

基于双分支多特征融合网络的运动想象脑电信号分类方法

于欣琪, 郑钟月, 张依林, 赵菁菁, 张丽艳*

大连交通大学轨道智能工程学院电子与通信工程系, 辽宁 大连

收稿日期: 2026年6月29日; 录用日期: 2026年7月21日; 发布日期: 2026年7月30日

摘要

基于运动想象的脑机接口(MI-BCI)为重度瘫痪患者提供了独立于外周神经和肌肉的人机交互途径, 在运动功能康复领域具有重要应用价值。然而, 脑电信号的会话间变异性大、判别性能量特征提取困难等问题, 给MI-BCI系统的信号解码带来了严峻挑战。针对上述问题, 本文提出一种双分支多特征融合网络(DBMF-Net), 通过多分支结构联合学习EEG信号的一阶时空特征、二阶协方差特征及时频域特征。具体而言, 该模型利用多尺度卷积与Transformer编码器提取时空特征, 引入基于黎曼几何的SPD协方差嵌入机制建模通道间二阶统计关系, 并通过短时傅里叶变换分支捕获时频域判别信息, 最终将三类特征融合用于运动想象分类。在BCI Competition IV 2a和2b数据集上的实验结果表明, DBMF-Net的平均分类准确率分别达到85.87%和89.39%, 优于现有基准方法, 验证了该方法在被试内会话间脑电解码任务中的有效性与鲁棒性。

关键词

运动想象, 脑机接口, 多特征融合, SPD协方差, 时频分析

DBMF-Net: Motor Imagery EEG Signals Classification Method Based on a Dual-Branch Multi-Feature Fusion Network

Xinqi Yu, Zhongyue Zheng, Yilin Zhang, Jingjing Zhao, Liyan Zhang*

Department of Electronic Communications, School of Railway Intelligent Engineering, Dalian Jiaotong University, Dalian Liaoning

Received: June 29, 2026; accepted: July 21, 2026; published: July 30, 2026

*通讯作者。

文章引用: 于欣琪, 郑钟月, 张依林, 赵菁菁, 张丽艳. 基于双分支多特征融合网络的运动想象脑电信号分类方法[J]. 图像与信号处理, 2026, 15(3): 426-440. DOI: 10.12677/jisp.2026.153038

Abstract

The motor imagery-based brain-computer interface (MI-BCI) provides an interaction pathway independent of peripheral nerves and muscles for individuals with severe paralysis, holding significant application value in the field of motor function rehabilitation. However, challenges such as large inter-session variability of EEG signals and the difficulty in extracting discriminative energy features pose serious obstacles to signal decoding in MI-BCI systems. To address these issues, this paper proposes a Dual-Branch Multi-Feature Fusion Network (DBMF-Net), which jointly learns first-order spatiotemporal features, second-order covariance features, and time-frequency domain features of EEG signals through a multi-branch structure. Specifically, the model extracts spatiotemporal features using multi-scale convolution and a Transformer encoder, introduces an SPD covariance embedding mechanism based on Riemannian geometry to model second-order statistical relationships between channels, and captures frequency-domain discriminative information through a short-time Fourier transform branch. Finally, the three types of features are fused for motor imagery classification. Experimental results on the BCI Competition IV 2a and 2b datasets demonstrate that DBMF-Net achieves average classification accuracies of 85.87% and 89.39%, respectively, outperforming existing baseline methods and validating the effectiveness and robustness of the proposed method in intra-subject inter-session EEG decoding tasks.

Keywords

Motor Imagery, Brain-Computer Interface, Multi-Feature Fusion, SPD Covariance, Time-Frequency Analysis

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

脑机接口(Brain-Computer Interface, BCI)通过建立独立于周围神经和肌肉的直接通路,实现人脑与外部设备的信息交换,为严重瘫痪但保有认知意识的患者提供了全新的交互可能[1]。随着人工智能与可穿戴技术的发展,BCI已广泛应用于创新医疗场景,使患者能够凭意念控制假肢、轮椅及光标等设备,显著改善了肌萎缩侧索硬化症(ALS)、中风及脊髓损伤患者的生活质量[2]。在各类BCI实现方式中,基于脑电图(Electroencephalogram, EEG)的非侵入式BCI因安全性高、成本低、无需手术等优势,成为最广泛应用的技术路线。常见范式包括稳态视觉诱发电位(SSVEP)[3]、P300诱发电位[4]和运动想象(MI)[5]等,其中MI作为一种内源性自发EEG范式最受关注。MI-BCI系统在受试者想象特定动作时采集脑电信号,通过解码识别想象内容并转化为设备控制指令。其神经生理学基础在于,不同MI任务引发运动相关脑区振荡活动及脑网络模式的显著差异,表现为 μ 节律(8~13 Hz)和 β 节律(13~30 Hz)等频带的ERD/ERS动态变化。

MI-EEG信号的分类性能在很大程度上取决于特征提取的质量,因此特征提取方法一直是该领域的研究核心。围绕这一问题,研究者先后提出了基于机器学习(Machine Learning, ML)和深度学习(Deep Learning, DL)的两类解码方法。

在机器学习方法中,公共空间模式(Common Spatial Pattern, CSP)[6]及其衍生算法应用最为广泛,包括公共空间谱模式(CSSP)[7]、子带公共空间模式(SBCSP)[8]和滤波器组公共空间模式(FBCSP)[9]等。这

类方法通过最大化类间方差与最小化类内方差来增强 MI-EEG 信号的可区分性,并结合线性判别分析(LDA)等分类器完成最终识别[10]。此外,CSP 及其变体还被集成至更复杂的解码框架中,如 Dagdevir 等[11]将 CSP 与截断阈值经验模式分解相结合用于 MI-EEG 解码。然而,上述方法均依赖先验知识与手工特征工程,特征表达能力有限,难以充分挖掘 EEG 信号中隐含的复杂非线性结构。

深度学习方法凭借端到端的自动特征学习能力,为上述问题提供了有效地解决途径。Schirrneister 等[12]设计了 ShallowConvNet 和 DeepConvNet 等卷积神经网络架构,实现了从原始 EEG 信号中直接解码运动想象任务。Lawhern 等[13]提出的 EEGNet 通过深度可分离卷积同时提取频域与时域特征,在多种 BCI 范式中表现出良好的泛化能力。Xie 等[14]提出了基于 Transformer 的解码框架,通过多种位置编码模块有效捕捉 EEG 信号的时空依赖关系。

综上,现有方法多局限于一阶时空特征的提取,对 EEG 信号中时-空-频多域耦合关系及高阶统计结构的建模能力不足。为此,本文提出一种基于 CNN 与 Transformer 的双分支多特征融合网络(DBMF-Net)。该方法在提取一阶时空特征的基础上,引入基于黎曼几何的 SPD 协方差[15]嵌入机制以建模通道间二阶统计关系,并增设基于短时傅里叶变换(STFT)的时频域分支以捕获时频判别信息,最终通过多特征融合实现运动想象分类。

