

DOI:10.19651/j.cnki.emt.2417039

用于运动想象脑电信号分类的深度学习网络^{*}

胡博 谢俊 刘俊杰 王虎

(新疆大学智能制造现代产业学院 乌鲁木齐 830017)

摘要: 运动想象(MI)脑电信号由于包含较长、连续的特征值以及其本身较强的个体差异性和较低的信噪比,导致其识别较为困难。本研究提出一种结合卷积神经网络(CNN)与 Transformer 的模型,旨在有效解码和分类运动想象脑电信号。该方法以原始多通道运动想象脑电信号作为输入,首先在第1个时间卷积层对信号的时域进行卷积操作,随后在第2个空间卷积层对信号的空域进行卷积操作,从而学习整个一维时间和空间卷积层的局部特征。其次,通过沿时间维度的平均池化层平滑时间特征,并将每个时间点的所有特征通道传递到注意力机制中,以提取局部时间特征中的全局相关性。最后,采用基于全连接层的简单分类器模块对脑电信号进行分类预测。通过在公开的 BCI 竞赛数据集 IV-2a 和数据集 IV-2b 上的实验验证,结果显示该模型可以有效分类 MI 脑电信号,平均分类准确率可达 80.95% 和 84.79%,相比于 EEGNet 网络,平均分类准确率分别提升了 6.45% 和 4.31%,有效的提高了运动想象诱发电位信号的脑-机接口性能。

关键词: 运动想象;注意力机制;脑电信号;脑-机接口

中图分类号: TN91 **文献标识码:** A **国家标准学科分类代码:** 310.61;510.99

Deep learning networks for motor imagery EEG signal classification

Hu Bo Xie Jun Liu Junjie Wang Hu

(College of Intelligent Manufacturing Modern Industry, Xinjiang University, Urumqi 830017, China)

Abstract: Motor imagery (MI) EEG signals are more difficult to recognize due to the inclusion of long, continuous eigenvalues as well as their own strong individual variability and low signal-to-noise ratio. In this study, we propose a model that combines a convolutional neural network (CNN) with a Transformer, aiming to effectively decode and classify motor imagery EEG signals. The method takes the original multichannel motor imagery EEG signals as input, and learns the local features of the entire one-dimensional temporal and spatial convolutional layers by firstly performing a convolutional operation on the temporal domain of the signals in the first temporal convolutional layer, and then subsequently performing a convolutional operation on the null domain of the signals in the second spatial convolutional layer. Next, the temporal features are smoothed by averaging the pooling layers along the temporal dimension and passing all the feature channels at each time point to the attention mechanism to extract the global correlations in the local temporal features. Finally, a simple classifier module based on a fully connected layer is used to classify the EEG signals for prediction. Through experimental validation on the publicly available BCI competition dataset IV-2a and dataset IV-2b, the results show that the model can effectively classify MI EEG signals with average classification accuracies of up to 80.95% and 84.79%, which is an improvement of 6.45% and 4.31% in comparison to the EEGNet network, respectively, and effectively improves motor imagery evoked potential signals of the brain-computer interface performance.

Keywords: motor imagery; attentional mechanisms; electroencephalogram signals; brain-computer interface

0 引言

脑-机接口(brain-computer interface, BCI)能够在人脑

与外界之间建立通信或控制通路,通过解码各种脑活动产生的脑电图(EEG)信号增强人脑与外部通信和交互的能力^[1]。脑-机接口已广泛应用于医学、神经生物学、心理学

收稿日期:2024-10-05

^{*} 基金项目:国家自然科学基金(52365040)项目资助

等领域,在治疗脑卒中、辅助治疗精神疾病等方面已成为一种有前景的工具。

运动想象脑-机接口实验范式要求受试者想象身体不同部位的运动,而不是实际的运动^[2]。当人在想象某种运动行为时,相关的脑运动皮层会产生神经元反应,通过减少大脑疾病导致的运动意图与运动感觉反馈之间的差距,将有助于促进运动功能的恢复。然而,运动想象脑电图(motor imagery electroencephalogram, MI-EEG)信号非平稳、低信噪比和非线性的特点带来了脑电解码上的困难^[3]。因此,对脑电信号进行分类就显得尤为重要^[4]。一般来说,运动想象脑电信号分类识别的方法可以分为两大类:传统方法和深度学习方法^[5]。传统的机器学习运动想象脑电图(MI-EEG)分类方法通常包括特征提取和特征分类,例如,用于特征提取的滤波器组公共空间模式(FBCSP)^[6]、快速傅立叶变换(FFT)、小波变换等,用于特征分类的支持向量机(SVM)^[7]、线性判别分析(LDA)的监督学习方法^[8]以及K最近邻(KNN)^[9]的无监督学习方法。但是,传统的方法在解码运动想象脑电信号时存在局限性,大多数传统的特征提取方法都是针对特定任务设计的,这意味着针对不同的脑机接口范式,特征提取依赖于特定的先验知识,因此其泛化能力有限^[10],这使得传统方法效率低下。

