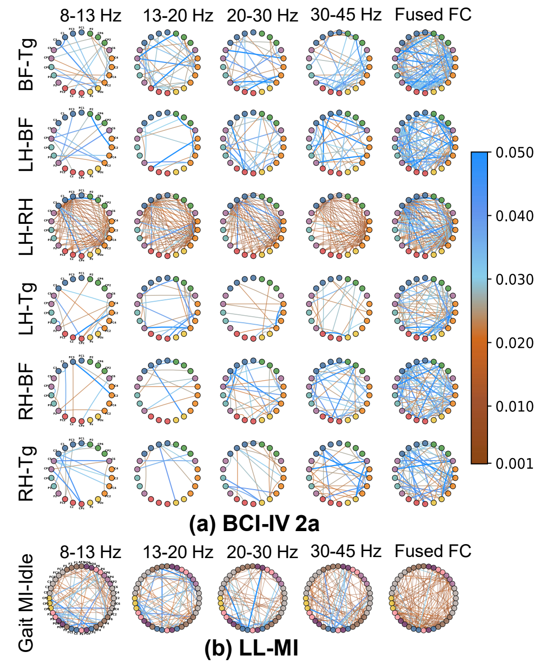

Figure 6: Comparison of single-band vs. fused FC for (a) BCI-IV 2a and (b) LL-MI datasets. Circular plots show FC patterns: nodes are EEG channels; lines denote significant between-group connectivity differences, with line color indicating p-value.

5 Ablation Studies

To verify the effectiveness of multi-frequency FC fusion, ablation studies are conducted comparing the performance of individual frequency bands against the fused FC. As shown in Figure 6 (sparsity level fixed at 0.05 for both datasets), while individual bands exhibit certain advantages in capturing specific inter-group differences, the fused FC comprehensively integrates unique discriminative features from all bands, leading to significantly improved overall classification performance.

Dataset	Scale	Accuracy(%)	F1 Score(%)	Kappa
BCI-IV 2a	S_w	78.16 ± 6.45	75.82 ± 6.18	0.6850
	S_h	76.88 ± 6.90	74.50 ± 6.60	0.6624
	S_l	79.40 ± 6.25	77.15 ± 5.95	0.7018
	$S_w + S_h$	80.35 ± 6.20	78.08 ± 5.90	0.7115
	$S_w + S_l$	80.90 ± 6.15	78.68 ± 5.85	0.7200
	$S_h + S_l$	79.95 ± 6.18	77.75 ± 5.88	0.7062
	$S_w + S_h + S_l$	**81.72 ± 6.07**	**79.41 ± 5.78**	**0.7364**
LL-MI	S_w	92.45 ± 1.48	92.33 ± 1.68	0.8542
	S_h	91.80 ± 1.65	91.72 ± 1.85	0.8420
	S_l	93.20 ± 1.40	93.15 ± 1.62	0.8680
	$S_w + S_h$	94.00 ± 1.35	93.96 ± 1.58	0.8805
	$S_w + S_l$	94.35 ± 1.30	94.30 ± 1.52	0.8865
	$S_h + S_l$	93.65 ± 1.38	93.58 ± 1.60	0.8740
	$S_w + S_h + S_l$	**94.69 ± 1.2**	**94.81 ± 1.39**	**0.8938**

Table 2: Performance comparison of different brain region fusion methods on BCI-IV 2a and LL-MI datasets.

The decoding network further improves recognition performance by fusing FC across whole-brain (S_w), hemispheric (S_h), and local region (S_l) scales. Ablation experiments confirm the effectiveness of this multi-region selection mechanism, showing that models using each scale alone or in partial combinations are outperformed by the full three-scale fusion (Table 2). Specifically, on the BCI-IV 2a dataset, the combined S_w, S_h, and S_l fusion yields the best results across all metrics, while achieving 94.69% accuracy on the LL-MI dataset. These findings demonstrate that the proposed coarse-to-fine feature fusion strategy substantially enhances model performance.