2. 相关工作

2.1. 整体网络架构

为充分挖掘运动想象脑电信号中不同层级的判别信息,本文构建了一种双分支网络结构。时空域分支从原始 EEG 信号中提取一阶时空特征与二阶统计特征,STFT 时频域分支从时频表示中提取频谱特征,两条分支协同互补,实现对 MI-EEG 信号的多层级联合建模。整体模型结构如图 1 所示。

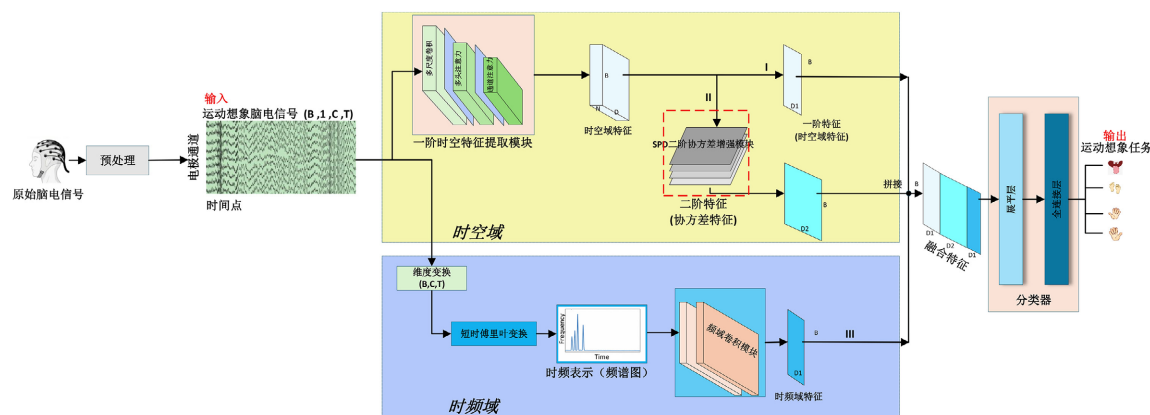


Figure 1. Overall model architecture
图 1. 整体模型架构

DBMF-Net 由时空域分支和 STFT 时频域分支构成双分支框架。时空域分支基于 CNN 与 Transformer 提取 EEG 信号的一阶时空特征,并通过 SPD 协方差级联模块结合黎曼几何学习获取二阶统计特征;STFT 时频域分支通过短时傅里叶变换提取时频特征,以增强模型对 μ 和 β 频段判别信息的建模能力。随后,将两类分支的特征在特征维度上进行拼接融合,输入分类器完成运动想象类别预测。

2.2. 时空域分支

时空域分支是 DBMF-Net 的核心特征提取通路,负责从原始 EEG 信号中学习时空模式与通道间的高

阶统计依赖关系。该分支由三个级联模块组成：首先通过多尺度时空卷积与 Transformer 编码器提取一阶时空特征，随后经 ECA 通道注意力机制对特征进行自适应加权，最后通过 SPD 协方差模块对通道间的二阶统计关系进行建模。各模块依次衔接，形成从局部时空特征到全局统计结构的层级化特征学习流程。

2.2.1. 一阶时空特征提取模块

一阶时空特征提取模块的结构如图 2 所示。该模块首先通过多尺度时间卷积捕获不同频段的时序信息，继而利用空间卷积融合多通道空间特征，最后结合 Transformer 自注意力机制与 ECA 通道注意力机制建模全局时序依赖关系。

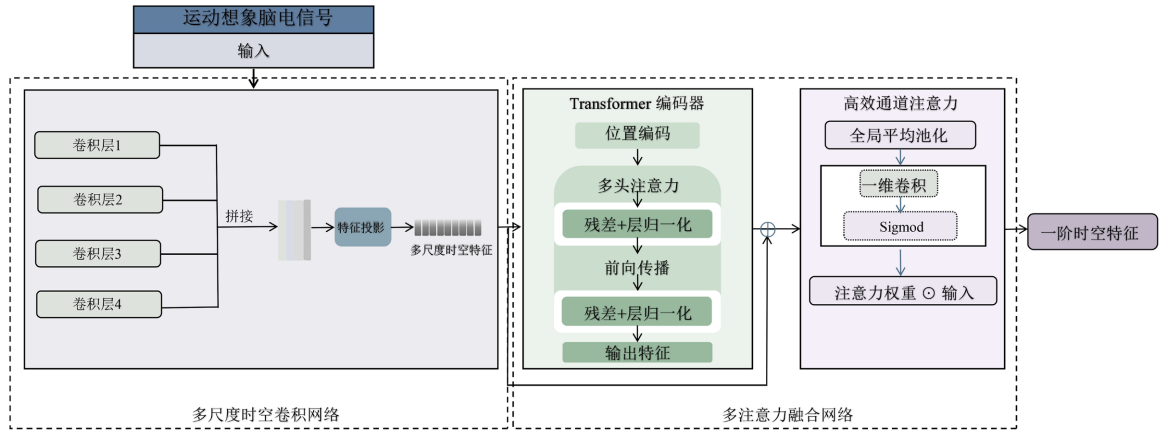


Figure 2. First-order spatiotemporal feature extraction module

图 2. 一阶时空特征提取模块

(1) 多尺度时空卷积网络

EEG 信号包含不同频段的神经营荡活动，其中 μ 节律(8~13 Hz)和 β 节律(13~30 Hz)与运动想象任务密切相关。运动想象过程中，感觉运动皮层会产生事件相关去同步化(ERD)和事件相关同步化(ERS)现象，表现为特定频段功率的显著变化[16]-[18]。为捕获不同频段的判别信息，本模块设置四个并行时间卷积积分支，卷积核尺寸分别为 1×16 、 1×32 、 1×64 和 1×128 ，以不同时间感受野提取多尺度时序特征。多尺度时空卷积网络如图 2 左部分所示，各分支卷积仅沿时间维度进行，在保留通道空间结构的同时有效建模信号的时间动态。二维卷积运算可表示为：

$$Y_k(b, j, c, t) = \sum_{i=1}^{C_{in}} \sum_{m=0}^{K_h-1} \sum_{n=0}^{K_w-1} W_k(j, i, m, n) X(b, i, c+m, t+n) \quad (1)$$

其中 W_k 表示第 k 个卷积分支的卷积核权重， K_h 和 K_w 分别表示卷积核的高度和宽度， C_{in} 表示输入特征通道数。

本研究所使用的 EEG 数据采样率为 250 Hz，即每个采样点约对应 4 ms 的时间分辨率。因此，当卷积核长度分别为 16、32、64 和 128 时，其对应的时间感受野约为 64 ms、128 ms、256 ms 和 512 ms。从信号处理角度来看，时间窗口长度决定了卷积操作能够感知的最低频率范围，其近似关系为：

$$f_{min} \approx \frac{1}{T_k} \quad (2)$$

其中， T_k 表示卷积核对应的时间窗口长度。因此，不同尺度卷积核通过不同时间感受野提取多频段时序特征。随后，采用尺寸为 $(C \times 1)$ 的空间卷积建模 EEG 通道间的空间相关性，并利用分组卷积增强空间特征表达能力。为进一步挖掘时空信息，在空间卷积后引入尺寸为 (1×16) 的时间卷积进行特征增强。最后，