深度学习(deep learning, DL)模型^[11]具有自动特征提取和丰富的特征表示能力,可以直接接收预处理后的脑电数据,无需特征提取即可建立端到端模型。例如卷积神经网络(convolutional neural networks, CNN)、长短期记忆网络(long short-term memory, LSTM)等^[12]。CNN 在计算机视觉任务中表现出了出色的表征能力。正如预期的那样,改进的 CNN 模型 ConvNet^[13]在脑电分类任务上取得了与传统算法相当的性能,学习了卷积层的判别特征。Sakhavi 等^[14]通过引入一种新的数据时态表示,并利用 CNN 架构进行分类,提出了一种新的 MI 分类框架。Sun 等^[15]提出了一种 SE-CNN 网络,该网络使用稀疏频谱时间分解算法提取时频特征,相较于传统的分类算法,SE-CNN 在准确性和 Kappa 值方面均显示出了优越性。Lawhern 等^[16]提出了一种引入深度卷积和可分离卷积的神经网络模型 EEGNet,该模型旨在学习如何最优地结合特征映射,为 MI-EEG 的解码提供了一个很好的思路。然而上述方法仍存在问题,卷积神经网络(CNN)在学习过程中受到卷积核大小的限制,这导致它主要学习局部接受域的特征而忽略了全局信息和长距离依赖关系。较大的卷积核可能会限制深层 CNN 模型的有效利用,而较小的卷积核则会限制 CNN 模型感受野的覆盖范围^[17]。Tortora 等^[18]利用长短期记忆(LSTM)深度神经网络来捕获 EEG 分类中的时间特征。然而,这样的模型无法并行训练,并且忽略了中间层的信息,计算得出的依赖关系影响在几个时间步之后很快丢失。

近几年,Vaswani 等^[19]首次提出了 Transformer 模型,

在自然语言处理(natural language processing, NLP)和图像处理领域引起了不小的轰动。Sun 等^[20]将 Transformer 和 3D-CNN 模型的特征并行起来提高了 MI-EEG 的分类准确性。Liu 等^[21]提出时间和通道注意力卷积网络,提升了对脑电信号的分类准确率。然而,这种模型在捕捉局部时间和空间特征方面存在不足,忽略了对局部特征的学习,以及在小样本数据学习中的过拟合倾向,而这也是 EEG 解码所必需的。注意力机制在 Transformer 模型中扮演着重要角色,不仅在计算长序列特征相关性和实现模型可视化与可解释性方面表现出色,而且在处理长期依赖关系时展现出更为卓越的性能^[22]。但是传统的多头注意力机制中各个头部通常独立运作,这导致了头部之间缺乏交互,限制了模型整体性能的提升潜力。然后不同的头部可能会学到相似或冗余的信息,这种冗余不仅浪费计算资源,还可能导致模型效率低下。

为了解决上述问题,本文提出了一种 CNN 与 Transformer 相结合的 MI-EEG 分类新方法,命名为 Talkformer。首先,使用时间和空间卷积分别捕获局部时间和空间特征。然后再利用平均池化层对时间特征片段进行切分,将其输入到注意力机制中,提取局部时间特征中的全局相关性,这样既降低了模型复杂度,又消除了冗余信息,增强了模型的表达能力和学习效率。最后,使用简单的全连接层对脑电数据进行分类。实验结果表明,这种方法显著提高了运动想象(MI)脑电信号分类的准确性。

1 数据集与预处理

1.1 数据集

本文使用了由奥地利格拉茨工业大学提供的 BCI 竞赛 IV-2a 数据集^[23]和 BCI 竞赛 IV-2b 数据集作为实验数据基础,BCI-2a 包含了 9 名受试者在进行左手、右手、双脚和舌头 4 种运动想象任务时的脑电图(EEG)信号。每个受试者的实验数据分为两组,即训练集和测试集。两组实验数据中分别包含了 6 组小实验,每次小实验又包含了 48 个测试(即 4 类动作,每类动作各有 12 次,但顺序随机出现,共 $4 \times 12 = 48$ 次),则每组实验数据各包含有 288(6 组 $\times$ 48 个 = 288 次)次测试。实验采集装置共包括 25 个信号通道,其中有 22 个通道用于记录脑电图(EEG)信号,而另外 3 个通道则捕捉眼动图(EOG)信号,数据信号分类仅涉及 EEG 信号,EOG 信号所对应的 3 个通道数据被排除在外。数据使用 250 Hz 的采样率进行记录,实验按照图 1(a)所示的计时流程来进行记录。