6 Conclusion

This work presents MFMSRNet, an interpretable end-to-end MI decoding framework that integrates kernelized FC construction, attention-based Riemannian fusion, and multi-scale geometric feature learning. By modeling EEG FC as SPD matrices and performing adaptive fusion directly on the Riemannian manifold, the proposed method effectively captures non-linear dependencies and preserves the intrinsic geometric structure of brain networks. Extensive experiments on two public MI datasets demonstrate that the proposed framework consistently outperforms state-of-the-art baselines in both intra- and inter-subject settings. Moreover, comprehensive explainability analyses, including small-world topology, frequency-specific attention patterns, and region-level connectivity strength, reveal neurophysiologically meaningful representations aligned with established motor control mechanisms.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No.62372338) and the Key Research and

Development Projects in Hubei Province (No.2023BCB133).

References

[Ajmeria *et al.*, 2022] Rahul Ajmeria, Mayukh Mondal, Reya Banerjee, Tamesh Halder, Pallav Kumar Deb, Debasish Mishra, Pravanjan Nayak, Sudip Misra, Surjya Kanta Pal, and Debashish Chakravarty. A critical survey of EEG-based BCI systems for applications in industrial internet of things. *IEEE Communications Surveys & Tutorials*, 25(1):184–212, 2022.

[Altaheri *et al.*, 2022] Hamdi Altaheri, Ghulam Muhammad, and Mansour Alsulaiman. Physics-informed attention temporal convolutional network for EEG-based motor imagery classification. *IEEE Transactions on Industrial Informatics*, 19(2):2249–2258, 2022.

[Altaheri *et al.*, 2023] Hadeer Altaheri, Ghulam Muhammad, Mansour Alsulaiman, et al. Deep learning techniques for classification of electroencephalogram (EEG) motor imagery (MI) signals: a review. *Neural Computing and Applications*, 35:14681–14722, 2023.

[Atz *et al.*, 2021] Kenneth Atz, Francesca Grisoni, and Gisbert Schneider. Geometric deep learning on molecular representations. *Nature Machine Intelligence*, 3:1023–1032, 2021.

[Bassett and Bullmore, 2017] Danielle S. Bassett and Edward T. Bullmore. Small-world brain networks revisited. *The Neuroscientist*, 23(5):499–516, 2017.

[Bronstein *et al.*, 2017] Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: Going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

[Dai *et al.*, 2019] Mengyu Dai, Zhengwu Zhang, and Anuj Srivastava. Analyzing dynamical brain functional connectivity as trajectories on space of covariance matrices. *IEEE Transactions on Medical Imaging*, 39(3):611–620, 2019.

[Dan *et al.*, 2022] Tianping Dan, Zhaohua Huang, Hongmin Cai, Paul J Laurienti, and Guorong Wu. Learning brain dynamics of evolving manifold functional mri data using geometric-attention neural network. *IEEE Transactions on Medical Imaging*, 41(10):2825–2836, 2022.

[Dissanayake *et al.*, 2022] Tharindu Dissanayake, Tharindu Fernando, Simon Denman, Sridha Sridharan, and Clinton Fookes. Geometric deep learning for subject independent epileptic seizure prediction using scalp EEG signals. *IEEE Journal of Biomedical and Health Informatics*, 26(2):527–538, 2022.

[Fei *et al.*, 2025] Yanhong Fei, Yingjie Liu, Chentao Jia, Zhengyu Li, Xian Wei, and Mingsong Chen. A survey of geometric optimization for deep learning: From euclidean space to riemannian manifold. *ACM Computing Surveys*, 57(5):1–37, 2025.

[Fu *et al.*, 2025] Rongrong Fu, Linmeng Shen, Bin Lu, Mengpu Cai, Guilin Wen, Junxiang Chen, and Chengcheng Hua. A convolutional transformer network with adaptation learning modules for enhancing motor imagery classification. *Expert Systems with Applications*, 269:126381, 2025.

[Gallego-Molina *et al.*, 2022] Nicolás J Gallego-Molina, Andrés Ortiz, Francisco J Martínez-Murcia, Marco A Formoso, and Almudena Giménez. Complex network modeling of EEG band coupling in dyslexia: An exploratory analysis of auditory processing and diagnosis. *Knowledge-based Systems*, 240:108098, 2022.

[Gerken *et al.*, 2023] Jan E. Gerken, Jimmy Aronsson, Oscar Carlsson, Hampus Linander, Fredrik Ohlsson, Christoffer Petersson, and Daniel Persson. Geometric deep learning and equivariant neural networks. *Artificial Intelligence Review*, 56(12):14605–14662, 2023.

