

DOI:10.16652/j.issn.1004-373x.2025.18.024

引用格式:谷学静,周记帆,郭志斌.融合注意力的双分支时空卷积脑电识别网络[J].现代电子技术,2025,48(18):159-164.

融合注意力的双分支时空卷积脑电识别网络

谷学静^{1,2}, 周记帆^{1,2}, 郭志斌^{1,2}

(1. 华北理工大学 电气工程学院, 河北 唐山 063000; 2. 唐山市数字媒体工程研究中心, 河北 唐山 063000)

摘要: 针对运动想象脑电信号特征表达能力弱、分类识别准确率低等问题,提出一种融合注意力机制的双分支时空卷积运动想象识别网络。为解决数据量少的问题,提出一种改进的滑动窗口技术进行数据增强;然后将数据输入到特征提取网络中,采用双分支结构同时提取时间和空间特征。时间分支上,采用多尺度时序卷积提取不同尺度下的时间特征;空间分支上,使用深度可分离卷积提取深层次的空间特征。最后,采用注意力机制为融合特征动态分配权重。与其他方法的对比结果表明,所提出的网络模型具有较好的分类性能和泛化能力。

关键词: 脑电信号; 运动想象; 多尺度时序卷积; 双分支结构; 数据增强; 注意力机制

中图分类号: TN911.7-34

文献标识码: A

文章编号: 1004-373X(2025)18-0159-06

Dual-branch spatio-temporal convolution EEG recognition network integrating attention

GU Xuejing^{1,2}, ZHOU Jifan^{1,2}, GUO Zhibin^{1,2}

(1. College of Electrical Engineering, North China University of Science and Technology, Tangshan 063000, China;

2. Tangshan Digital Media Engineering Technology Research Center, Tangshan 063000, China)

Abstract: In allusion to the problems of weak feature expression of motor image (MI) electroencephalography (EEG) signals and low accuracy of classification and recognition, a dual-branch spatio-temporal convolutional motor imagery recognition network integrating attention mechanism is proposed. In order to address the problem of few data, an improved sliding window technology is proposed for the data augmentation. The data is fed into the feature extraction network, and the dual-branch structure is used to simultaneously extract the feature in both temporal and spatial aspects. On the temporal branch, multi - scale temporal convolution is used to extract temporal features at different scales. On the spatial branch, deep spatial features are used to extract depth - separable convolution. The attention mechanism is used to dynamically assign weights to the fused features. The comparison results with other methods show that the proposed network model has better classification performance and generalization ability.

Keywords: EEG signal; motor imagery; multi - scale temporal convolution; dual - branch structure; data augmentation; attention mechanism

0 引言

脑机接口(Brain-Computer Interface, BCI)系统可以将脑电信号解码成计算机设备可识别的指令,进而实现大脑活动与外部设备的交互^[1-2]。BCI的应用领域广泛,包括临床诊断、康复治疗和脑功能研究等。BCI的应用首先需要获取脑电(Electroencephalography, EEG)信号,常见的EEG信号实验范式包括运动想象(Motor Imagery, MI)、稳态视觉诱发电位(Steady State Visual Evoked Potential, SSVEP)和P300^[3]。MI是通过想象而

不进行实际动作向大脑发送指令,通过解析MI信号可以实现对大脑认知活动的精确识别。BCI中信号处理部分是重要组成之一,主要包括预处理、特征提取、特征分类三个方面。传统MI-EEG分类算法主要是提取信号的时频域特征、空域特征以及非线性特征,然后使用单域特征或多域融合特征输入到分类器当中进行分类^[4]。时频域特征提取算法主要包括小波包变换^[5]、功率谱密度^[6]和傅里叶变换^[7]等。共空间模式^[8]作为一种常用的空域特征方法,通过矩阵的对角化,找到一组最优空间滤波器并进行投影,使得两类信号的方差值差异最大化,从而得到具有较高区分度的特征向量。虽然传统机器学习方法在二分类领域取得了不错的成效,但这

收稿日期:2024-08-14

修回日期:2024-10-18

基金项目:唐山市沉浸式虚拟环境基础仿真团队项目(18130221A)