在 BCI-2b 数据集中,对 9 名健康受试者进行了运动想象(MI)脑电信号的采集,涉及两种任务类型:左手和右手的运动想象。数据采集使用了 3 个 EEG 电极,并设置采样率为 250 Hz。整个采集过程分为 5 个阶段,其中前两个阶段每阶段收集了 120 个样本,而后 3 个阶段则各收集了 160 个或 120 个样本。实验选择[3, 7]s 时间段内的 EEG 时间

点作为样本,实验按照图 1(b)所示的计时流程来进行记录。

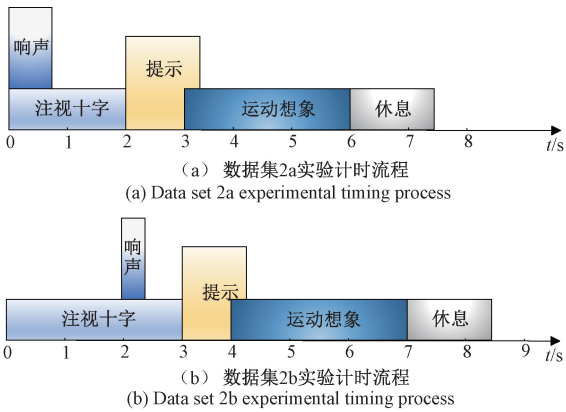


图 1 数据集 2a 和数据集 2b 运动想象范式示意图

Fig. 1 Schematic diagram of the motion imagery paradigm for dataset 2a and dataset 2b

1.2 数据预处理

使用 Chebyshev II 类型的滤波器,对数据进行带通滤波以降低高频和低频噪声,只保留 4~40 Hz 的频率成分,这是因为这个频率范围内的信号与运动想象任务高度相关。

使用 Matlab 代码从数据集中批量提取和处理脑电图

(EEG)信号数据,遍历每个试验或时间点,从每个试验中提取出一个特定时间窗口的 EEG 信号,并将其与对应的标签关联起来,用于进一步的数据分析。

本文采用了时域分割和重构(S&R)来生成新的数据^[24],将同一类别的训练样本平均分成 n 个片段,然后在保持原有时间顺序的情况下随机拼接,在每次迭代中生成与批处理相同大小的增强数据。这种增强策略特别适合于时间序列数据,如脑电图(EEG)信号,可以帮助模型学习到更广泛的数据分布特性,从而提高其泛化能力。

采用 z 分数进行标准化,以减少波动和非平稳性。

$$x_o = \frac{x_i - \mu}{\sqrt{\sigma^2}} \tag{1}$$

式中: X_i 和 X_0 分别表示带通滤波后的数据和标准化后的输出。 μ 和 σ^2 表示均值和方差,用训练数据计算,直接用于测试数据。

2 模型结构与评价指标

如图 2 所示,模型的总框架包括 3 个部分:卷积模块、注意力模块、全连接层分类器。首先,对脑电数据进行时间卷积和空间卷积来捕获局部时间和空间特征,用平均池化层在时间维度上对信号进行下采样,然后,通过注意力机制层进一步学习全局时间依赖性,最后使用全连接层进行分类。