[He and Wu, 2020] He He and Dongrui Wu. Transfer learning for brain–computer interfaces: A euclidean space data alignment approach. *IEEE Transactions on Biomedical Engineering*, 67(2):399–410, 2020.

[Hu *et al.*, 2025] Chen Hu, Rui Wang, Xiaoning Song, Tao Zhou, Xiao-Jun Wu, Nicu Sebe, and Ziheng Chen. A correlation manifold self-attention network for EEG decoding. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence (IJCAI)*, pages 5372–5380, 2025.

[Ju and Guan, 2022] Ce Ju and Cuntai Guan. Tensor-cspnet: A novel geometric deep learning framework for motor imagery classification. *IEEE Transactions on Neural Networks and Learning Systems*, 34(12):10955–10969, 2022.

[Ju and Guan, 2024] Ce Ju and Cuntai Guan. Graph neural networks on SPD manifolds for motor imagery classification: A perspective from the time–frequency analysis. *IEEE Transactions on Neural Networks and Learning Systems*, 35(12):17701–17715, 2024.

[Kalaganis *et al.*, 2022] Fotis P Kalaganis, Nikos A Laskaris, Vangelis P Oikonomou, Spiros Nikopolopoulos, and Ioannis Kompatsiaris. Revisiting riemannian geometry-based EEG decoding through approximate joint diagonalization. *Journal of Neural Engineering*, 19(6):066030, 2022.

[Lawhern *et al.*, 2018] Vernon J. Lawhern, Amelia J. Solon, Nicholas R. Waytowich, Stephen M. Gordon, Chou P. Hung, and Brent J. Lance. EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of Neural Engineering*, 15(5):056013, 2018.

[Liu *et al.*, 2016] Chuanwei Liu, Yunfa Fu, Jun Yang, Xin Xiong, Huiwen Sun, and Zhengtao Yu. Discrimination of motor imagery patterns by electroencephalogram phase synchronization combined with frequency band energy. *IEEE/CAA Journal of Automatica Sinica*, 4(3):551–557, 2016.

[Liu *et al.*, 2025] Yuan Liu, Zhuolan Gui, De Yan, Zhuang Wang, Ruisi Gao, Ningxin Han, Junying Chen, Jialing Wu, and Dong Ming. Lower limb motor imagery EEG dataset based on the multi-paradigm and longitudinal-training of stroke patients. *Scientific Data*, 12(1):314, 2025.

[Mane *et al.*, 2020] Ravikiran Mane, Neethu Robinson, A. Prasad Vinod, Seong-Whan Lee, and Cuntai Guan. A multi-view CNN with novel variance layer for motor imagery brain computer interface. In *The 42nd Annual Inter-*

national Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pages 2950–2953, 2020.

[Neyland *et al.*, 2021] Blake R Neyland, Christina E Hugenschmidt, Robert G Lyday, Jonathan H Burdette, Laura D Baker, W Jack Rejeski, Michael E Miller, Stephen B Kritchevsky, and Paul J Laurienti. Effects of a motor imagery task on functional brain network community structure in older adults: Data from the brain networks and mobility function (B-NET) study. *Brain Sciences*, 11(1):118, 2021.

[Paillard *et al.*, 2025] Joseph Paillard, Jörg F. Hipp, and Denis A. Engemann. Green: A lightweight architecture using learnable wavelets and riemannian geometry for biomarker exploration with EEG signals. *Patterns*, 6(3):101182, 2025.

[Pan *et al.*, 2025] Lincong Pan, Kun Wang, Yongzhi Huang, Xinwei Sun, Jiayuan Meng, Weibo Yi, Minpeng Xu, Tzyy-Ping Jung, and Dong Ming. Enhancing motor imagery EEG classification with a riemannian geometry-based spatial filtering (RSF) method. *Neural Networks*, 188:107511, 2025.

[Qin *et al.*, 2024] Yiyang Qin, Banghua Yang, Sixiong Ke, Peng Liu, Fenqi Rong, and Xinxing Xia. M-FANet: Multi-feature attention convolutional neural network for motor imagery decoding. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 32:401–411, 2024.

[Rudin and Cynthia, 2019] Rudin and Cynthia. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5):206–215, 2019.