些方法往往依赖于特征工程,会造成特征的遗漏或出现冗余特征,特别在面临高维数据时容易受到维度的影响,难以进一步挖掘EEG信号中的深层次特征。

随着深度学习技术的快速发展,尤其是卷积神经网络(Convolution Neural Network, CNN)在图像和语音识别领域的成功应用,研究者开始将CNN应用到脑电信号识别领域^[9]。CNN采用端到端的模式,可以从MI数据中自动提取特征,无需人工干预,有效提高了模型的泛化能力和准确率,不仅减少了机器学习中复杂的预处理步骤,而且处理高维数据时具有优势,一定程度上提高了模型的效率。文献[10]提出了一个参数量较小的EEGNet模型,在多个BCI范式中表现出了良好的泛化能力。文献[11]提出了DTP-Net,首次在神经网络中应用时间金字塔,能够有效提取多尺度特征并降低噪声。文献[12]提出了一种基于图形卷积神经网络的深度学习框架,以增强不同类型MI任务期间原始信号的解码能力。文献[13]提出了EEG-inception模型,通过将处理图像的模型应用到EEG时间序列中,提取脑电信号的时间特征,并在网络中添加残差块以防止学习退化。但是,此方法受限于卷积核大小,CNN只能从局部感受野接收信息,忽略了长程时间依赖关系。文献[14]提出了基于时间注意力的ATCNet模型,通过添加注意力机制的时间卷积网络进行轻量化设计,提高了MI的分类性能。虽然以上方法在解码性能上有所提高,但未能将注意力机制与EEG信号的时空特征结合起来。

考虑到MI-EEG信号中的多模态信息,本文提出了一种融合注意力机制的双分支时空卷积脑电特征识别网络。该网络采用双分支结构,分别通过多尺度时序卷积层、深度可分离卷积层同时捕捉MI信号中的时空信息,并结合注意力机制(Attention Mechanism)来增强关键特征的学习,从而提高分类的准确性,最后在BCI竞赛IV2a数据集和HGD数据集上进行验证。

1 数据集与预处理

1.1 数据集介绍

本文采用的第一个数据集是来自2008年BCI竞赛的IV2a数据集。实验共有9名被试,每个被试在不同天做了2个session,每个session有6个runs,每个runs有48个trials,一共288次测试。在每一次测试中受试者面对屏幕,测试开始时会伴随着提示音,屏幕上会出现一个固定的十字。2 s后,屏幕上会出现一个指向左、右、上、下的箭头,分别对应左手运动、右手运动、双脚运动和舌头运动,提示受试者进行运动想象。数据采集MI的时间为4 s,每个受试者根据箭头提示完成想象任务直到十

字消失,之后进行短暂的休息。本文实验中数据全部使用4种标签数据,2a数据集实验范式图如图1所示。

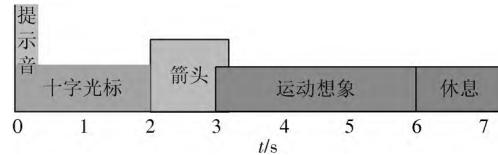


图1 2a数据集实验范式

本文的第二个数据集是HGD数据集,该数据集一共有14名被试,数据共有4种标签,分别为左手、右手、双脚和休息。每个session有13个runs,每个runs有80个trials,一共1 040次实验,具体实验细节参考文献[15]。

1.2 数据预处理

受试者在进行感觉运动相关活动时,会出现事件相关去同步现象和事件相关同步现象,在实验过程中通常利用这两个现象下特定频段的信号对MI信号进行识别。2a数据集一共有25个通道,其中22个为EEG通道,3个为EOG通道,信号采样频率为250 Hz。HGD数据集一共有128个通道,采样频率为500 Hz。EEG信号中常常包含来自头部运动和眼动的伪影,本文使用独立成分分析去除伪影部分。为了提高EEG信号的信噪比,使用4阶Butterworth滤波器对8~30 Hz的信号进行带通滤波处理和50 Hz的陷波滤波处理。两个数据集在每次实验中均使用4.5 s时间窗口。

2 方法实现

BCI系统需要建立一个实时性好、准确率高、泛化能力强的模型,从而满足在实际应用过程中的需求。虽然传统机器学习模型在二分类方面取得了很高的分类结果,但这些模型依赖人工提取特征,以及在处理大规模数据时存在效率不高的问题。针对上述问题,本文提出的模型可以从时域和空域充分提取特征信息,以实现对不同类别的准确识别。本文所提模型主要由三部分组成,网络总体框架如图2所示。首先是数据增强(Data Augmentation, DA)部分,为模型创建多样化数据,防止模型过拟合;然后是双分支时空网络模块,同时对时间和空间两方面进行特征提取,随后将特征进行融合并输入到注意力层,以对不同特征分配权重;最后经过平均池化层和Dropout层,将结果输入到全连接层进行分类。