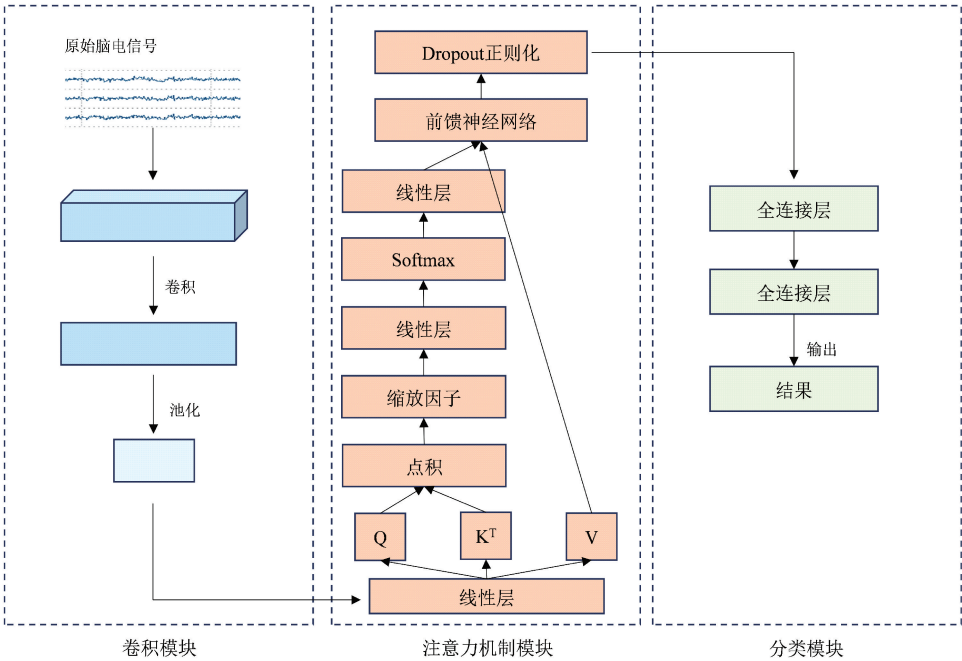


图 2 模型结构

Fig. 2 Model structure

2.1 卷积模块

脑电数据以四维张量的形式输入,第 1 层为卷积层,有 40 个大小为 (1,25) 的卷积核,其步长为 (1,1),这意味着

卷积是在时间维度上执行的,目的是提取时间序列上的特征。第 2 层也是卷积层,有 40 个大小为 (22,1) 的核,步长为 (1,1),其中 22 是 EEG 数据的电极通道数。这一层作为

一个空间滤波器来学习不同电极通道之间相互作用的表示。随后采用批量归一化操作加速训练并提高模型稳定性,使用指数线性单元(elu)作为激活函数,为网络添加非线性能力,帮助模型捕捉复杂的特征。第三层是沿时间维度的平均池化层,内核大小为(1,75),步长为(1,15)。该池化层平滑了时间特征,既避免了过拟合,又降低了计算复杂度。最后,重新排列卷积模块的特征映射,压缩电极通道维度,并将卷积通道维度与时间维度进行转置。

2.2 注意力机制模块

由于神经活动是连贯的,所以 EEG 信号时空特征的上下联系对于增强运动想象脑电解码的效果尤为关键。该模块使用 Talking-Heads Attention 来学习 EEG 特征的全局时间依赖性,以补充卷积模块中有限的感受域。具体如下,输入序列会被分别线性投影成多组查询(query, Q)、键(key, K)和值(value, V),对 Q 和 K 使用点积来计算注意力权重,添加比例因子来避免梯度消失,将结果通过 softmax 函数进行归一化之后,得到权重矩阵,也就是注意力得分,然后将注意力得分对 V 用点积进行加权。THA 通过在计算得到注意力得分之后以及将不同头部的输出合并之前引入额外的线性变换,来增强不同头部间的交互,大大提高了模型的表达能力。这个过程可以表述为:

$$THAttention(Q, K, V) = Softmax\left(A \times \left(\frac{QK^T}{\sqrt{d_k}}\right) \times B\right)V \quad (2)$$

式中: d_k 是键 K 的维度,这个缩放因子用于控制内积的大小,以防止梯度消失。 A 和 B 是两个附加的矩阵,是作为模型参数通过学习得到的,它们作用在多头注意力的得分上,用于在不同头之间调节注意力分配,它们的形状通常与注意力机制中的头数 h 有关。

最后,将所有头的输出拼接起来,并通过最终的线性投影:

$$TalkHead(Q, K, V) = Concat(head_1, \dots, head_h)W^0 \quad (3)$$

$$Head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \quad (4)$$

式中: $Concat$ 是矩阵的拼接函数, $Head_i$ 是第 i 个 head 的输出结果, W^0 是输出权重矩阵。 W_i^Q, W_i^K, W_i^V 是可学习的权重矩阵,对应于每个头的查询、键和值的投影。

THA 通过增加模型内部不同注意力头部之间的交互,改进了模型对信息的整合能力,允许模型更全面地识别 EEG 信号中的复杂模式。这种头部间的动态信息交换大大提高了分类精度,使模型能够更精确地区分各种类型的运动想象任务。同时,THA 还提高了模型的适应性和灵活性,使其能够根据不同个体的差异来调整注意力分配,从而在多变的数据环境中做出准确预测。

2.3 分类模块

采用 1 个全局平均池化层和 3 个全连接层作为分类器模块,全局平均池化层利用平均池化方法,将特征向量序

列被转换为均值表示,每一个均值对应到一个类别标识或是一个输出单元,再经过两个全连接层到第 3 层时形成最终的四类输出。

2.4 训练过程与方法

在本文的注意力机制中,注意力头数设为 10。为了防止模型训练过程中的过拟合,设置丢弃率为 0.5,同时使用了 Adam 优化器,学习率设置为 0.000 2,该优化器因其简洁的实现方式、高效的计算性能以及较低的内存需求而广泛应用并且它同样适用于大数据集和高维空间。本文的方法在 Python 3.11 中的 PyTorch 库中使用 Geforce 4090 GPU 实现的,选用交叉熵作为整个框架的损失函数,该损失函数适用于多分类网络,如式所示,它衡量的是模型预测概率分布与真实标签的概率分布之间的差异,这就使得模型在训练过程中,通过优化算法(如梯度下降)调整参数,以最小化交叉熵损失。

$$L(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(\hat{y}_{i,c}) \quad (5)$$