将各卷积分支提取的多尺度特征在通道维度进行融合，其表达形式为：

$$F = \text{Concat}(F_1, F_2, F_3, F_4) \quad (3)$$

在多尺度时空卷积网络提取到的时空特征拼接后，为了将通道维度上融合的高维特征映射到网络后续模块所需的固定嵌入维度(embedding size)，本文引入了一个特征投影层(projection layer)。具体而言，该层首先通过 Rearrange 操作将卷积输出的四维张量 $x \in R^{B \times C_{\text{conv}} \times H \times W}$ 重排为三维张量 $R^{B \times (H \cdot W) \times C_{\text{conv}}}$ ，即将空间维度展平，使每个时间步的多尺度特征作为一个向量表示。随后，通过线性映射层将每个时间步的多尺度特征向量投影到固定的嵌入维度 emb_size，从而得到适合 Transformer 等后续模块处理的特征表示：

$$F_{\text{proj}} = \text{Linear}(\text{Flatten}(F_{\text{conv}})) \in R^{B \times T \times \text{emb_size}} \quad (4)$$

(2) 多注意力融合网络

在运动想象任务中，特征的判别信息既蕴含于时间序列的全局结构中，也体现在不同特征通道的重要性差异上。Transformer 的多头自注意力机制(MHSA)擅长建模全局时序依赖关系，而通道注意力机制则能够自适应强化关键特征维度。二者具有明显的互补性。

基于这一思想，本文设计了一种多注意力融合策略(Multi-Attention Fusion)。多尺度卷积特征网络如图 2 右部分所示，通过融合 Transformer 的多头注意力机制(Multi-Head Attention Mechanism, MHA)与高效通道注意力机制(Efficient Channel Attention, ECA) [19]，在提升特征表达能力的同时保持较低的计算复杂度，从而增强模型对脑电信号动态模式的建模能力。将特征输入 Transformer 编码器以捕获 EEG 信号中的全局时序依赖关系：

$$X_t = \text{TransformerEncoder}(X) \quad (5)$$

Transformer 编码器由多个堆叠的编码块组成，每个编码块包含多头自注意力模块与前馈网络。在多头注意力计算中，输入特征首先通过线性变换映射为查询(Query)、键(Key)和值(Value)：

$$Q = XW_Q, K = XW_K, V = XW_V \quad (6)$$

随后通过缩放点积注意力计算序列中不同位置之间的相关性：

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{D}}\right)V \quad (7)$$

为了增强模型表达能力，该注意力计算被划分为多个子空间进行并行建模，即多头注意力机制：

$$\text{MHA}(X) = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h)W_O \quad (8)$$

其中 h 表示注意力头数， W_O 为输出投影矩阵。通过这种方式，模型能够在不同子空间中捕获多样化的时序依赖关系，从而更有效地建模脑电信号中的复杂动态模式。

在获得 Transformer 输出特征后，为了避免过度依赖注意力结构而削弱卷积网络提取的局部时空特征，本文采用残差连接方式对两种特征进行融合：

$$X_f = X + X_t \quad (9)$$

该策略能够同时保留卷积网络提取的局部时空特征以及 Transformer 建模得到的全局时序依赖信息，从而形成更加完整和稳定的特征表示。为进一步增强关键特征的表达能力，本文在多尺度特征融合后引入 ECA (Efficient Channel Attention) 注意力机制进行自适应加权。与 CBAM 等传统注意力机制相比，ECA 通过一维卷积直接建模局部跨通道交互关系，能够以较低的计算开销提升特征表征能力。因此，ECA 在

保证参数效率的同时，更适用于脑电信号分析任务。首先在序列维度上对融合特征进行平均池化，获得通道描述向量：

$$z = \frac{1}{N} \sum_i^N X_f(i) \quad (10)$$

随后利用一维卷积建模通道之间的局部交互关系，并通过 Sigmoid 激活函数生成通道权重：

$$s = \sigma(\text{Conv1D}(z)) \quad (11)$$

最终，通过通道加权方式对输入特征进行重新标定：

$$X_{out} = X_f \odot s \quad (12)$$

其中 $\odot$ 表示逐元素乘法。

通过多注意力融合策略，模型能够同时建模 EEG 信号的长程时序依赖和关键特征维度的重要性。其中，Transformer 用于捕获运动想象过程中跨时间尺度的神经动态变化，而 ECA 则自适应强化与 μ 节律和 β 节律相关的判别性特征。两种注意力机制的协同作用提升了模型对运动想象脑电模式的表征能力和分类性能。

2.2.2. SPD 二阶协方差增强模块

为进一步挖掘通道间的统计依赖关系，本文在 CNN + Transformer 框架基础上引入基于对称正定(SPD)流形的二阶协方差嵌入模块。该模块通过协方差矩阵建模特征通道间的二阶相关性，并结合注意力机制优化协方差表示，以增强特征的判别性。SPD 二阶协方差模块结构如图 3 所示。

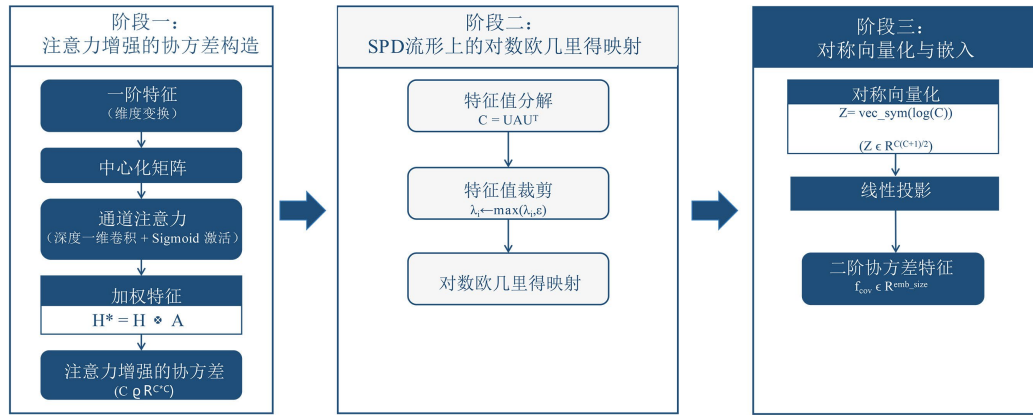


Figure 3. Structure diagram of the SPD second-order covariance module
图 3. SPD 二阶协方差模块结构图