[Shi *et al.*, 2025] Yuxuan Shi, Aimin Jiang, Ju Zhong, Min Li, and Yanping Zhu. Multiclass classification framework of motor imagery EEG by riemannian geometry networks. *IEEE Journal of Biomedical and Health Informatics*, 29(2):935–947, 2025.

[Song *et al.*, 2022] Yonghao Song, Qingqing Zheng, Bingchuan Liu, and Xiaorong Gao. EEG conformer: Convolutional transformer for EEG decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:710–719, 2022.

[Su *et al.*, 2025] Haolong Su, Gege Zhan, Yifang Lin, Lu Wang, Jie Jia, Lihua Zhang, Zhongxue Gan, and Xiaoyang Kang. Analysis of brain network differences in the active, motor imagery, and passive stoke rehabilitation paradigms based on the task-state EEG. *Brain Research*, 1846:149261, 2025.

[Suh and Hyung, 2021] Kim Suh, Yoon-Je and Byung Hyung. Riemannian embedding banks for common spatial patterns with EEG-based SPD neural networks. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:1107–1116, 2021.

[Tanamachi *et al.*, 2023] Kenya Tanamachi, Wataru Kuwahara, Megumi Okawada, Shun Sasaki, and Fuminari Kaneko. Relationship between resting-state functional connectivity and change in motor function after motor imagery intervention in patients with stroke: A scoping review. *Journal of NeuroEngineering and Rehabilitation*, 20(1):159, 2023.

[Tangermann *et al.*, 2012] Michael Tangermann, Klaus-Robert Müller, Ad Aertsen, Niels Birbaumer, Christoph Braun, Clemens Brunner, Robert Leeb, Carsten Mehring, Kai J Miller, Gernot R Müller-Putz, et al. Review of the BCI competition IV. *Frontiers in Neuroscience*, 6:55, 2012.

[Tibermacine *et al.*, 2024] Imad Eddine Tibermacine, Samuele Russo, Ahmed Tibermacine, Abdelaziz Rabehi, Bachir Nail, Kamel Kadri, and Christian Napoli. Riemannian geometry-based EEG approaches: A literature review. *arXiv preprint arXiv:2407.20250*, 2024.

[Tuzel *et al.*, 2007] Oncel Tuzel, Fatih Porikli, and Peter Meer. Human detection via classification on riemannian manifolds. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2007.

[Varoquaux *et al.*, 2010] Gaël Varoquaux, Alexandre Gramfort, Jean-Baptiste Poline, and Bertrand Thirion. Brain covariance selection: Better individual functional connectivity models using population prior. In *Advances in Neural Information Processing Systems (NIPS)*, volume 23, pages 1–9, 2010.

[Wang *et al.*, 2024] Rui Wang, Chen Hu, Ziheng Chen, Xiaojun Wu, and Xiaoning Song. A grassmannian manifold self-attention network for signal classification. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI)*, pages 5099–5107, 2024.

[Wang *et al.*, 2025] Tao Wang, Rui Mao, Shuang Liu, Erik Cambria, and Dong Ming. Explainable multi-frequency and multi-region fusion model for affective brain-computer interfaces. *Information Fusion*, 118:102971, 2025.

[Wen *et al.*, 2024] Huan Wen, Yucun Zhong, Lin Yao, and Yueming Wang. Neural correlates of motor/tactile imagery and tactile sensation in a BCI paradigm: A high-density EEG source imaging study. *Cyborg and Bionic Systems*, 5:0118, 2024.

[Xu and Yang, 2025] Biao Xu and Guanci Yang. Interpretability research of deep learning: A literature survey. *Information Fusion*, 115:102721, 2025.

[Yger *et al.*, 2017] Florian Yger, Maxime Berar, and Fabien Lotte. Riemannian approaches in brain-computer interfaces: A review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(10):1753–1762, 2017.

[Zhang and Etemad, 2023] Guangyi Zhang and Ali Etemad. Spatio-temporal EEG representation learning on riemannian manifold and euclidean space. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 8(2):1469–1483, 2023.

[Zhang and Li, 2025] Haodong Zhang and Hongqi Li. Transformer-based EEG decoding: A survey. *arXiv preprint arXiv:2507.02320*, 2025.

[Zhang *et al.*, 2023] Dongxue Zhang, Huiying Li, and Jingmeng Xie. MI-CAT: A transformer-based domain adaptation network for motor imagery classification. *Neural Networks*, 165:451–462, 2023.