式中: C 是类别总数, $y_{i,c}$ 是一个指示器,当样本 i 属于类 C 时为 1,否则为 0。模型的预测效果随着交叉熵损失值的降低而提升。

设置实验的训练总迭代次数为 2 000 次,本文使用 ELU (exponential linear unit, ELU) 激活函数,在 Transformer 模型中,ELU 激活函数可以通过其非零负区域梯度和输出均值接近零的特性,有助于改善层间信息传递、减少训练偏差,促进模型的稳定训练和深层表示学习,如式(6)所示。

$$ELU(x) = \begin{cases} x, & x > 0 \\ \alpha(e^x - 1), & x \leq 0 \end{cases} \quad (6)$$

2.5 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。

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。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

式中: TP 是模型正确预测为正类的样本数量,即实际为正类且被模型正确分类的样本数; TN 是模型正确预测为负类的样本数量,即实际为负类且被模型正确分类的样本数; FP 是模型错误预测为正类的样本数量,即实际为负类但被模型误分类为正类的样本数; FN 是模型错误预测为负类的样本数量,即实际为正类但被模型误分类为负类的样本数。

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。

$$kappa = \frac{p_o - p_e}{1 - p_e} \tag{8}$$

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 and 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。

ROC 曲线(receiver operating characteristic curve, ROC)是一种用于评估二分类模型性能的图形工具,通过展示模型在不同阈值下的真阳性率(TPR)与假阳性率(FPR)之间的权衡,来衡量分类器的识别能力。AUC 值(area under the curve, AUC)是 ROC 曲线下方的面积,数值在 0~1 之间,反映了模型的总体判别能力。AUC 值越接近 1,表示模型的区分能力越强;若 AUC 值为 0.5,模型的表现相当于随机猜测。ROC 曲线和 AUC 值特别适用于不平衡数据集或多分类任务中模型性能的比较与选择。

3 实验结果与分析

3.1 结果与其他方法对比

使用公开数据集 BCI Competition IV 2a 和 2b 进行实

验结果的对比,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。例如,获得 BCICompetition IV 冠军的 FBCSP,它通过多频带处理来改善特征的分类性能;ConvNet 和 EEGNet 在许多基于 CNN 的端到端框架的 EEG 数据集上显示了出色的结果;C2CM 将 FBCSP 的特征输入到 CNN 模型中,结合了传统特征提取方法和深度学习方法的优势;FBCNet 通过空间过滤多视图数据提取频谱-空间特征。表 1 展示了本研究方法与其他方法在数据集 2a 上的实验结果对比,所有对照方法的参数和实验结果均来自相关文献。可以看到,本文的方法相较于依赖传统特征提取的 FBCSP 显著提高了 13.2%的准确率($p<0.01$)。结果同样显示,其他深度学习方法,如 ConvNet 和 EEGNet,也表现优于 FBCSP,表明基于 CNN 的方法在特征表示能力上较强。然而,这些基于 CNN 的方法由于感知域的限制,仅关注局部特征而忽略了全局相关性,这可能影响脑电图序列的解码精度。与此不同,本文的方法通过在原始 CNN 架构中集成 Transformer,封装了局部和全局的依赖关系。因此,Talkformer 在大多数的受试者中取得了更好的结果,在平均准确率和 kappa 值上都有显著提升($p<0.05$)。FBCNet 有效地结合了手工特征和深度模型的思想,但除了受试者 2 和受试者 5 之外,其他情况下仍未能超越本文的方法,这再次证明了同时利用局部和全局特征的有效性。

表 1 本文方法与其他方法在数据集 2a 上的准确率与 Kappa 值

Table 1 Accuracy and Kappa value of this paper's method and other methods on dataset 2a

受试者	FBCSP ^[25]		ConvNet ^[13]		EEGNet ^[16]		C2CM ^[14]		FBCNet ^[26]		本文方法	
	Acc	Kappa	Acc	Kappa	Acc	Kappa	Acc	Kappa	Acc	Kappa	Acc	Kappa
S01	76.00	0.680 0	76.39	0.685 2	85.76	0.810 1	87.50	0.833 3	85.42	0.805 6	88.32	0.828 7
S02	56.50	0.420 0	55.21	0.402 8	61.46	0.486 1	65.28	0.537 1	60.42	0.472 3	58.33	0.444 4
S03	81.25	0.750 0	89.24	0.856 5	88.54	0.847 2	90.28	0.870 4	90.63	0.875 1	92.36	0.898 1
S04	61.00	0.480 0	74.65	0.662 0	67.01	0.560 1	66.67	0.555 6	76.39	0.685 2	86.25	0.815 7
S05	55.00	0.400 0	56.94	0.425 9	55.90	0.412 0	62.50	0.500 0	74.31	0.657 5	65.54	0.527 7
S06	45.25	0.270 0	54.17	0.388 9	52.08	0.361 1	45.49	0.273 2	53.82	0.384 3	66.71	0.564 8
S07	87.25	0.830 0	92.71	0.902 8	89.58	0.861 1	89.58	0.861 1	84.38	0.791 7	96.18	0.949 1
S08	81.25	0.750 0	77.08	0.694 4	83.33	0.777 7	83.33	0.777 7	79.51	0.726 8	88.58	0.833 3
S09	70.75	0.610 0	76.39	0.685 2	86.81	0.824 1	79.51	0.726 8	80.90	0.745 3	86.24	0.833 3
Average	67.75	0.576 7	72.53	0.633 7	74.50	0.659 9	74.46	0.659 5	76.20	0.682 6	80.95	0.743 9