2.1 数据增强

EEG数据采集难度大且耗时长,导致EEG训练数据集有限。在深度学习中,数据增强是一种在不实质性增加数据的情况下,让有限的数据产生更多的数据价

值,以此来提高模型的泛化能力和鲁棒性的技术。通过增加数据集的大小,可以减少模型对训练数据的过度拟合,防止网络学习一些不相关的特征,从而在测试集上获得更好的性能。本文采用滑动窗口(Sliding Window, SW)技术来实现数据增强。首先确定时间窗口大小,太

小会无法捕捉所有相关信息,太大可能会包含不相关信息,故本文将设计好的窗口在序列数据上进行滑动,步长设置为10,每次以固定步长移动,通过随机组合不同的滑动窗口的数据来合成新的样本,以此增加数据集的大小。滑动窗口技术原理图如图3所示。

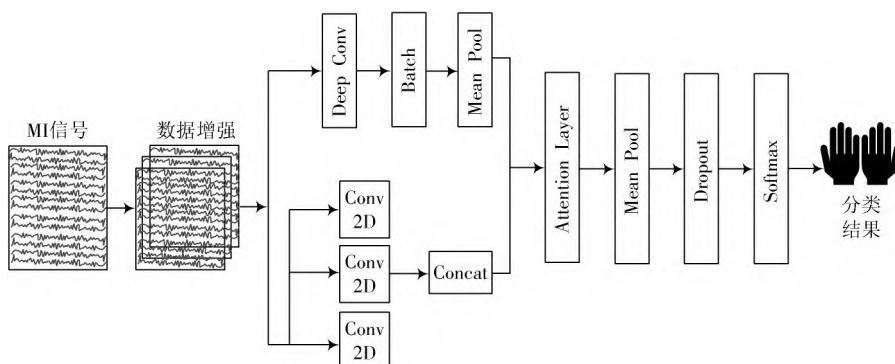


图2 本文模型网络框架

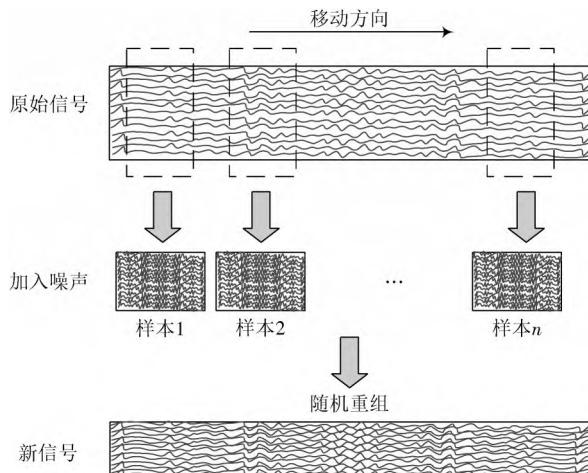


图3 滑动窗口技术原理

2.2 多尺度时序卷积

MI-EEG作为一种时间序列,具有非周期性、非线性、时变复杂等特点,在时间方面具有丰富的信息,但单一尺度的卷积核感受野太小,无法充分提取MI特征。本文设计了多尺度时序卷积模块(Multi-scale Temporal Convolution Module, MTCM),通过并行设计3个不同的二维卷积核,在不同尺度上对数据进行处理,从而增加模型对输入的感受野,以捕捉数据中的局部和全局特征。激活函数选择ReLU。

MTCM中的3个卷积核大小分别为 1×15 、 1×25 和 1×51 ,使模型能够更好地理解和捕捉时间序列数据中的长期时间依赖关系和时域上的信息。最后得到不同尺度下的特征图,将不同尺度的特征图进行拼接,形成一个多尺度特征张量。本文所提模型中MTCM架构如图4所示。