(1) Attention 增强的中心化协方差构造

经过维度变换的一阶特征形状为：

$$H \in R^{B \times C \times 1 \times T}$$

其中： B 为批量大小， C 为特征通道数， T 为时间维度长度。

为了消除时间维度均值的影响，首先构造中心化矩阵：

$$I_c = \frac{1}{T} I - \frac{1}{T^2} \mathbf{1} \mathbf{1}^T \quad (13)$$

其中 I 为单位矩阵， $\mathbf{1}$ 为全 1 向量， $\mathbf{1}^T$ 为向量的转置。该矩阵实现时间维度的去均值操作，其作用等价

于对特征矩阵进行中心化处理。在此基础上,引入 Riemannian Attention 机制[20]对特征进行自适应加权。首先通过通道维度的深度可分离卷积提取注意力权重:

$$A = \sigma(f(H)) \quad (14)$$

其中: $f(\cdot)$ 表示基于一维卷积的注意力映射函数, $\sigma(\cdot)$ 为 Sigmoid 激活函数, $A \in R^{B \times C \times 1 \times T}$ 。

随后对输入特征进行加权:

$$\tilde{H} = H \odot A \quad (15)$$

其中 $\odot$ 表示逐元素乘法。在此基础上构造加权中心化协方差矩阵:

$$C = \tilde{H} \mathbf{I}_c \tilde{H}^T + H \mathbf{I}_c H^T + \epsilon \mathbf{I} \quad (16)$$

(2) SPD 流形结构与 Log-Euclidean 映射

SPD 矩阵集合构成一个非欧氏黎曼流形。在该流形上,矩阵间的距离与运算不能直接采用欧氏空间中的线性操作,否则会破坏其几何结构[21]。为在保持流形结构的同时降低计算复杂度,本文采用 Log-Euclidean [22]度量进行映射。具体步骤如下:

首先对协方差矩阵进行特征值分解:

$$C = U \Lambda U^T \quad (17)$$

其中: U 为特征向量矩阵, $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_c)$, 随后对特征值进行对数映射:

$$\log(C) = U \log(\Lambda) U^T \quad (18)$$

其中 $\log(\Lambda) = \text{diag}(\log(\lambda_1), \dots, \log(\lambda_c))$,

为确保数值稳定性,对特征值进行下界截断,其中 ε 为极小正数。

$$\lambda_i \leftarrow \max(\lambda_i, \varepsilon)$$

(3) 对称向量化与特征嵌入

由于对称矩阵包含冗余信息,仅需保留上三角(含对角线) [23]元素即可完整表示矩阵。本文定义对称向量化操作:

$$z = \text{vec}_{\text{sym}}(\log(C)) \quad (19)$$

则: $z \in R^{C(C+1)/2}$, 该步骤显著降低了特征维度,同时避免信息冗余。

随后通过线性映射实现特征压缩: $f_{\text{cov}} = w_z$

其中: $f_{\text{cov}} \in R^{\text{embsize}}$, embsize 为具体的特征向量维度,该表示作为二阶统计嵌入向量,用于后续多分支特征融合

2.3. STFT 时频域分支

脑电信号具有显著的非平稳特性,其频率成分随时间动态变化。在运动想象任务中, θ 节律(4~8 Hz)、 μ 节律(8~13 Hz)和 β 节律(13~30 Hz)表现出明显的 ERD/ERS 现象[24], 仅依赖时域建模难以充分刻画其频谱判别信息。为此,本文设计基于短时傅里叶变换(STFT)的时频域分支[25], 从时频表示中提取判别特征,与时空域分支形成互补。其结构如图 4 所示。

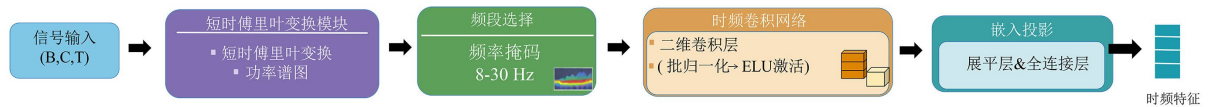


Figure 4. Structure diagram of the STFT branch

图 4. STFT 分支结构图

首先对输入信号沿时间维度进行短时傅里叶变换, 将一维时间序列映射为二维时频表示。STFT 的定义如下:

$$S(f, \tau) = \left| \sum_t x(t) w(t - \tau) e^{-j2\pi f t} \right|^2 \quad (20)$$

其中 $w(\cdot)$ 表示滑动窗函数, f 表示频率, τ 表示时间帧位置。在实现中, 本文采用基于 GPU 的 STFT 计算方式, 并取其幅值平方作为功率谱(power spectrogram):

$$S = |\text{STFT}(x)|^2 \quad (21)$$

从而得到时频特征表示:

$$S \in R^{B \times C \times T \times T'}$$

其中 F 表示频率维度, T' 表示时间帧数。

为了突出与运动想象任务相关的判别性频率成分, 本文引入频率掩码, 仅保留 8~30 Hz 范围内的频率信息:

$$S' = S \odot M \quad (22)$$

其中 M 为频率掩码矩阵。该频带覆盖 μ 节律及 β 节律范围, 有助于增强与运动想象任务相关的节律特征, 同时减少与任务无关频段成分的干扰。

在获得时频表示后, 采用二维卷积神经网络对其进行特征提取。卷积操作在时间-频率二维平面上进行, 能够同时建模频谱结构及其随时间的动态变化。本文采用两层卷积结构, 每层均由卷积、批归一化及 ELU 激活函数组成, 从而逐步提取高层语义特征。随后, 通过自适应平均池化将特征压缩至全局表示:

$$F_{tf} \in R^{B \times f_2} \quad (23)$$

然后通过线性映射将其投影至统一嵌入空间:

$$F_{freq} = \text{Linear}(F_{tf}) \in R^{B \times D} \quad (24)$$

其中 D 表示嵌入维度(embedding size)。

综上, 所提出的 STFT 时频域分支能够从时频角度对脑电信号进行建模, 与时空域分支形成有效互补: 前者侧重刻画脑电信号的频谱结构及节律特征, 后者侧重建模时序动态与空间相关性。从神经生理角度来看, 该分支能够显式表征 μ 节律和 β 节律的能量变化, 从而更有效地捕获运动想象过程中 ERD/ERS 相关模式, 提升模型的判别能力。

2.4. 特征融合与分类

在获得时空域分支的两类特征表示(一阶时空特征、SPD 二阶协方差特征)与 STFT 时频域分支的时频特征后, 本文在特征维度上进行拼接融合, 将三类特征向量拼接为统一的联合表示。融合后的特征向量通过线性层映射至分类空间, 并经 Softmax 函数输出各类别的预测概率, 实现对运动想象任务的最终分类。

$$f_{fusion} = [f_{time} | f_{cov} | f_{freq}] \quad (25)$$

其中, f_{time} 、 f_{cov} 和 f_{freq} 分别表示一阶时空特征、SPD 二阶协方差特征和 STFT 时频特征。融合后的特征向量输入分类器进行最终的运动想象类别预测。通过该双分支协同建模框架, 模型能够在统一架构下同时学习时空模式、通道间统计依赖与频谱结构信息, 从而实现对 MI-EEG 信号的多层级联合表达。