表 2 展示了本研究方法与其他方法在数据集 2b 上的实验结果对比,所有对照方法的参数和实验结果均来自相关文献。可以看到,运动想象二分类任务呈现出与数据集 2a 相似的趋势,本文的方法与 FBCSP 相比提高了 4.79%的准确率($p<0.05$)。与仅使用 CNN 架构端到端方法的 ConvNet 和 EEGNet 相比,准确率分别提高了 5.42%($p<0.05$)和 4.31%($p<0.05$)。在几乎所有的被试上,本文方

法的平均准确率和 kappa 仍然高于 SPCNN,进一步验证了模型的有效性。

图 3(a)和图 3(b)分别展示了在数据集 2a 的 S03 样本集和数据集 2b 的 S07 样本集上应用本研究方法所得到的训练损失与准确率曲线。从图中可以观察到,随着训练损失的逐渐减少,准确率相应提升,特别是在最初的 200 个训练 epoch 期间,准确率迅速增加。这表明模型在初期阶

表 2 本文方法与其他方法在数据集 2b 上的准确率与 Kappa 值

Table 2 Accuracy and Kappa value of this paper's method and other methods on dataset 2b

受试者	FBCSP ^[25]		ConvNet ^[13]		EEGNet ^[16]		SPCNN ^[27]		本文方法	
	Acc	Kappa	Acc	Kappa	Acc	Kappa	Acc	Kappa	Acc	Kappa
S01	70.00	0.400 0	76.56	0.531 2	75.94	0.518 8	73.61	0.472 2	81.50	0.630 0
S02	60.36	0.207 2	50.01	0.000 2	57.64	0.152 8	75.74	0.514 8	66.36	0.327 2
S03	60.94	0.218 8	51.56	0.031 2	58.43	0.168 6	77.78	0.555 6	63.75	0.275 0
S04	97.50	0.950 0	96.88	0.937 6	98.13	0.962 6	93.24	0.864 8	96.31	0.926 2
S05	93.12	0.862 4	93.13	0.862 6	81.25	0.625 0	91.22	0.824 4	90.75	0.815 0
S06	80.63	0.612 6	85.31	0.706 2	88.75	0.775 0	81.94	0.638 8	89.33	0.786 6
S07	78.13	0.562 6	83.75	0.675 0	84.06	0.681 2	84.03	0.680 6	88.93	0.778 6
S08	92.50	0.850 0	91.56	0.831 2	93.44	0.868 8	84.21	0.684 2	92.38	0.847 6
S09	86.88	0.737 6	85.62	0.712 4	89.69	0.793 8	76.39	0.527 8	93.76	0.875 2
Average	80.00	0.600 1	79.37	0.587 5	80.48	0.616 3	82.02	0.640 4	84.79	0.695 7

