

控制与决策

Control and Decision

基于动态卷积和注意力机制的多域特征融合与运动想象解码

张美晨, 李明爱

引用本文:

张美晨, 李明爱. 基于动态卷积和注意力机制的多域特征融合与运动想象解码[J]. *控制与决策*, 2025, 40(6): 1873-1882.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2024.0942>

您可能感兴趣的其他文章

Articles you may be interested in

[周围神经MicroCT图像中神经束轮廓获取算法的改进](#)

An improved approach to obtain contours of fascicular groups from MicroCT images of peripheral nerve
控制与决策. 2021, 36(7): 1601-1610 <https://doi.org/10.13195/j.kzyjc.2019.1664>

[基于稀疏化神经网络的浮选泡沫图像特征选择](#)

Selection method for froth image characters based on sparse neural network
控制与决策. 2021, 36(7): 1627-1636 <https://doi.org/10.13195/j.kzyjc.2019.1788>

[Anchor-free的尺度自适应行人检测算法](#)

Anchor-free scale adaptive pedestrian detection algorithm
控制与决策. 2021, 36(2): 295-302 <https://doi.org/10.13195/j.kzyjc.2020.0124>

[基于多尺度特征表示的行人再识别](#)

Multi-scale feature representation for person re-identification
控制与决策. 2021, 36(12): 3015-3022 <https://doi.org/10.13195/j.kzyjc.2020.0952>

[结合注意力机制的循环神经网络复述识别模型](#)

Recurrent neural networks based paraphrase identification model combined with attention mechanism
控制与决策. 2021, 36(1): 152-158 <https://doi.org/10.13195/j.kzyjc.2019.0638>

基于动态卷积和注意力机制的多域特征融合与运动想象解码

张美晨, 李明爱[†]

(北京工业大学 信息科学技术学院, 北京 100124)

摘要: 基于卷积神经网络 (CNN) 进行运动想象 (MI) 解码是智能康复的一个研究热点, 而目前的解码方法难以针对脑电信号的个体差异性动态深度挖掘其蕴含的时-频-空特征, 影响解码效果. 鉴于此, 提出一种基于动态卷积和注意力机制的 CNN 模型 (DCAMNet). 首先, 使用滤波器组对每导原始脑电信号进行多频带划分, 同时输入特征提取模块; 然后, 由动态卷积块动态计算注意力权重, 获得有价值的时频信息, 再依次经空间卷积块和时间注意力块学习来挖掘空间信息和时间相关性, 以实现个性化时-频-空特征提取和融合; 最后, 由分类模块完成 MI 解码. 基于公开的 BCI Competition IV Dataset 2a 数据集对 9 名受试者进行 4 分类十折交叉验证实验, 取得了 79.17% 的平均准确率和 0.788 的 F_1 值. 实验结果表明, DCAMNet 能够自适应地关注和增强受试者个性化的特征, 可实现多域特征提取和融合, 相对于现有的流行方法在解码精度和泛化性能上具有一定优势.

关键词: 脑机接口; 脑电信号; 运动想象; 动态卷积; 注意力机制; 特征融合

中图分类号: TN911.7; TP202⁺.7; R318 文献标志码: A

DOI: 10.13195/j.kzyjc.2024.0942

引用格式: 张美晨, 李明爱. 基于动态卷积和注意力机制的多域特征融合与运动想象解码 [J]. 控制与决策, 2025, 40(6): 1873-1882.

Multi-domain feature fusion and motor imagery decoding based on dynamic convolution and attention mechanism

ZHANG Mei-chen, LI Ming-ai[†]

(School of Information Science and Technology, Beijing University of Technology, Beijing 100124, China)

Abstract: Motor-imagery (MI) decoding based on convolutional neural networks (CNNs) is a research hotspot in intelligent rehabilitation. However, current decoding methods struggle to dynamically explore the temporal-spectro-spatio features of electroencephalogram (EEG) signals, which vary across subjects and affect decoding performance. Therefore, we propose a dynamic convolution and attention mechanism based convolutional neural network (DCAMNet). First, a filter bank divides raw EEG signals from each channel into multi-spectral bands, which are simultaneously input into the feature extraction module. Then, the dynamic convolution block calculates attention weights to obtain valuable temporal-spectral information, which is passed through the spatial convolution block and the temporal attention block to learn spatial information and temporal correlations, achieving discriminative temporal-spectro-spatio feature extraction and fusion. Finally, the classification module completes the MI decoding. Based on the public BCI Competition IV Dataset 2a, a four-class ten-fold cross-validation experiment on nine subjects achieves an average accuracy of 79.17% and an F_1 score of 0.788. The results show that the DCAMNet can adaptively focus on and enhance discriminative features caused by inter-subject variability, achieving multi-domain feature extraction and fusion. Compared to current popular methods, it has advantages in decoding accuracy and generalization performance.

Keywords: brain-computer interface; electroencephalography; motor imagery; dynamic convolution; attention mechanisms; feature fusion

0 引言

脑卒中以其高发病率和高致残率成为全球主要

健康问题之一. 脑卒中后患者运动皮质区神经元损

伤, 约 3/4 的存活者伴有肢体运动控制障碍. 脑-机接

收稿日期: 2024-08-08; 录用日期: 2025-01-14.

基金项目: 国家自然科学基金项目 (62173010).

责任编辑: 彭木根.

[†]通信作者. E-mail: limingai@bjut.edu.cn.

口 (BCI)^[1-2] 可以实现大脑与外部设备间的直接通信和控制, 特别是运动想象 (MI)^[3-4] 作为一种重要的 BCI 范式, 在运动功能障碍康复训练中具有潜在的应用前景, 而基于运动想象脑电图 (MI-EEG) 准确解码患者的运动意图尤为重要。

深度学习, 特别是卷积神经网络 (CNNs)^[5], 通过权重共享以更少的参数捕获 MI-EEG 中更高层次的特征, 展现出强大的特征提取和分类能力, 是 MI-EEG 重要的解码工具。如 Koles^[6] 提出的共空间模式 (CSP) 算法, 使用空间滤波器组, 提取了一组 EEG 中最大与最小方差时间序列, 通过主成分分析展示其空间拓扑, 从而提取了 EEG 的不同成分^[7]; Ang 等^[8] 提出了滤波器组共空间模式 (FBCSP), 使用多个切比雪夫 II 型滤波器, 通过 CSP 算法进行空间滤波并完成了分类。这些方法虽然简单且计算效率高, 但是, 在设计空间滤波器前, 特定的带通滤波器在一定程度上牺牲了灵活性和自适应性。受到 CSP 算法的启发, Schirmermeister 等^[9] 提出了浅层卷积网络 (ShallowConvNet) 和深度卷积网络 (DeepConvNet), 卷积层的数量由 2 层增加至 5 层, 从原始 EEG 中解码与 MI 任务相关的信息; Ma 等^[10] 提出了滤波器组卷积网络 (FBCNet), 使用多视角数据表示, 通过空间滤波提取频谱和空间特征, 并使用方差层有效地聚合了 EEG 的时域信息; Lawhern 等^[11] 提出了 EEGNet, 分别使用卷积核大小为 1×64 、 $C \times 1$ (C 为 EEG 通道数) 的深度卷积学习 EEG 的时空信息, 再通过可分离卷积优化组合, 构建了一个适用于 EEG 的模型; Santamaria-Vazquez 等^[12] 提出了启发式 EEGNet (EEG-Inception), Inception 块对每个 EEG 通道进行 3 种不同时间尺度的处理 (分别为 64×1 、 32×1 和 16×1), 以捕捉不同尺度上特征间的依赖关系。