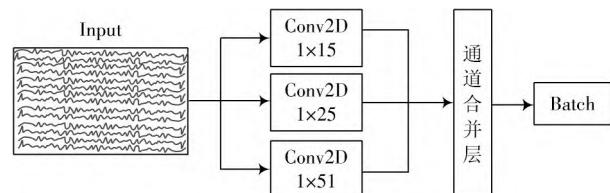


图4 多尺度时序卷积模块架构

2.3 深度可分离卷积

在空间分支方面,本文采用深度可分离卷积作为该卷积层的主要组件,进行空间方面特征提取,以更好地捕捉通道之间和空间位置之间的特征关系。在MI分类过程中,一次MI任务与多个通道有关,因此在特征提取时,需要进行逐通道卷积来捕捉重要信息。深度可分离卷积是一种将标准卷积分解为深度卷积和逐点卷积两个步骤的卷积操作,这样的分解使得网络能够在保持特征图尺寸不变的情况下有效地减少参数数量,从而提高模型的效率和泛化能力。深度卷积对于每个输入通道,使用一个独立的卷积核进行卷积操作,对输入的MI数据进行空间特征提取;逐点卷积将深度卷积的输出作为输入,将组合所有深度卷积的输出,从而捕捉不同通道之间的相互关系。本文在空间分支中使用了二维卷积操作,特征图大小为32,卷积层之后是归一化层,激活函数选择ReLU函数,采用平均池化层以减少特征维度和增加模型的感受野。平均池化是沿着核大小(1, 50)的时间维度进行的,经过双分支时空卷积处理后,将输出特征输入到自注意力机制模块,以提高模型对输入数据的抽象能力和表达能力。

2.4 注意力机制

注意力机制模拟了人类对外界刺激的注意力分配机制,旨在自动学习并强调特征中的关键信息,从而增强分类性能。通过对时空信息中的不同特征进行动态调整,以便网络更加关注那些对于运动想象识别任务而言更为关键的特征。自注意力机制中的注意力权重公式如式(1)所示,每个特征的权重是由查询(Query)、键(Key)、值(Value)组成,通过 Q 、 K 、 V 这3个矩阵计算每个注意力头的权重。

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

式中:Softmax是激活函数; $\sqrt{d_k}$ 是 Q 、 K 矩阵的列数,即向量维度。

3 实验及结果分析

3.1 实验相关设置

本文所有实验在基于Linux的服务器上运行,处理

器为Intel Xeon Platinum 8255C CPU 2.5 GHz,配备NVIDIA GTX 2080 Ti GPU。使用PyTorch 1.7和Python 3.7搭建深度学习模型,使用MNE库对数据进行预处理。在训练过程中采用10折交叉验证,数据集被随机打乱,80%划分为训练集,20%划分为测试集。模型Epoch参数设置为500,Batch size大小设置为10,学习率设置为0.001,优化函数为Adam,激活函数为ReLU函数。训练采用早停法,当连续20轮没有改进且损失变化小于阈值,便停止训练,保存最优模型。

3.2 实验结果分析

为了对模型进行全面评估,本文采用准确率(Acc)和Kappa系数这两个指标进行综合评价,将本文模型与FBCSP^[16]、EEGNet、DeepConvNet^[17]、FBCNet^[18]、ATCNet等5种分类方法进行对比。每个被试均实验10次,结果取平均。2a数据集分类结果如表1所示。

表1 2a数据集分类结果

模型	被试10次分类结果的准确率/%									9名被试 平均分类 准确率/%	Kappa系数
	1	2	3	4	5	6	7	8	9		
FBCSP	79.5	55.1	88.3	59.9	50.6	61.5	65.8	81.2	73.7	68.4	0.58
EEGNet	85.9	62.2	93.7	70.1	52.5	68.4	71.6	81.4	68.5	72.7	0.65
DeepConvNet	78.3	52.4	95.8	58.7	53.1	64.5	68.6	92.2	73.6	70.8	0.63
FBCNet	86.4	56.7	96.9	78.4	63.5	59.0	82.5	86.8	81.0	76.8	0.68
ATCNet	88.6	67.2	94.9	72.0	66.7	68.2	87.8	94.4	85.6	80.6	0.73
本文模型	88.1	67.9	96.2	75.5	84.4	82.6	90.3	95.0	88.6	85.4	0.80

表1中,第2~10列表示当前被试在不同模型下的分类准确率,最后一列为平均Kappa值,每个被试最高结果用加粗表示。从表中可以看出,本文方法在每名被试上均取得了最佳或者次佳准确率。FBCSP是一种基于CSP改进的机器学习分类方法,该方法通过频带划分寻找最优滤波器,从而提取空间特征;再采用支持向量机进行分类,最终平均分类准确率只有68.4%,各项指标均低于深度学习方法,准确率比本文方法低17%。EEGNet作为经典模型,平均分类准确率只有72.7%。DeepConvNet采用深度卷积网络对MI数据进行特征提取,平均分类准确率达到了70.8%,但可能是卷积层数过深,导致过度拟合。FBCNet通过多个窄带滤波器获得不同频带的MI数据,通过Depthwise卷积层获得空间信息,但其在2a数据集上平均准确率只有76.8%,比本