3. 实验与结果分析

3.1. 实验设置与环境

本文模型的训练是在 Geforce RTX 4060 8GB 内存 GPU 上进行的, 仅使用 EEG 通道数据进行分析, 去除 EOG 通道。实验遵循 BCI Competition IV 官方评估协议, 采用数据集提供的训练集(Session T)和测试集(Session E)进行评估。其中, BCI IV-2a 数据集的训练集和测试集均包含 288 个试次; BCI IV-2b 数据集的训练集包含 400~440 个试次, 测试集包含 280~320 个试次。训练过程中, 进一步将训练集按 7:3 划分为训练集和验证集, 以监控模型的泛化性能。

3.2. 数据集描述与预处理

本文采用 BCI Competition IV-2a 和 2b 数据集进行实验。2a 数据集包含 9 名受试者执行左手、右手、双脚和舌头四类 MI 任务的 EEG 数据, 使用 22 个电极采集, 采样率为 250 Hz, 实验范式如图 5(a)所示。2b 数据集包含 9 名受试者执行左右手两类 MI 任务的 EEG 数据, 使用 3 个电极采集, 采样率为 250 Hz, 实验范式如图 5(b)所示。

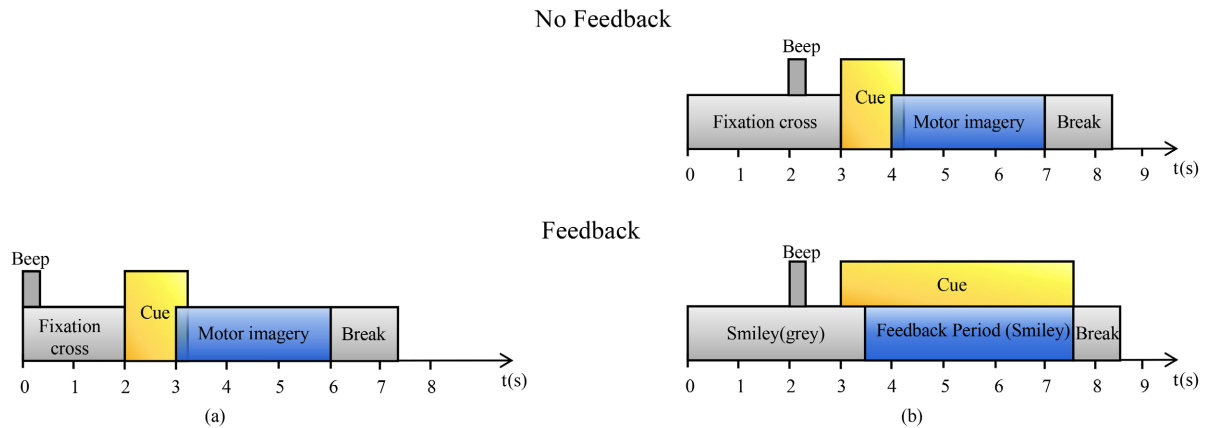


Figure 5. Schematic diagram of the motor imagery paradigm: (a) Dataset 2a, (b) Dataset 2b

图 5. 运动想象范式示意图 (a) 数据集 2a, (b) 数据集 2b

本文采用 Python 对 EEG 数据进行批量提取与预处理, 并将信号与对应标签进行关联。对于 BCI IV-2a 数据集, 选取[2 s, 6.5 s]时间窗口; 对于 BCI IV-2b 数据集, 考虑视觉反馈与无视觉反馈实验的差异, 选取[3 s, 7.5 s]时间窗口。两种数据集均包含 1126 个采样点。随后采用 Z-score 标准化对 EEG 信号进行归一化处理, 以减小数据波动和非平稳性的影响, 提高模型训练的稳定性与效率。

3.3. 评价指标

采用分类准确率(Acc)、一致性(Kappa 值)、混淆矩阵来评价模型的性能。此外, 为进一步验证模型间性能差异的统计显著性, 本文采用 Wilcoxon 符号秩检验进行配对比较。准确率被定义为一个分类模型在经过训练集的训练后, 在测试集上正确分类的样本数量与测试集中样本总数量之间的比值。这个指标用于评估模型在新数据上的表现, 即它能多大程度上正确识别出测试集中的样本。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (26)$$

一致性 Kappa 值是一种统计量, 可以用于评估模型性能, 也可以衡量分类的效果。Kappa 值的范围

从-1~1。值为 1 表示完全一致；值为 0 表示观察到的一致性与随机一致性相同；负值表示一致性比随机还差。在实践中，Kappa 值高于 0.6 通常被认为是好的，表明较高的一致性。

$$\text{kappa} = \frac{P_o - P_e}{1 - P_e}$$

(27)

式中： P_o 表示实验的准确率， P_e 表示随机猜测的准确率。

混淆矩阵是一种特别的表格布局，广泛应用于机器学习和统计分类中，用以可视化算法性能，特别是在多分类问题中。它是一种方阵，其中每一行表示实际的类别，而每一列表示预测的类别。通过混淆矩阵，可以直观地看到模型在各个类别上的分类效果，识别出哪些类别容易被误分类，以及误分类的具体情况。

3.4. 结果对比

使用公开数据集 BCICompetitionIV 2a 和 2b 进行实验结果的对比，2a 数据集为四分类，2b 数据集为二分类，同时选用了一些经典和先进的基于深度学习的方法，来评估该模型在运动想象任务中的表现。

表 1 展示了各模型在 BCI Competition IV-2a 数据集上的分类结果。在被动依赖设置下，DBMF-Net 取得了 85.87% 的平均分类准确率、0.8117 的 Kappa 值和 7.59 的标准差，在所有对比方法中表现最佳。相比 Deep ConvNet 和 EEGNet，准确率分别提高 13.65% 和 12.72%；相比 CTNet 和 TCANet，仍分别提升 3.35% 和 2.81%，表明所提模型具有更强的特征表征能力。进一步分析各受试者结果发现，DBMF-Net 在大多数受试者上均取得较优性能，尤其在基线性能较低的受试者上表现出更明显的提升，体现了其对复杂 EEG 信号的良好鲁棒性和判别能力。

Table 1. Comparison of motor imagery classification accuracy of different models on BCI-2a
表 1. 不同模型在 BCI-2a 运动想象分类准确率比较

Subject	Deep ConvNet (2017)	EEGNet (2018)	EEGInception (2020)	CTNe (2024) t	MSCFormer (2025)	TCANet (2025)	DBMF-Net -
A01	78.13	79.55	70.14	90.97	86.46	88.89	92.36
A02	45.14	61.11	53.82	73.61	61.46	70.14	80.90
A03	85.42	88.54	70.49	96.53	93.75	92.71	95.14
A04	67.01	71.53	68.40	84.72	80.90	79.86	85.07
A05	77.43	71.18	73.26	77.08	77.78	77.78	77.78
A06	53.13	59.03	53.82	64.24	69.44	74.31	72.22
A07	84.46	71.53	68.75	86.11	91.32	92.71	92.36
A08	78.13	80.56	72.22	84.38	83.68	85.76	88.54
A09	79.17	75.35	68.75	85.07	78.47	85.42	88.54
Avg acc	72.22	73.15	66.63	82.52	80.36	83.06	85.87*
S.D.	13.06	10.53	7.44	9.61	10.21	8.02	7.59
Kappa	0.6857	0.6718	0.5550	0.7670	0.7382	0.7742	0.8117