MI-EEG 是复杂的高维信号, 具有丰富的时间、空间和频谱特征^[13], 有效的 MI 信息通常出现在特定的通道位置、频段和时间间隔, 且不同受试者的 EEG 存在个体差异, 而单尺度卷积核的特征提取能力受限。Zhang 等^[14] 提出了一种包含多尺度特征的分层分割块 (Hierarchical-Split), 同时提出了分层分割残差网络 (HS-ResNet), 在一个残差块内包含许多分层分割和连接, 提取更丰富特征的同时避免了冗余信息; Michielli 等^[15] 提出了一种新的基于长短期记忆单元 (LSTM) 的递归神经网络架构 (RNN), 从单通道 EEG 中提取了 55 个时域和频域特征; Li 等^[16] 提出了一种通道投影混合尺度卷积神经网络 (CP MixedNet), 学习初级时空特征后, 使用混合尺度卷积块来捕获混合尺度的时序信息, 扩大了网络接收

域的同时减少了训练参数; Cao 等^[17] 提出了 IFBCLNet, 将滤波器组与 CNN 和 LSTM 块相结合, 实现了时空域和频域的特征提取。然而, 复杂深度神经网络的超参数过多, 结构过于复杂, 模型需要大量数据集进行训练, 否则会陷入过拟合和局部最优^[18]。

近年来, 注意力机制在 CNN 中得到了广泛应用。注意力机制使得模型能够有选择地识别目标特征, 忽略一些对分类无用的信息, 可以引入至多尺度、多分支网络中。Hu 等^[19] 提出了挤压-激励注意力模块 (SENet), 自动学习通道权重, 显著提高了特征提取效果; Park 等^[20] 提出了瓶颈注意力模块 (BAM) 和卷积块注意力模块 (CBAM)^[21], 有效结合了空间和通道注意力; Zhang 等^[14] 提出了高效金字塔切分注意力网络 (EPSANet), 在不显著增加计算负担的情况下提高了分类性能; Li 等^[22] 提出了一种时间-频谱的挤压-激励特征融合网络 (TS-SEFFNet), 同时, 利用深度时域卷积 (DT-Conv) 和多频谱卷积 (MS-Conv) 提取了高维时间信息和频谱特征, 并基于 SENet 有效地提取了最具鉴别力的特征和抑制信息量较少的特征。最近的研究建议在网络内核层次上使用注意力机制^[23]。Barmpas 等^[24] 提出了基于动态卷积的深度学习框架, 使用因果推理刻画 MI 任务中由受试者差异引起的所有可能的分布偏移, 注意力网络学习不同受试者间的差异性, CNN 中的每个卷积块使用 K 个并行可训练的卷积层, 其深度模型在跨受试者实验中获得更好的泛化性能。可见, 利用动态卷积和注意力机制应对受试者个体差异, 并确保特征的完整性对于提升 MI 解码性能具有很好的研究价值。

鉴于此, 本文提出一种基于动态卷积和注意力机制的 CNN 模型, 记为 DCAMNet。首先, 利用动态卷积块动态调整卷积核参数, 获得有判别性的时频信息; 然后, 依次基于空间卷积块和时间注意力块相继学习空间信息和时间依赖关系, 生成更具判别力的时-频-空多域特征, 以增强网络对个体差异的自适应性, 改善 MI 解码性能; 最后, 基于公共数据集验证所提出网络的有效性。

1 研究方法

1.1 整体模型

所提出 DCAMNet 主要由 3 部分组成: 多视角数据表示模块、特征提取模块以及分类模块, 充分提取 MI-EEG 的频率、空间和时间特征。具体而言, 首先, 使用滤波器组对原始 EEG 进行多频带划分, 结合后输入至动态卷积块, 进行时频特征提取, 其中

SENet 在 MI-EEG 的频谱维度上突出所有不同频段有价值的频谱信息, 时间卷积层将时频特征最优地融合在一起, 使得所提出模型能够关注到时频维度上最有价值的特征; 然后, 空间卷积块进行空间特征提取, 沿通道维度上学习空间信息和时频信息; 最后,

将时间序列分割为一系列非重叠时间窗口, 通过时间注意力块挖掘时间相关性, 聚合所有窗口的特征并输入至全连接层执行 MI 任务的概率预测. 网络的整体结构如图 1 所示, 网络参数设置具体细节如表 1 所示.

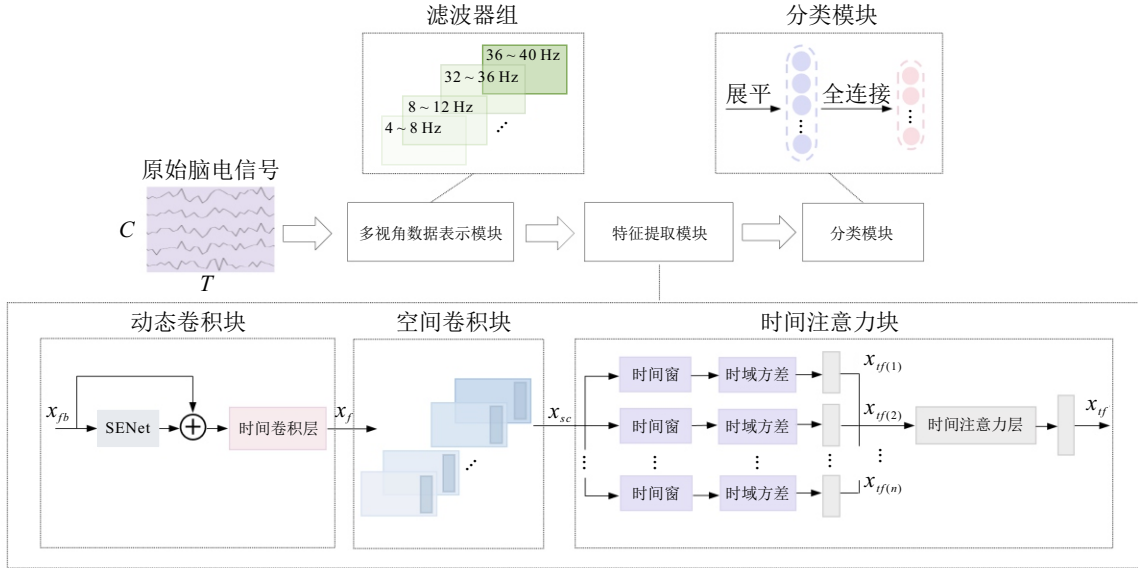


图1 DCAMNet 网络整体框架

表1 DCAMNet 网络参数设置

网络层	滤波器组数	尺度	参数	输出
输入层	-	-	-	(C, T)
滤波器组	-	-	-	(N_b, C, T)
动态卷积层	-	-	-	(N, C, T_1)
空间卷积层	m	$(C, 1)$	$m \times N \times C$	$(m, 1, T_1)$
时间滑窗层	-	-	-	$(m, T_1/w)$
时间注意力层	h	T_1/w	$h \times (T_1/w)$	$(m/h, h, 1)$
展开层	-	-	-	$(m/h \times h)$
全连接层	-	-	$N_c \times (m/h \times h)$	N_c