文方法低了8.6%。ATCNet是一个基于注意力的时间卷积网络,平均准确率为80.6%,高于其他未添加注意力机制的模型。本文方法所有被试的平均分类准确率为85.4%,与其他深度学习方法相比,所提网络拥有更好的结果,但被试3分类准确率结果最高为96.2%,被试2结果略差,为67.9%。原因可能是在脑电信号采集过程中存在噪声干扰,影响采集的信号质量,在训练过程中无法得到最佳特征。Kappa系数用来衡量模型一致性,数值越高代表越与实际情况相同,本文Kappa系数为0.80,高于其他算法,可知模型具有良好的一致性。

3.3 消融实验分析

为了解滑动窗口技术和注意力机制在整个模型中的作用,分别去除DA、线性注意力机制(LA)和同时去除LA和DA,将本文网络作为基本网络与上述方案进行

实验对比,每名被试的分类结果如图5所示。从图5中可以看出,DA与注意力机制在提高分类性能方面效果显著,有数据增强相比于不添加数据增强的平均分类准确率提高了3.1%,说明数据增强能够有效避免网络过拟合,提高网络性能与泛化能力。在添加注意力机制之后,总体平均准确率提高了8.6%,表明注意力机制可以更好地帮助模型聚焦重要的部分,忽略无关信息。

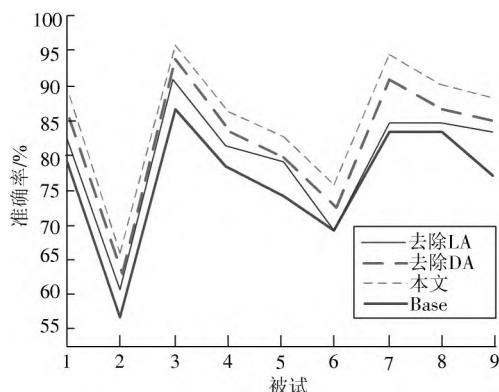


图5 消融实验结果

3.4 模型稳定性分析

为进一步验证本文方法的稳定性与泛化能力,将本文网络应用在HGD数据集上,所得准确率结果如图6所示。通过图6中结果可知,被试3分类准确率最高为99.3%,被试11最低分类准确率为80.6%,这可能是信号采集质量不佳,无法充分提取特征所致。综合计算可知,平均分类准确率为91.3%,由此可见,本文算法具有良好的分类能力和泛化能力。

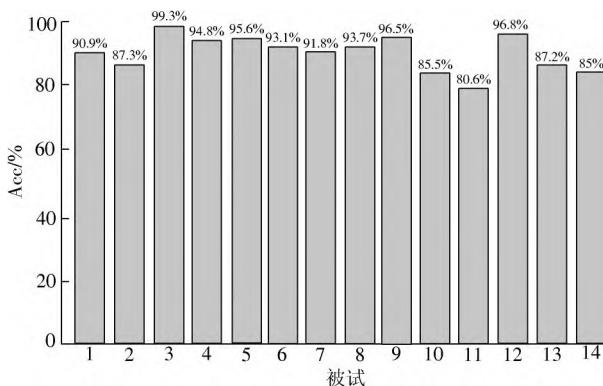


图6 HGD数据集分类结果

为了更加直观地了解本文网络在HGD数据集中上4个类别的分类结果,绘制了如图7所示的混淆矩阵,其中纵轴表示真实标签,横轴代表模型预测标签,图中数值为每个任务的分类准确率,对角线结果代表该任务的召回率。从图中可以看出,右手分类结果最好,为97%。综上所述,本文网络模型对不同数据集和分类任务具有良好的自适应性和鲁棒性。

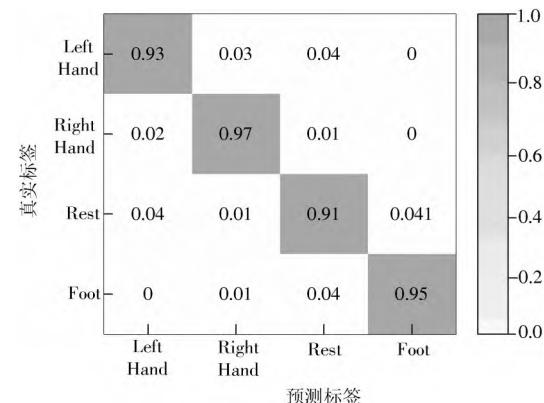


图7 HGD分类混淆矩阵

4 结语

本文提出了一种融合注意力机制的双分支时空卷积网络,用于MI-EEG识别。针对脑电信号数据缺乏的问题,采用了滑动窗口的方法进行数据增强。本文采用双分支结构同时提取不同尺度下的时间特征和深度空间特征,然后通过自注意力层进一步识别关键特征,得到具有时空信息增强的特征图。最后在BCI竞赛IV2a和HGD数据集上进行验证,平均分类准确率分别为85.4%和91.3%。实验结果表明,相比于其他深度学习方法,本文模型具有更好的分类性能、稳定性和泛化能力。在未来研究中,希望将现有成果为脑机接口技术的发展和应用带来更多的突破和进步,为医疗和康复领域做出更大的贡献。