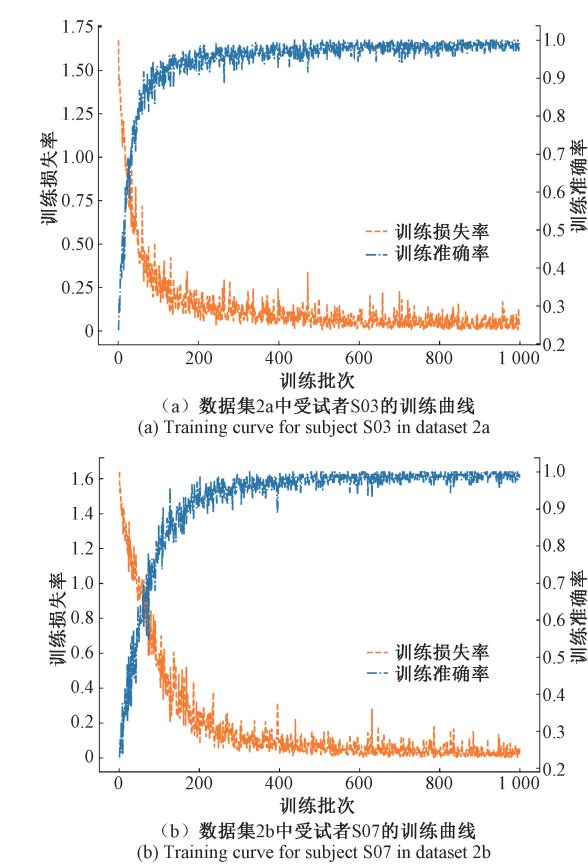


图 3 模型的训练损失和准确率曲线

Fig. 3 Training loss and accuracy curves of the model

段能够有效地学习数据特征,快速提高分类性能。当训练达到约 500 个 epoch 后,训练损失趋于稳定,这表明本文提出的方法在经过一段时间的学习后,已经能够较好地适应训练数据,并找到了较优的特征表示方法,从而表现出较高的准确性和稳定性。

为了深入探讨本方法在运动想象(MI)脑电信号分类中的性能表现,图 4 展示了数据集 2a 中受试者 S03 的分类结果混淆矩阵。该矩阵的横轴显示了所采用方法预测出的运动想象类别,而纵轴则展示了实际的运动想象类别。混淆矩阵最右侧的颜色条通过其色度的深浅来表示不同的分类正确率,其中,矩阵的主对角线上的数据点表示各个类别被正确分类的概率,而非对角线上的数据则表示了将某个类别错误分类到其他类别的情况。图 4 所示结果表明,在受试者 S03 上,模型表现出了较高的一致性。

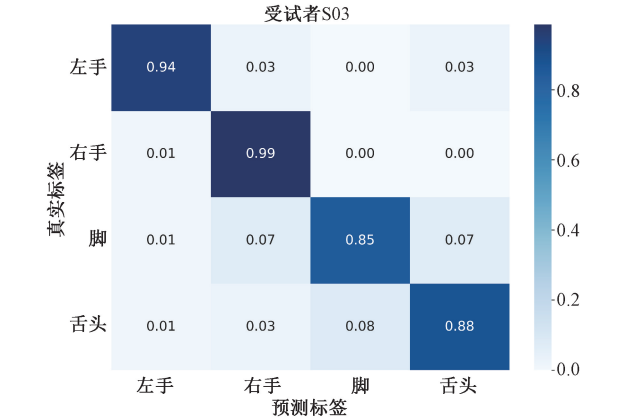


图 4 数据集 2a 中受试者 S03 的混淆矩阵

Fig. 4 Confusion matrix for subject S03 in dataset 2a

图 5 展示了数据集 2a 中受试者 S07 在 4 个类别上的 ROC 曲线和对应的 AUC 值。通过分析图中的数据,可以看出,模型在左手、右手、脚和舌头这 4 个类别上的分类性能表现出色。其中,左手的 AUC 值为 0.98,右手的 AUC 值为 0.97,脚的 AUC 值也为 0.97,这些高 AUC 值表明模型在这些类别上具有极高的识别准确性,且不同类别之间的表现相当一致,说明模型能够稳定地区分这些运动想象类别。尽管舌头类别的 AUC 值为 0.94,略低于其他类别,

但其识别能力仍然相当出色,展示了模型在这一类别上的较强区分能力。整体来看,模型在 4 个运动想象类别的多分类任务中,展现了优异的区分能力,具有高度准确性和稳定性,适用于需要高精度分类的应用场景中。