注: C 为 EEG 通道数, T 为时间采样点, N_b 为频带滤波器个数, N 为时间滤波器个数, T_1 为经动态卷积后的时间采样点, m 为空间滤波器个数, w 为窗口长度, h 为时间注意力层通道数, N_c 为类别数.

1.2 多视角数据表示

不同 MI 类别对应的感觉运动节律模式可能会不同, 本文通过一组带通滤波器 $F = \{f_i\}_{i=1}^{N_b}$ 将原始 EEG 数据 $X \in \mathbb{R}^{C \times T}$ 分为多个局部窄带 EEG 信号, 建立多频带特征表示 $x_{fb} \in \mathbb{R}^{N_b \times C \times T}$ 作为特征提取模块的输入. 其中: N_b 为带通滤波器的个数, 每个频带表示一个经滤波后在特定窄带内的 EEG 信号; C 、 T 分别为 EEG 通道数和时间采样点数. 滤波器组 F 由 9 个带宽为 4 Hz, 过渡带宽为 2 Hz, 阻带纹波为 30 dB 的切比雪夫 Type II 滤波器组成, 具体频率范围覆盖为 4 ~ 40 Hz (即 4 ~ 8 Hz, 8 ~ 12 Hz, ...,

36 ~ 40 Hz).

1.3 特征提取模块

特征提取模块结合多种机制来处理 EEG 信号, 主要由动态卷积块、空间卷积块和时间注意力块 3 部分组成, 能够充分提取频率、空间和时间特征.

1) 动态卷积块: 多视角特征表示 x_{fb} 作为输入信号, 经过 SENet 以突出所有不同频段有价值的频谱信息, 时间卷积层将时频特征最优地融合在一起. 模块通过动态卷积学习 MI-EEG 在不同频率下的特征, 并通过 SENet 识别最有助于分类的频段, 从而区分 MI 任务. 动态卷积块框架如图 2 所示.

动态卷积块使用 K 个并行卷积核 (W_k, b_k) 对每个输入 x_{fb} 进行动态聚合, 如下所示:

$$\begin{aligned} \tilde{W}(x_{fb}) &= \sum_{k=1}^K \pi_k(x_{fb}) \tilde{W}_k, \\ \tilde{b}(x_{fb}) &= \sum_{k=1}^K \pi_k(x_{fb}) \tilde{b}_k. \end{aligned} \quad (1)$$

其中: π_k 为随输入 x_{fb} 变化的注意力权重, 通过 SENet 计算; $\pi_k(x_{fb}) \in [0, 1]$; $\sum_{k=1}^K \pi_k(x_{fb}) = 1$.

对于一个 MI 输入 x_{fb} , SENet 首先使用全局平均池化层对输入数据的频谱信息进行压缩. 然后传递至两个全连接层: 第 1 个全连接层用于激活修正

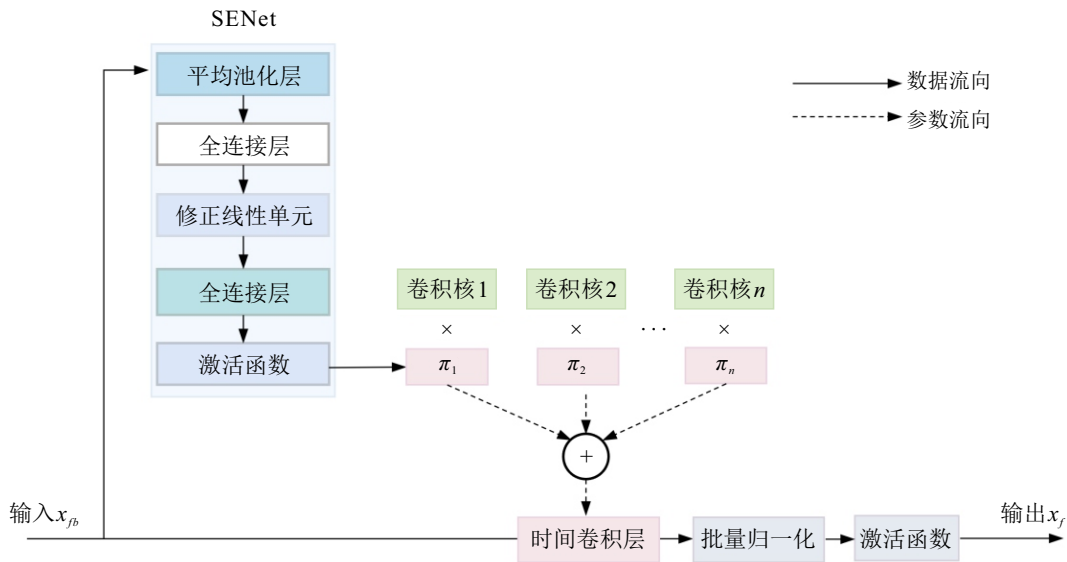


图2 动态卷积块框架

线性单元 (ReLU), 第 2 个全连接层使用激活函数 (Softmax) 对注意力权重进行归一化处理, 从而生成注意力权重, 并基于温度系数 τ_d 计算 Softmax, 以控制注意力权重的稀疏性, 如下所示:

$$\pi_k = \frac{\exp(z_k/\tau_d)}{\sum_j \exp(z_j/\tau_d)}, \quad (2)$$

其中 z_k 为第 2 个全连接层的输出. 生成注意力权重 π_k 后, 计算聚合卷积, 并输入至一个批量归一化 (BN) 层, 使用 Softmax 激活, 该卷积层输出一个时频特征 $x_f \in \mathbb{R}^{N_b \times C \times T_1}$.

2) 空间卷积块: MI-EEG 通常由放置于多个脑区的电极记录, 空间信息学习旨在对 EEG 数据的不同通道进行空间信息编码. 模块在空间维度上 (即不同 EEG 通道间) 进行卷积来捕捉 MI 任务涉及的特定脑区活动, 如中央前回区域相关的运动皮层活动等, 进而识别 MI 任务激活的大脑区域, 有助于提取不同脑区的特征. 空间卷积块将时频特征 x_f 输入一个具有 m 个输出通道的空间卷积层, 该层作为所有通道上的 m 个空间滤波器. 卷积核大小设置为 $C \times 1$, 以整合所有 EEG 通道中的不同信息, 也有助于整合输入的时频信息. 在空间卷积层后, 进一步应用 BN 层和指数线性单元 (ELU). 因此, 空间卷积块的整个过程输出 m 个时间序列 $x_{sc} \in \mathbb{R}^{m \times 1 \times T_1}$.