注:本文通讯作者为周记帆。

参 考 文 献

- [1] VÄRBU K, MUHAMMAD N, MUHAMMAD Y. Past, present, and future of EEG-based BCI applications [J]. Sensors, 2022, 22(9): 3331.
- [2] 任玲玲,王力,黄学文,等.基于注意力与多尺度的4类脑电信号解码[J].科学技术与工程,2022,22(34):15180-15187.
- [3] 陈黎,龚安民,丁鹏,等.基于欧式空间-加权逻辑回归迁移学习的运动想象EEG信号解码[J].南京大学学报(自然科学),2022,58(2):264-274.
- [4] FAN C C, YANG H, HOU Z G, et al. Bilinear neural network with 3 - D attention for brain decoding of motor imagery movements from the human EEG [J]. Cognitive neurodynamics, 2021, 15: 181-189.
- [5] 陈万忠,王晓旭,张涛.基于可调Q因子小波变换的识别左右手运动想象脑电模式研究[J].电子与信息学报,2019,41(3): 530-536.
- [6] 赵德春,沈利豪,蒋宇皓,等.基于共空间模式和功率谱密度的脑电信号分类[J].科学技术与工程,2023,23(10):4272-4278.
- [7] 衡玮,徐自强,陈璐瑶,等.基于脑电波监测智能助睡眠耳机设

- 计[J].计算机应用与软件,2020,37(8):219-226.
- [8] LI Y, WANG Q, WANG T, et al. Feature extraction of EEG signals based on local mean decomposition and fuzzy entropy [J]. International journal of pattern recognition and artificial intelligence, 2020, 34(12): 2058017.
- [9] KHADEMI Z, EBRAHIMI F, KORDY H M. A review of critical challenges in MI - BCI: from conventional to deep learning methods [J]. Journal of neuroscience methods, 2023, 383: 109736.
- [10] 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.
- [11] PEI Y, XU J, CHEN Q, et al. DTP - Net: learning to reconstruct EEG signals in time-frequency domain by multi-scale feature reuse [J]. IEEE journal of biomedical and health informatics, 2024, 28(5): 12.
- [12] HOU Y M, JIA S Y, LUN X M, et al. GCNs - net: a graph convolutional neural network approach for decoding time - resolved EEG motor imagery signals [J]. IEEE transactions on neural networks and learning systems, 2024, 35(6): 7312-7323.
- [13] ZHANG C, KIM Y K, ESKANDARIAN A. EEG-inception: an accurate and robust end-to-end neural network for EEG-based motor imagery classification [J]. Journal of neural engineering, 2021, 18(4): 046014.
- [14] INGOLFSSON T M, HERSCHE M, WANG X, et al. EEG - TCNet: an accurate temporal convolutional network for embedded motor-imagery brain-machine interfaces [C]// 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC). [S.1.]: IEEE, 2020: 2958-2965.
- [15] ALTAHERI H, MUHAMMAD G, ALSULAIMAN M. Physics-informed attention temporal convolutional network for EEG-based motor imagery classification [J]. IEEE transactions on industrial informatics, 2022, 19(2): 2249-2258.
- [16] ANG K K, CHIN Z Y, WANG C, et al. Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b [J]. Frontiers in neuroscience, 2012, 6: 21002.
- [17] SCHIRRMEISTER 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.
- [18] MANE R, CHEW E, CHUA K, et al. FBCNet: a multi-view convolutional neural network for brain-computer interface [EB/OL]. [2023-01-12]. <https://arxiv.org/abs/2104.01233>.

作者简介:谷学静(1972—),女,河北唐山人,博士研究生,教授,研究方向为虚拟现实、脑机接口。

周记帆(2000—),男,河南新乡人,硕士研究生,研究方向为脑机接口及模式识别。

郭志斌(1999—),男,山东泰安人,硕士研究生,研究方向为运动想象识别。