注：*表示 DBMF-Net 与次优方法(TCANet)的差异经 Wilcoxon 符号秩检验显著(p < 0.05)。

Table 2. Comparison of motor imagery classification accuracy of different models on BCI-2b
表 2. 不同模型在 BCI-2b 运动想象分类准确率比较

Subject	Deep ConvNet (2017)	EEGNet (2018)	EEGInception (2020)	CTNet (2024)	MSCFormer (2025)	TCANet (2025)	DBMF-Net -
B01	79.06	75.94	77.81	78.75	75.00	82.50	79.06
B02	65.71	66.07	66.07	71.07	68.75	70.71	73.57
B03	82.19	85.31	85.62	84.38	80.00	86.88	87.50
B04	97.50	98.44	98.12	97.19	98.44	97.81	96.56
B05	95.31	94.38	98.12	97.81	95.94	94.69	96.25
B06	80.62	84.38	87.81	87.81	85.00	87.19	90.94
B07	91.25	91.25	91.25	94.06	93.75	92.50	93.13
B08	92.19	94.69	95.31	94.69	94.38	95.94	96.25
B09	89.06	87.50	90.00	90.63	88.44	88.44	91.25
Avg acc	85.88	86.44	87.69	88.49	86.61	88.52	89.39
S.D.	9.98	10.19	10.35	9.03	10.30	8.32	8.11
Kappa	0.7176	0.7288	0.7538	0.7697	0.7322	0.7703	0.7878

在 BCI Competition IV-2b 数据集上的实验结果如表 2 所示。DBMF-Net 取得了 89.39% 的平均分类准确率、0.7878 的 Kappa 值和 8.11 的标准差，在所有对比方法中达到最优。与 Deep ConvNet、EEGNet、EEGInception、MSCFormer、CTNet 和 TCANet 相比，准确率分别提升 3.51%、2.95%、1.70%、2.78%、0.90% 和 0.87%，进一步验证了所提模型的有效性。同时从各受试者结果可以看出，模型在大多数受试者上均取得稳定提升，尤其是在 B02、B03 和 B06 等难分类受试者上表现更为明显，说明模型在处理低信噪比 EEG 信号时具有更强的鲁棒性。混淆矩阵如图 6 所示，DBMF-Net 在各类别上均取得了较高的分类准确率。

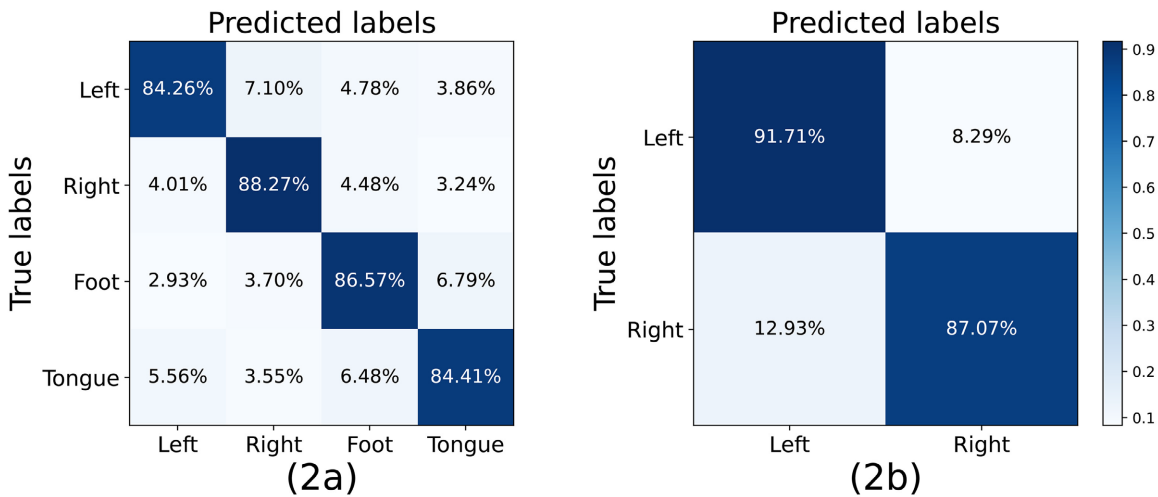


Figure 6. Confusion matrices on the BCI IV 2a and 2b datasets
图 6. BCI IV 2a、2b 数据集上混淆矩阵

对于 IV-2a 数据集，由于通道数较多且信号模式复杂，模型性能提升主要受益于多尺度时空特征建

模和二阶统计结构建模。其中，多尺度卷积与 Transformer 分别负责提取局部动态特征和长程时序依赖，而协方差分支进一步增强了对通道间相关性的刻画能力。相比之下，IV-2b 数据集通道数较少，运动想象相关的 μ 节律和 β 节律具有更强的判别性，因此 STFT 时频域分支对性能提升贡献更为显著，而协方差分支的作用相对减弱。尽管 DBMF-Net 在 IV-2b 数据集上取得了最高分类准确率，但相较于 TCANet 的性能提升未达到统计显著水平(Wilcoxon, $p = 0.250$)，这可能与数据集通道数较少及模型间性能差异有限有关。

总体而言，DBMF-Net 通过融合一阶时空特征、时频特征和二阶协方差特征，实现了对不同类型 EEG 数据的有效建模，在两类数据集上均表现出良好的适应性与稳定性。

3.5. 消融实验

为了验证所提出模型中关键模块的有效性，本文设计了消融实验，分别分析协方差(Cov)模块与短时傅里叶变换(STFT)时频域分支对模型性能的贡献。在所有实验中均采用相同的训练策略与参数设置，以保证公平比较。实验结果如表 3 所示。

Table 3. Ablation experiment on the Cov module and STFT module

表 3. Cov 模块、STFT 模块的消融实验

模型	Cov	STFT	平均准确率(%)		Kappa	
			2a	2b	2a	2b
Baseline	×	×	80.24	88.13	0.7366	0.7625
Baseline + Cov	√	×	85.33	89.12	0.8045	0.7824
Baseline + STFT	×	√	80.86	88.81	0.7449	0.7763
Baseline + Cov + STFT	√	√	85.87	89.39	0.8117	0.7878

Table 4. Impact of different attention mechanism combinations on model performance

表 4. 不同注意力机制组合对模型性能的影响

注意力机制组合	平均准确率(%)		Kappa	
	2a	2b	2a	2b
MHA	84.68	89.02	0.7957	0.7804
MHA + CBAM	85.49	88.81	0.8065	0.7762
MHA + ECA (本文)	85.87	89.39	0.8117	0.7878