3) 时间注意力块: 不同 MI 任务可能在持续时间和信号波动方面表现出不同的特征, 通过时间维度上的轻量卷积来提取时间序列特征, 如信号的波动或时间模式, 并通过计算信号的方差来捕捉任务期间信号强度的波动, 可以帮助理解 MI 任务的时序特征, 如每个任务的持续时间和频率模式变化. 时间

注意力块将时间序列 x_{sc} 分割为一系列非重叠时间窗口, 通过深度可分离卷积层来挖掘时间相关性, 聚合后输入至分类模块. 时间注意力块框架如图 3 所示.

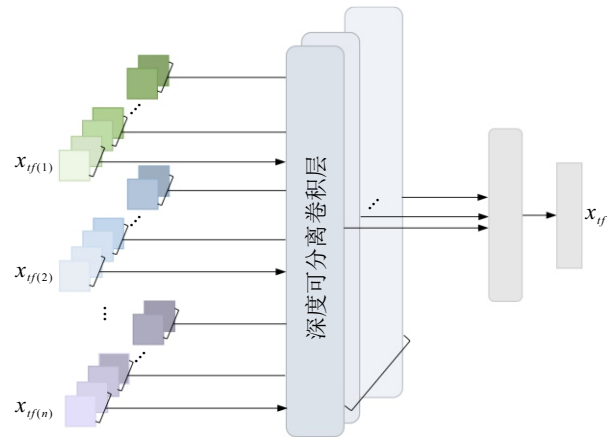


图3 时间注意力块框架

为了捕捉不同时间段 EEG 信号呈现出的不同 MI 相关模式, 采用窗口长度为 w 的 n 个不重叠时间窗口, 对沿时间维度的输入 x_{sc} 进行分割. 其中: $n = \lfloor T/w \rfloor$, $\lfloor \cdot \rfloor$ 表示向下取整操作. 采用时间方差层计算第 k 个时间窗口中大小为 w 的第 i 个时间序列沿时间维度的方差, 以提取每个时间窗口中的判别特征, 计算如下所示:

$$x_{tf}^{(k)}(i, 1, 1) = \frac{1}{w} \sum_{t=w \times k}^{(k+1) \times w - 1} (x_{sc}(i, 1, t) - \mu(i, 1, k))^2, \quad (3)$$

其中 $\mu(i, 1, k)$ 为第 k 个时间窗口内 $x_{sc}(i, 1, t)$ 的时间均值. 每个时间窗口生成的特征表示为 $\{x_{tf}^{(1)}, x_{tf}^{(2)}, \dots, x_{tf}^{(n)}\}$, $x_{tf}^{(i)} \in \mathbb{R}^{m \times 1 \times 1}$. 沿通道维度将所有特征 $\{x_{tf}^{(1)}, x_{tf}^{(2)}, \dots, x_{tf}^{(n)}\}$ 分为 m/h 组, 每组具有 h 个通道.

同组的特征输入至具有 h 个输出通道的深度可分离卷积层, 独立地对每个通道执行卷积操作, 每组分配自注意力权重, 在时间维度上通过 Softmax 进行归一化. 不同组中计算的特征被连接起来生成输出特征 $x_{tf} \in \mathbb{R}^{m \times 1 \times 1}$. 时间注意力块计算输出通道 c 的公式如下所示:

$$x_{tf}(c, 1, 1) = \sum_{j=1}^n W((c \bmod h), 1, j) x_{tf}^{(j)}(c, 1, 1). \quad (4)$$

这里: $W \in \mathbb{R}^{h \times 1 \times n}$ 为深度可分离卷积的权重, “mod” 为取模运算. 最后将输出特征 x_{tf} 用于分类.

1.4 分类模块

模块在上述特征提取的基础上进行分类, 将提取的频谱、空间和时间 3 个维度的 EEG 特征共同构成对 MI 任务的全面表示, 最终的分层学习不同 MI 任务的特征模式与类别间的映射关系, 从而得到最终的解码结果. 具体操作如下: 首先将特征 x_{tf} 展平为一维特征向量, 然后将该向量送入全连接层, 最后将输出最大值的标签作为最终结果.

2 运动想象实验数据

2.1 数据集介绍

选用 BCI Competition IV Dataset 2a 公开数据集进行实验研究. 该数据集由 Graz University of Technology 提供, 是广为采用并用于运动想象任务研究中的数据集, 包含 9 名健康受试者 (用 $S_1 \sim S_9$ 表示) 的 EEG 数据, 包括 4 类不同的 MI 任务: 左手、右手、双脚、舌头. EEG 由 25 个 Ag/AgCl 电极记录, 3 个 EOG 电极不用于解码, 采样率为 250 Hz, 并在 0.5 ~ 100 Hz 间进行带通滤波. 每个受试者进行两组实验. 每组实验有 288 个试次, 每次实验包含 48 次 MI 任务 (4 个类别各进行 12 次). 一组实验数据用于模型训练, 另一组则用于评估性能. 每次实验持续 7 s, 具体流程如下.

step 1: 0 ~ 2 s: 休息阶段. 屏幕上显示一个固定的十字 (+), 受试者保持放松.

step 2: 2 ~ 3.25 s: 准备阶段. 屏幕上显示一个箭头提示符, 指示下一步任务的方向 (左手、右手、双脚、舌头).

step 3: 3 ~ 6 s: 执行阶段. 受试者根据提示符进行相应的 MI 任务.

2.2 预处理方法

对数据进行以下预处理以去除与 MI 任务无关的信息.

1) 删除 3 个 EOG 通道信号, 仅使用 22 个 EEG

通道, 即通道数目 $C = 22$;

2) 使用五阶巴特沃斯带通滤波器对原始 EEG 数据进行 4 ~ 40 Hz 的带通滤波, 以截取与 MI 任务相关波段;

3) 截取相对于 MI 任务开始时的相同长度时间窗口 $[-0.5 \text{ s}, 4 \text{ s}]$ 来提取每个试次的 EEG 数据, 对应采样点数 $T = 1125$, 确保提取信号主要包含与 MI 任务相关的活动.

3 实验与结果

本节在公开的 BCI Competition IV Dataset 2a 脑电图数据集上进行跨受试者实验, 以验证所提出方法的有效性. 首先, 通过消融研究验证 DCAMNet 的有效性; 然后, 通过一系列对比实验分析模型参数对于整体表现的影响; 最后, 将所提出方法与其他先进的 MI 识别方法进行性能比较.

3.1 实验设置

本文实验在 PyTorch v1.5.1 框架下进行, Python 版本为 3.7.0. 实验环境系统为 Windows 操作系统, 处理器配置为 Intel(R) Xeon(R) CPU E5-2683 v3 @ 2.00 GHz. 选用 Adam 优化器, 用于优化目标函数, 学习率设置为 0.001.

为了克服数据划分对于实验结果的影响, 采用十折交叉验证对每位受试者的解码性能进行评估. 具体而言, 将数据集划分为 10 个子集, 每次选择其中 1 个子集作为测试集, 其余 9 个子集作为训练集. 在整个评估过程中, 数据的划分方式保持顺序和类别平衡不变. 最终通过对 10 次评估结果的平均值进行计算, 以获得每位受试者的整体解码表现. 训练时 batch 值设定为 64, 若在连续 50 个 epoch 内准确度没有提升, 则停止训练, 且最大 epoch 设置为 500, 以减少冗余的训练过程, 加快训练速度.

3.2 评价指标

选择准确率 (Accuracy)、 F_1 值 (F_1 -score)、Kappa 系数 (Cohen's Kappa) 和计算时间作为所提出网络整体性能的评价指标. 准确率表示正确分类的比例; F_1 值为精确率 (Precision) 与召回率 (Recall) 的调和平均数, 尤其为类不平衡时提供更全面的性能评估; Kappa 系数是一个用于衡量分类器一致性或分类结果可靠性的重要指标. 准确率、 F_1 值和 Kappa 系数计算公式如下所示:

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}}, \quad (5)$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (6)$$

$$\text{Kappa} = \frac{P - P_e}{1 - P_e}, \quad (7)$$

$$P_0 = \frac{\sum_{i=1}^n a_{ij}}{N}, \quad (8)$$

$$P_e = \frac{\sum_{i=1}^n a_{i+} \times a_{+i}}{N^2}. \quad (9)$$

其中: Precision为精确率; Recall为召回率; TP和TN分别表示预测为正类的正样本和预测为负类的负样本; FP和FN分别表示预测为正类的负样本和预测为负类的正样本; 在Kappa系数的计算中, $a_{i+} = \sum_j a_{ij}$ 表示真实类别为*i*的样本数量, 即混淆矩阵中第*i*行的总和, $a_{+j} = \sum_i a_{ij}$ 表示预测类别为*j*的样本数量, 即混淆矩阵中第*j*行的总和, 这里*n*为类别数, *N*为样本总数.

3.3 实验结果与分析

所提出网络包含多视觉数据表示、动态卷积块、空间卷积块和时间注意力块, 每个模块均对网络整体性能有不同的影响. 此外, 每个模块中包含不同的可训练参数, 这些参数的大小也是影响模型分类精度的重要因素. 因此, 为了验证所提出网络的有效性和合理性, 进行一系列的消融和对比实验.

3.3.1 基于不同频段划分的多视觉数据表示对于解码性能的影响

MI-EEG 解码常用的频率范围为4 ~ 40 Hz, 不同频段的选择和分割方式会影响分类性能, 因此, 基于不同频段划分的多视角数据表示对 DCAMNet 解码性能的影响进行评估, 频段划分取值范围为{2, 4, 6, 9}, 如表2所示. 评估指标为可训练参数的数量、平均准确率和 F_1 值.

表2 基于不同频段划分方式的多视觉数据表示对于 DCAMNet 解码性能的影响

频段划分数	可训练参数数量	平均准确率/%	F_1 值
2	16896	57.02	0.548
4	63360	66.55	0.653
6	140800	69.17	0.681
9	142912	79.17	0.788

表2结果表明, 随着频段划分数量的增加, 模型的解码性能逐渐提升, 表现在可训练参数的数量、平均准确率和 F_1 值的线性增长. 频段划分数从2增加至9, 模型的平均准确率从57.02%提高至79.17%,

F_1 值也从0.548提升至0.788. 这表明更精细的频段划分能够更好地捕捉 EEG 信号中的频域特征, 增强了模型对运动想象任务的区分能力. 特别是在频段划分数为9时, DCAMNet 的解码性能达到了最优状态. 这可能是因为更细粒度的频段划分能够提供更多的频率信息, 从而令动态卷积模块更有效地捕获不同运动想象任务下的脑电特征. 此外, 较高的频段划分数可能更有助于网络在不同频率上发现更具区分力的特征, 从而提高分类的准确性和泛化能力.

总体而言, 基于不同频段划分的多视角数据表示能够有效提升 DCAMNet 对 MI 任务的解码性能, 而频段划分数为9时模型表现最优, 因此, 选用9个频段划分数作为动态卷积模块的输入.

3.3.2 基于不同注意力机制的动态卷积块对于解码性能的影响

多种注意力机制, 如 CBAM、EPSANet、CA (coordinate attention) 模块等, 已在深度学习得到了较多应用, 下面通过实验来验证本文在动态卷积块中选择 SENet 的正确性和有效性. 在保持其余模块参数相同的情况下, 保证选用不同注意力机制的动态卷积块的参数控制在一定的范围内, 基于不同注意力机制的动态卷积块对 DCAMNet 解码性能的影响进行实验和评估, 评估指标为可训练参数数量、平均准确率和 F_1 值, 结果如表3所示.

表3 基于不同注意力机制的动态卷积块对于 DCAMNet 解码性能的影响

注意力机制	可训练参数数量	平均准确率/%	F_1 值
CBAM	132406	64.78	0.642
EPSANet	131328	70.10	0.689
CA	142912	77.04	0.767
SENet	142912	79.17	0.788

表3结果表明, 在动态卷积块替换的4种注意力网络中, SENet 效果最为显著, 在不增加更多的可训练参数数量的前提下, 达到了79.17%的平均准确率和0.788的平均 F_1 值. 本文认为导致这一结果的可能原因是, SENet 能够通过自适应地调整各通道的权重, 更好地突出不同频段与通道间的相关性, 从而帮助模型在处理多频段的运动想象任务时, 能够精确捕捉关键特征. 此外, SENet 的结构相对简单, 在引入较少额外计算开销的同时, 显著提升了解码性能; 相比之下, CBAM 虽然在时序信号处理方面表现较好, 但是, 在多频段特征提取中略显不足; EPSANet 擅长捕捉空间与通道维度的相关性, 但是, 其复杂的结构可能会导致特征融合效果不如 SENet; CA 虽然在复杂空间关系的捕捉上有所提升, 但是, 其在动态

卷积块中的表现未能达到与 SENet 同样的效果. 总之, SENet 在动态卷积块中表现出更好的通道特征提取能力, 使其更适合于所提出模型. 因此, 在模型的动态卷积块设计中, 选择了 SENet 网络来突出不同频段的信息.