为进一步验证通道注意力机制的必要性及其类型选择，本文设计了三组对比实验：仅使用多头注意力(MHA)、MHA 结合 CBAM 通道 - 空间注意力、以及 MHA 结合 ECA 轻量级通道注意力(本文方法)。如表 4 所示，在 BCI-2a 数据集上，仅使用 MHA 时准确率为 84.68%，加入 CBAM 后提升至 85.49% (+0.81%)，加入 ECA 后进一步提升至 85.87% (+1.19%)；在 BCI-2b 数据集上，ECA (89.39%)同样优于 CBAM (88.81%)和仅 MHA (89.02%)。上述结果表明，在多尺度时空卷积与 Transformer 全局注意力之后引入通道注意力机制是有必要的，能够进一步强化关键特征维度的表达。ECA 相比 CBAM 具有更优的性能，原因在于：经过多尺度卷积与 Transformer 融合后，特征已充分整合了时空判别信息，此时所需的注意力操作侧重于通道维度的自适应校准，而非空间结构建模。ECA 通过轻量级一维卷积实现跨通道交互建模，与该阶段的特征增强需求更为契合；而 CBAM 额外引入的空间注意力分支增加了计算冗余，可能对已融合的特征表示造成干扰，导致性能略有下降。

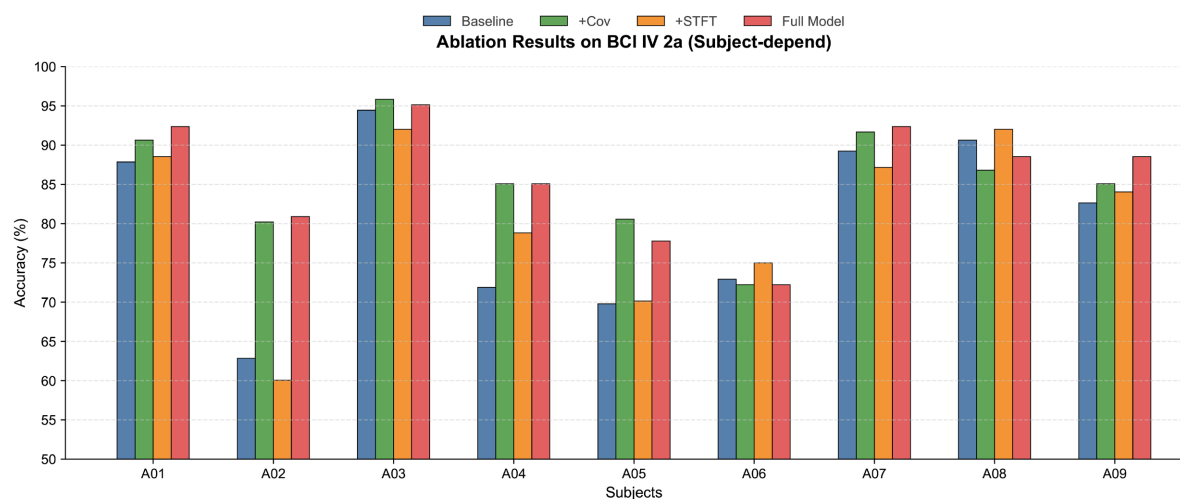


Figure 7. Ablation experiment for each subject on the 2a dataset

图 7. 在 2a 数据集上各被试者消融实验

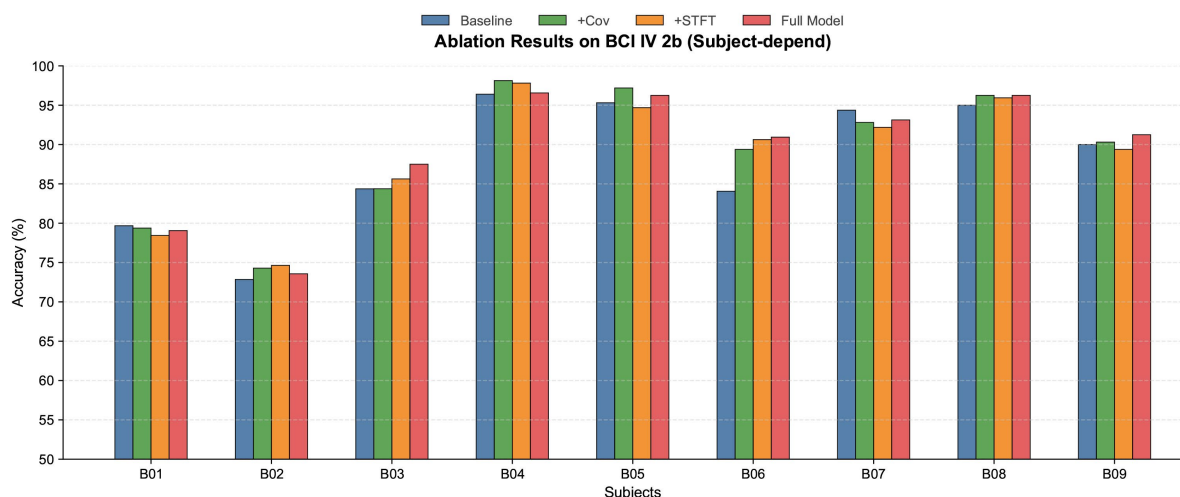


Figure 8. Ablation experiment for each subject on the 2b dataset

图 8. 在 2b 数据集上各被试者消融实验

图 7、图 8 展示了在 2a, 2b 数据集上, 不同消融模型在各被试上的分类准确率。从图中可以看出, Cov 模块与 STFT 分支在大多数被试上均带来性能提升, 说明所提出模块具有良好的稳定性与泛化能力。完整模型在绝大多数被试上取得最优或次优结果, 进一步验证了多分支特征融合的有效性。

4. 结论

本文提出了一种基于多特征融合的运动想象脑电分类模型 DBMF-Net。该模型通过双分支结构分别提取 EEG 信号的一阶时空特征、二阶协方差特征及时频域特征, 并引入黎曼几何学习对 SPD 协方差矩阵进行建模, 实现了多域特征的有效融合。

在 BCI Competition IV 2a 和 2b 数据集上的实验结果表明, DBMF-Net 的平均分类准确率分别达到 85.87% 和 89.39%, 优于现有基准方法。消融实验验证了协方差模块和 STFT 时频域分支各自的有效性, 二者融合后性能进一步提升。然而, 本文方法在通道数较少的数据集上优势有限, 未来工作将探索更轻量化的多特征融合策略, 并在更多类型的 BCI 数据集上验证模型的泛化能力。