3.3.3 网络消融实验

本实验比较了动态卷积块、空间卷积块、时间注意力块对于 DCAMNet 在跨受试者实验中解码性能的影响. 分别对动态卷积块 (D)、空间卷积块 (S)、时间注意力块 (T)、动态卷积块 + 空间卷积块 (DS)、动态卷积块 + 时间注意力块 (DT)、空间卷积块 + 时间注意力块 (ST)、动态卷积块 + 空间卷积块 + 时间注意力块 (DST) 7 种网络结构进行评估, 结果如图 4 所示. 图 4 结果表明, 动态卷积块、空间卷积块和时间注意力块均对网络分类准确率和泛化性能产生了不同影响, 当 3 个模块相结合为 DCAMNet 最优结构时, 网络性能综合达到最佳.

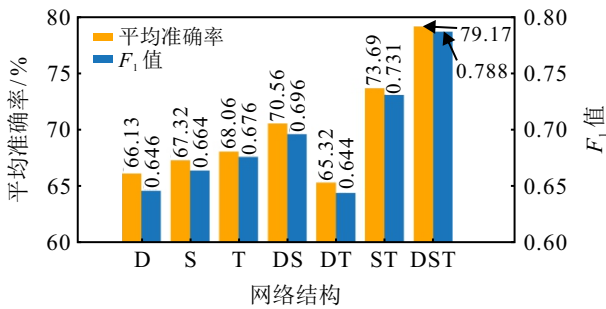


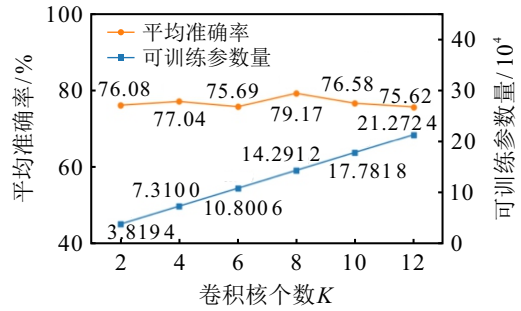
图4 DCAMNet 消融实验结果

3.3.4 网络参数优选实验

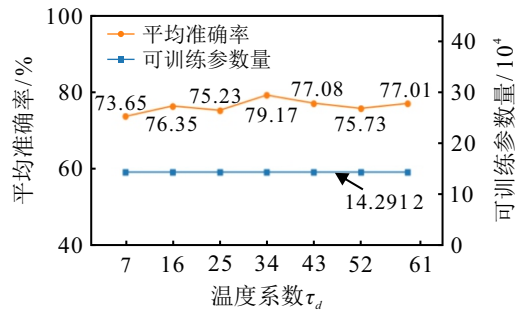
动态卷积中有两个重要的超参数: 卷积核个数 K 和温度系数 τ_d . 同时, 空间滤波器个数 m 以及时间窗口长度 w 也是影响网络性能的重要超参数. 在跨受试者实验中, 不同的参数取值决定了网络的复杂程度, 也影响了分类精度, 故本节经过一系列对比实验来阐明各参数确定的合理性和可靠性. 图 5 为在分别固定另外 3 个参数的情况下, 不同取值范围的参数 K 、 τ_d 、 m 和 w 对于网络整体解码性能的影响, 评估指标为平均准确率和可训练参数的数量.

1) 动态卷积块参数确定.

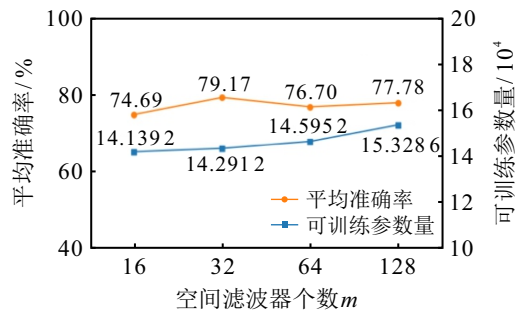
超参数 K 决定了动态卷积核的尺度. 令 $K = \{2, 4, 6, 8, 10, 12\}$, 得到的平均准确率和可训练参数数量结果如图 5(a) 所示. 由图 5(a) 可见: 随着 K 的增加, 网络可训练参数数量单调递增, 平均正确率则呈现波动趋势, 均值为 76.70%. 当 $K = 8$ 时, 动态卷积使得模型具有更高的表征能力, 平均准确率达到最高. 然而, 随着 K 值的增加, 模型的复杂度也随之增大, 网络更易出现过拟合现象, 准确性不再提高. 在此情况



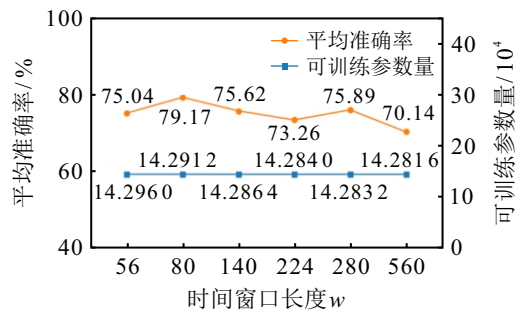
(a) 卷积核个数 K 的影响



(b) 温度系数 τ_d 的影响



(c) 空间滤波器个数 m 的影响



(d) 时间窗口长度 w 的影响

图5 参数 K 、 τ_d 、 m 和 w 对于平均准确率和可训练参数数量的影响

下, 同时优化所有卷积和注意力参数变得更具挑战性, 这也表明了 $K = 8$ 时参数量为 142912 是较为合适的.

温度系数 τ_d 决定了注意力权重的稀疏性. τ_d 可设置为一个固定值, 也可在训练初期将 τ_d 设置为较大的值, 然后, 在后期训练中逐渐减小, 这是动态卷积训练成功的关键因素. 令 $\tau_d = \{7, 16, 25, 34, 43, 52\}$, 得到的实验结果如图 5(b) 所示. 显然, 动态温度的设置可以使得模型取得良好的效果. 理论上, τ_d 越大,

注意力权重越均匀,使得每层中的所有卷积核均能够得到较好的优化,但是,平均准确率总体呈现先上升后下降的趋势.其中: $\tau_d = 34$ 时,平均准确率达到最高;当 $\tau_d > 34$ 时,平均准确率出现了下降趋势.

2) 空间卷积块参数确定.

在空间卷积块中空间滤波器个数 m 用于平衡网络中空间卷积的容量与计算效率,控制模型的空间和频谱学习能力.令 $m = \{16, 32, 64, 128\}$,实验结果如图5(c)所示.图5(c)结果显示, m 在很大程度上决定了可训练参数的总数,影响了模型的解码性能和计算效率.当 $m = 32$ 时,平均准确率相对最高;当 $m > 32$ 时,模型可训练参数增加,平均准确率反而呈现下降趋势;当 $m > 64$ 时,网络可训练参数急剧增加,网络复杂度随之增加,但是,平均准确率基本稳定.

3) 时间注意力块参数确定.

时间窗口长度 w 在时频特征提取和特征融合时起着重要作用,不同的 w 值从时间序列中捕获的信

息量不同,导致DCAMNet的整体性能有所变化.对于不同尺度的卷积核,令 $w = [50, 1125]$,对于网络性能的影响如图5(d)所示.由图5(d)可见:整体而言,平均准确率呈波动下降趋势;当 $w = 80$ 时,网络性能表现最好,整体性优于其他窗口长度;当 $w > 80$ 时,平均准确率呈波动下降趋势;当 $w = 560$ 时,过大的时间窗口尺寸会降低解码精度,这表明所提出网络在使用较大时间窗口长度时不能充分探索时间依赖关系.

3.3.5 跨受试者实验结果

基于DCAMNet对9位受试者进行跨受试者实验,准确率、 F_1 值、Kappa系数和计算时间(即网络训练时间)结果如表4所示.整体取得79.17%的平均准确率、0.788的平均 F_1 值和0.703的Kappa系数,体现了DCAMNet具有较好的分类准确性和泛化能力,且0.445s的平均计算时间表明模型具有较高的学习效率.

表4 DCAMNet 解码性能

所提模型	评估指标	受试者									平均值
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	
DCAMNet	准确率/%	87.50	59.03	92.71	76.39	68.06	61.81	90.97	86.11	89.93	79.17
	F_1 值	0.875	0.585	0.926	0.762	0.653	0.619	0.911	0.859	0.900	0.788
	Kappa系数	0.802	0.517	0.826	0.698	0.608	0.545	0.844	0.799	0.750	0.710
	计算时间/s	0.478	0.470	0.479	0.438	0.439	0.411	0.429	0.429	0.434	0.445

3.3.6 与其他流行方法对比实验

本文在DCAMNet主要组成部分的超参数优选对比实验中,验证了动态卷积块、空间卷积块和时间注意力块的参数取值对于MI-EEG解码性能的重要作用.实验结果表明,可训练参数过多、网络层数过深均会导致平均准确率下降.本节将通过与其他流行方法的对比,进一步展示所提出网络的有效性.

1) 与其他流行方法对比.

为了进一步验证所提出网络的性能优势,选择5个代表性的深度学习模型与所提出方法进行整体

表5 不同MI解码任务的深度学习模型与本文方法解码性能对比

发表时间	网络	平均准确率 \pm 方差/%	F_1 值 \pm 方差	Kappa \pm 方差
2012年	Deep ConvNet ^[9]	72.21 \pm 20.54	0.716	0.627
2021年	FBCNet ^[8]	72.60 \pm 12.39	0.762	0.683
2018年	EEGNet-8, 2 ^[11]	73.15	0.707	0.690
2021年	TS-SEFFNet ^[22]	74.71 \pm 11.32	0.757 \pm 0.131	0.663 \pm 0.151
2024年	IFBCLNet ^[17]	78.79 \pm 12.91	-	0.720
2024年	DCAMNet	79.17 \pm 12.45	0.788 \pm 0.128	0.710 \pm 0.118

比较,跨受试者的实验结果如表5所示.由表5结果可见,DCAMNet在平均准确率和泛化性能方面均具有一定优势,有效提升了MI-EEG的解码能力.

2) 网络预测相关性分析.

为了展示DCAMNet在不同频段下对MI任务进行预测时各EEG通道的相关性分布,对 S_3 测试集数据进行微分熵分析(DE),颜色变化反映不同区域的能量密度变化.图6为DCAMNet在不同频段下对MI任务进行预测时的EEG通道相关性分布.其中:颜色深浅表示相关性强弱,红色区域为高相关性,蓝色区域为低相关性.

图6中,以“左手”任务为例,在4~8Hz频段明显显示出中央脑区的高相关区域,表示该频段下的一些电极与“左手”任务高度相关.DCAMNet通过这些正相关区域提高了分类准确率.当这些电极在4~8Hz频段的振幅增大时,模型对“左手”任务的分类概率也随之增加,表明模型尤其关注了这些电极位置和该频段的信号特征.电极捕捉到了与“左手”运动想象相关的重要脑电活动,表明DCAMNet

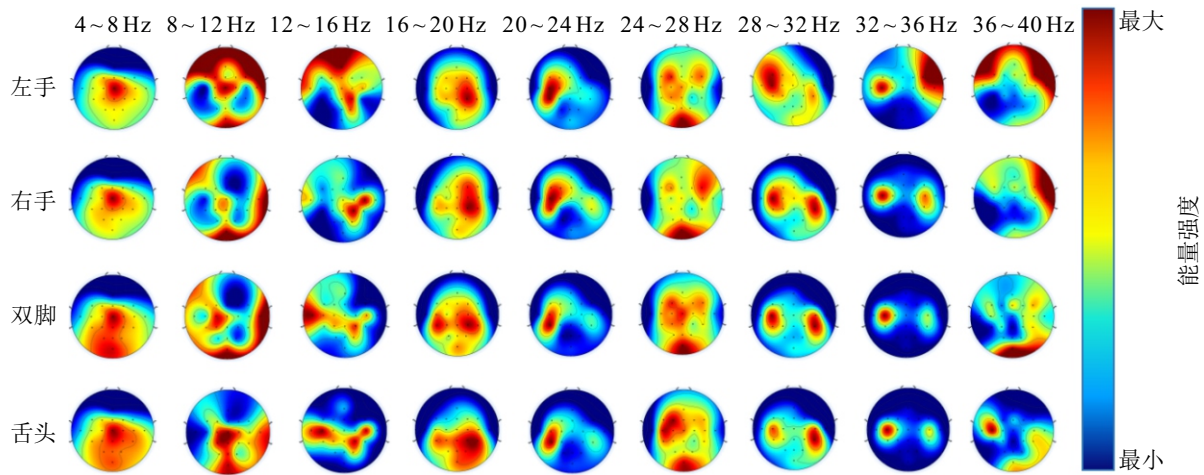


图6 S_3 在不同频段下的微分熵分布脑区能量

能够识别出特定脑区与 MI 任务的相关性。

3) 网络有效性分析。

DCAMNet 在模型设计上充分考虑了不同个体 MI-EEG 信号间在频率、空间和时间 3 个维度上的差异,通过引入动态卷积、空间卷积和时间注意力机制来增强模型的自适应能力和泛化能力。具体而言,针对不同个体 EEG 在频谱维度上的差异设计滤波器组,以确保每个频段的信息被有效提取,动态卷积块利用对输入信号的自适应学习动态调整卷积核,从而捕捉个体间 EEG 的时频特征差异;针对个体间 EEG 的空间分布模式差异设计空间卷积块,学习 EEG 通道维度上的相关空间信息,从而增强模型的自适应性和鲁棒性;针对个体在 MI 任务中的时序模式不同,利用时间注意力机制挖掘时间相关性,在不同时间窗口中聚焦最具判别力的时间特征,帮助模型自适应地关注不同个体的时间序列差异,能够进一步减小个体差异带来的不利影响。

第 3.3.3 节的网络消融实验结果表明,DCAMNet 通过多频段特征提取、动态卷积、空间卷积和时间注意力机制的协同作用有效地减小了个体差异对于模型分类性能的影响,展现出了 DCAMNet 在个体间差异性较大的 MI 任务中的优越性,具备较强的泛化能力。