参考文献

- [1] Kennedy, P.R. and Bakay, R.A.E. (1998) Restoration of Neural Output from a Paralyzed Patient by a Direct Brain Connection. *NeuroReport*, **9**, 1707-1711. <https://doi.org/10.1097/00001756-199806010-00007>
- [2] Acı, Ç.İ., Kaya, M. and Mishchenko, Y. (2019) Distinguishing Mental Attention States of Humans via an EEG-Based Passive BCI Using Machine Learning Methods. *Expert Systems with Applications*, **134**, 153-166. <https://doi.org/10.1016/j.eswa.2019.05.057>
- [3] Gao, Z., Dang, W., Liu, M., Guo, W., Ma, K. and Chen, G.R. (2021) Classification of EEG Signals on VEP-Based BCI Systems with Broad Learning. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **51**, 7143-7151. <https://doi.org/10.1109/tsmc.2020.2964684>
- [4] Li, Y., Pan, J., Wang, F. and Yu, Z. (2013) A Hybrid BCI System Combining P300 and SSVEP and Its Application to Wheelchair Control. *IEEE Transactions on Biomedical Engineering*, **60**, 3156-3166. <https://doi.org/10.1109/tbme.2013.2270283>
- [5] Jiao, Y., Zhang, Y., Chen, X., Yin, E., Jin, J., Wang, X., et al. (2019) Sparse Group Representation Model for Motor Imagery EEG Classification. *IEEE Journal of Biomedical and Health Informatics*, **23**, 631-641. <https://doi.org/10.1109/jbhi.2018.2832538>
- [6] Lemm, S., Blankertz, B., Curio, G. and Muller, K. (2005) Spatio-Spectral Filters for Improving the Classification of Single Trial EEG. *IEEE Transactions on Biomedical Engineering*, **52**, 1541-1548. <https://doi.org/10.1109/tbme.2005.851521>
- [7] Dornhege, G., Blankertz, B., Krauledat, M., Losch, F., Curio, G. and Muller, K. (2006) Combined Optimization of Spatial and Temporal Filters for Improving Brain-Computer Interfacing. *IEEE Transactions on Biomedical Engineering*, **53**, 2274-2281. <https://doi.org/10.1109/tbme.2006.883649>
- [8] Quadrianto, N., Cuntai, G., Dat, T.H. and Xue, P. (2007) Sub-Band Common Spatial Pattern (SBCSP) for Brain-Computer Interface. *Proceedings of the 2007 3rd International IEEE/EMBS Conference on Neural Engineering*, Kohala Coast, 2-5 May 2007, 2-7.
- [9] Ang, K.K., Chin, Z.Y., Wang, C., Guan, C. and Zhang, H. (2012) Filter Bank Common Spatial Pattern Algorithm on BCI Competition IV Datasets 2a and 2b. *Frontiers in Neuroscience*, **6**, Article No. 39. <https://doi.org/10.3389/fnins.2012.00039>
- [10] Balakrishnama, S. and Ganapathiraju, A. (1998) Linear Discriminant Analysis—A Brief Tutorial. Institute for Signal and Information Processing, 1-8.
- [11] Dagdevir, E. and Tokmakci, M. (2021) Truncation Thresholds Based Empirical Mode Decomposition Approach for Classification Performance of Motor Imagery BCI Systems. *Chaos, Solitons & Fractals*, **152**, Article ID: 111450. <https://doi.org/10.1016/j.chaos.2021.111450>
- [12] Schirrmester, R.T., Springenberg, J.T., Fiederer, L.D.J., Glasstetter, M., Eggenberger, K., Tangermann, M., et al. (2017) Deep Learning with Convolutional Neural Networks for EEG Decoding and Visualization. *Human Brain Mapping*, **38**, 5391-5420. <https://doi.org/10.1002/hbm.23730>
- [13] Lawhern, V.J., Solon, A.J., Waytowich, N.R., Gordon, S.M., Hung, C.P. and Lance, B.J. (2018) EEGNet: A Compact Convolutional Neural Network for EEG-Based Brain-Computer Interfaces. *Journal of Neural Engineering*, **15**, Article ID: 056013. <https://doi.org/10.1088/1741-2552/aace8c>
- [14] Xie, J., Zhang, J., Sun, J., Ma, Z., Qin, L., Li, G., et al. (2022) A Transformer-Based Approach Combining Deep Learning Network and Spatial-Temporal Information for Raw EEG Classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **30**, 2126-2136. <https://doi.org/10.1109/tnsre.2022.3194600>
- [15] Huang, Z. and Van Gool, L. (2017) A Riemannian Network for SPD Matrix Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, **31**, 2036-2042. <https://doi.org/10.1609/aaai.v31i1.10866>
- [16] Harandi, M., Salzmann, M. and Hartley, R. (2018) Dimensionality Reduction on SPD Manifolds: The Emergence of Geometry-Aware Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **40**, 48-62. <https://doi.org/10.1109/tpami.2017.2655048>
- [17] 周凯, 艾尔肯·亥木都拉. 基于ECA和三分支卷积融合网络的脑电信号解码研究[J]. 现代电子技术, 2024, 47(23): 89-97.
- [18] Jia, X., Song, Y., Yang, L. and Xie, L. (2022) Joint Spatial and Temporal Features Extraction for Multi-Classification of Motor Imagery EEG. *Biomedical Signal Processing and Control*, **71**, Article ID: 103247. <https://doi.org/10.1016/j.bspc.2021.103247>
- [19] 周成诚, 曾庆军, 杨康, 等. 基于高效通道注意力模块的运动想象脑电识别[J]. 计算机与现代化, 2023(12): 19-23.
- [20] Wang, R., Wu, X.J., Li, H., et al. (2023) Riemannian Self-Attention Mechanism for SPD Networks.

-
- <https://arxiv.org/abs/2311.16738>
- [21] Zhou, B., Wang, L., Xu, W. and Jiang, C. (2024) Multi-Scale Convolutional Attention and Riemannian Geometry Network for EEG-Based Motor Imagery Classification. *IEEE Access*, **12**, 79731-79740. <https://doi.org/10.1109/access.2024.3410036>
 - [22] Kalaganis, F.P., Georgiadis, K., Nousias, G., Oikonomou, V.P., Laskaris, N.A., Nikolopoulos, S., *et al.* (2025) Enhancing EEG-Based Neuromarketing with Attention Mechanism and Riemannian Features. 2025 *25th International Conference on Digital Signal Processing (DSP)*, Pylos, 25-27 June 2025, 1-5. <https://doi.org/10.1109/dsp65409.2025.11075055>
 - [23] Suh, Y. and Kim, B.H. (2021) Riemannian Embedding Banks for Common Spatial Patterns with EEG-Based SPD Neural Networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**, 854-862. <https://doi.org/10.1609/aaai.v35i1.16168>
 - [24] Pfurtscheller, G. (2001) Functional Brain Imaging Based on ERD/ERS. *Vision Research*, **41**, 1257-1260. [https://doi.org/10.1016/s0042-6989\(00\)00235-2](https://doi.org/10.1016/s0042-6989(00)00235-2)
 - [25] Lu, Y., Jiang, H. and Liu, W. (2017) Classification of EEG Signal by STFT-CNN Framework: Identification of Right-/Left-Hand Motor Imagination in BCI Systems. *Proceedings of the 7th International Conference on Computer Engineering and Networks—PoS (CENet2017)*, Volume 299, 1-8. <https://doi.org/10.22323/1.299.0001>