4 结论

本文提出了一种新的 MI-EEG 解码方法 (DCAMNet),应用动态卷积块、空间卷积块和时间注意力块充分提取和融合了 MI-EEG 多视角数据表示的时频域和空域信息,特别是通过学习不同时间段的时间依赖关系来获取判别性特征,增强了运动想象解码性能。消融实验结果表明了动态卷积块和时间注意力块对整体网络性能有显著贡献,参数优选的对比实验结果则显示了网络参数设置的合理性和

有效性。所提出网络在 BCI Competition IV Dataset 2a 数据集上获得了 79.17% 的平均准确率和 0.788 的平均 F_1 值,且在保证泛化性能的前提下,使得网络训练参数相对较少 (140 k)。同时,通过对比实验也直观地验证了所提出网络具有较好的特征提取能力。未来将重点研究如何刻画不同受试者间的神经生理差异性,以期进一步改善模型的泛化能力。

参考文献 (References)

- [1] 王行愚, 金晶, 张宇, 等. 脑控: 基于脑-机接口的人机融合控制[J]. *自动化学报*, 2013, 39(3): 208-221. (Wang X Y, Jin J, Zhang Y, et al. Brain control: Human-computer integration control based on brain-computer interface[J]. *Acta Automatica Sinica*, 2013, 39(3): 208-221.)
- [2] 郝矿荣, 原博炜, 陈磊, 等. 基于免疫堆叠降噪自编码机网络的运动想象脑电信号识别[J]. *控制与决策*, 2018, 33(4): 719-723. (Hao K R, Yuan B W, Chen L, et al. Recognition of motor imagery EEG based on immune stacked denoising auto-encoders network[J]. *Control and Decision*, 2018, 33(4): 719-723.)
- [3] Abiri R, Borhani S, Sellers E W, et al. A comprehensive review of EEG-based brain-computer interface paradigms[J]. *Journal of Neural Engineering*, 2019, 16(1): 011001.
- [4] 付荣荣, 李林玉, 孟云, 等. 基于因果稀疏优化的运动想象脑电意图解码研究[J]. *控制与决策*, 2025, 40(2): 699-704. (Fu R R, Li L Y, Meng Y, et al. Research on casual sparse optimization for improving motor imagery electroencephalogram decoding[J]. *Control and Decision*, 2025, 40(2): 699-704.)
- [5] Ju Y, Sun G Y, Chen Q H, et al. A model combining convolutional neural network and LightGBM algorithm for ultra-short-term wind power forecasting[J]. *IEEE Access*, 2019, 7: 28309-28318.
- [6] Koles Z J. The quantitative extraction and topographic

- mapping of the abnormal components in the clinical EEG[J]. *Electroencephalography and Clinical Neurophysiology*, 1991, 79(6): 440-447.
- [7] 胡莹, 刘燕, 程晨晨, 等. 基于自适应时频共空间模式结合卷积神经网络的多任务运动想象脑电分类[J]. *生物医学工程学杂志*, 2022, 39(6): 1065-1081. (Hu Y, Liu Y, Cheng C C, et al. Multi-task motor imagery electroencephalogram classification based on adaptive time-frequency common spatial pattern combined with convolutional neural network[J]. *Journal of Biomedical Engineering*, 2022, 39(6): 1065-1081.)
- [8] Ang K K, Chin Z Y, Zhang H H, et al. Filter bank common spatial pattern in brain-computer interface[C]. *IEEE International Joint Conference on Neural Networks*. Hong Kong, 2008: 2390-2397.
- [9] Schirrmester R T, Springenberg J T, Fiederer L D J, et al. Deep learning with convolutional neural networks for EEG decoding and visualization[J]. *Human Brain Mapping*, 2017, 38(11): 5391-5420.
- [10] Ma X Z, Chen W H, Pei Z C, et al. A temporal dependency learning CNN with attention mechanism for MI-EEG decoding[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2023, 31: 3188-3200.
- [11] Lawhern V J, Solon A J, Waytowich N R, et al. EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces[J]. *Journal of Neural Engineering*, 2018, 15(5): 056013.
- [12] Santamaria-Vazquez E, Martinez-Cagigal V, Vaquerizo-Villar F, et al. EEG-inception: A novel deep convolutional neural network for assistive ERP-based brain-computer interfaces[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2020, 28(12): 2773-2782.
- [13] Freer D, Yang G Z. Data augmentation for self-paced motor imagery classification with C-LSTM[J]. *Journal of Neural Engineering*, 2020, 17(1): 016041.
- [14] Zhang H, Zu K K, Lu J, et al. EPSANet: An efficient pyramid squeeze attention block on convolutional neural network[C]. *Proceedings of the 16th Asian Conference on Computer Vision on Computer Vision*. Macao, 2023: 541-557.
- [15] Michielli N, Acharya U R, Molinari F. Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals[J]. *Computers in Biology and Medicine*, 2019, 106: 71-81.
- [16] Li Y, Zhang X R, Zhang B, et al. A channel-projection mixed-scale convolutional neural network for motor imagery EEG decoding[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2019, 27(6): 1170-1180.
- [17] Cao J S, Li G H, Shen J H, et al. IFBCLNet: Spatio-temporal frequency feature extraction-based MI-EEG classification convolutional network[J]. *Biomedical Signal Processing and Control*, 2024, 92: 106092.
- [18] 李红利, 郭伟, 张荣华, 等. 基于改进深度森林的运动想象脑电分类方法[J]. *控制与决策*, 2023, 38(2): 468-474. (Li H L, Guo W, Zhang R H, et al. Classification method of motor imagery EEG based on improved deep forest[J]. *Control and Decision*, 2023, 38(2): 468-474.)
- [19] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, 2018: 7132-7141.
- [20] Park J, Woo S, Lee J Y, et al. Bam: Bottleneck attention module[C]. *Proceedings of British Machine Vision Conference*. Newcastle, 2018: 1-14.
- [21] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module[C]. *Proceedings of the European Conference on Computer Vision*. Munich, 2018: 3-19.
- [22] Li Y, Guo L H, Liu Y, et al. A temporal-spectral-based squeeze-and-excitation feature fusion network for motor imagery EEG decoding[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2021, 29: 1534-1545.
- [23] Yang B, Bender G, Le Q V, et al. CondConv: Conditionally parameterized convolutions for efficient inference[C]. *Proceedings of Advances in Neural Information Processing Systems 32*. Vancouver, 2019: 1307-1318.
- [24] Barmpas K, Panagakis Y, Bakas S, et al. Improving generalization of CNN-based motor-imagery EEG decoders via dynamic convolutions[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2023, 31: 1997-2005.

作者简介

张美晨 (2000-), 女, 硕士生, 主要研究方向为脑机接口, E-mail: randi@emails.bjut.edu.cn;

李明爱 (1966-), 女, 教授, 博士, 博士生导师, 主要研究方向为脑机接口、神经康复, E-mail: limingai@bjut.edu.cn.