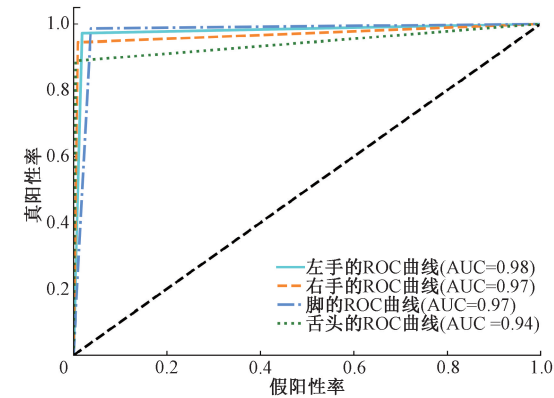


图 5 数据集 2a 中受试者 S07 的 ROC 曲线和 AUC 值

Fig. 5 ROC curves and AUC values for subject S07 in dataset 2a

3.2 消融实验

与传统的基于卷积神经网络(CNN)的方法相比,本文提出的 Talkformer 模型的关键改进是增加了基于注意力机制的模块来学习全局特征表示,从而更好地理解 and 分类复杂的数据结构,还增加了数据增强模块来提升模型对多样化数据分布的适应能力,从而增强其泛化性能,使其在处理新数据时仍能保持较高的准确率。为了验证 Talkformer 模型的有效性,本文对数据集 2a 进行了如下 3 个消融实验:同时去除注意力机制模块和数据增强模块(实验 1)、去除注意力机制模块(实验 2)、去除数据增强模块(实验 3)。如图 6 所示,当同时移除注意力机制模块和数据增强模块后,所有受试者的分类准确率显著下降,受试者 7 的准确率降幅最大,达到了 24.86%,而受试者 5 的准确率降幅最小,为 3.86%。总体而言,模型的平均准确率显著下降了 14.13%($p<0.01$)。当只移除注意力机制模块后,受试者 7 的准确率降幅最大,达到了 6.94%,而受试者 6 的准确率降幅最小,为 0.39%。总体而言,模型的平均准确率下降了 3.86%($p<0.01$)。此外,图 6 的实验结果也表明,数据增强策略对模型性能提升具有积极影响,与未采用数据增强的方法相比,模型的平均性能提升了 7.19%($p<0.01$),因此,在训练过程中引入数据增强策略,可以有效提升 Talkformer 模型的鲁棒性。

3.3 特征分布

t-SNE 是一种用于高维数据的降维和可视化技术。它通过保持高维数据中的局部结构,将数据映射到低维空间,使相似的数据点在低维空间中相互接近,便于观察和分析复杂数据的内在结构,从而揭示出高维数据中的复杂关系,该技术广泛应用于机器学习算法的数据输入、数据预处理以及数据可视化。数据集 2a 中受试者 S07 的特征

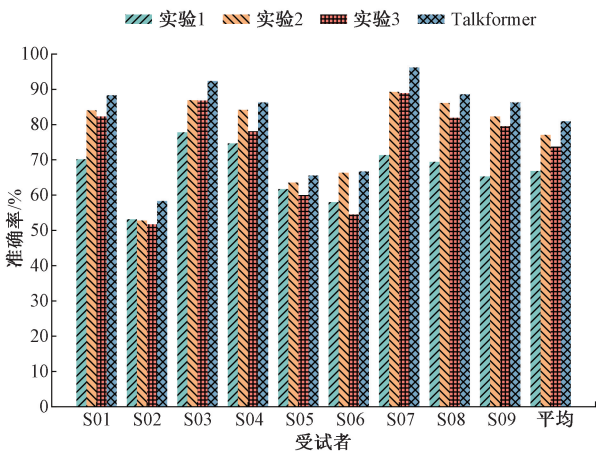
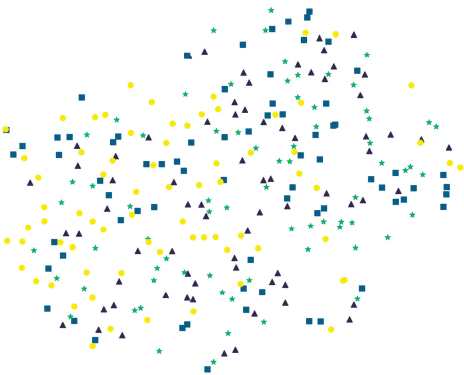


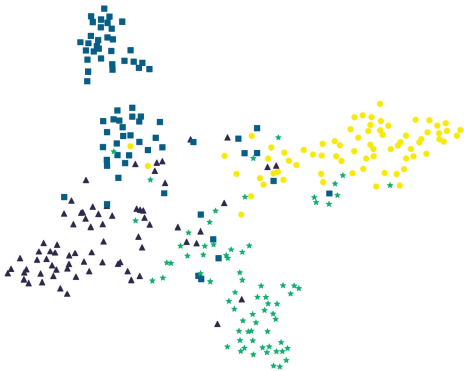
图 6 注意力机制模块和数据增强模块的消融实验

Fig. 6 Ablation experiments for the attention mechanism module and the data enhancement module

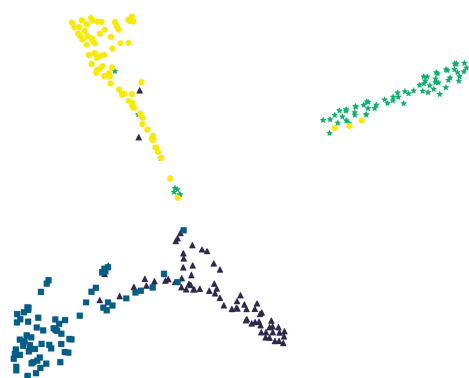
分布如图 7 所示,图 7(a)是原始脑电信号的测试数据分布,图 7(b)是不使用 Transformer 的脑电信号测试数据分布,图 7(c)是使用 Transformer 的脑电信号测试数据分布,可以清楚的观察到,在没有 Transformer 的帮助下,不同类别的特征是比较接近的。加入 Transformer 后,4 个类别的区分变得非常清晰。



(a) 原始脑电信号的测试数据分布
(a) Test data distribution of raw EEG signals



(b) 不使用transformer的脑电信号测试数据分布
(b) Distribution of EEG signal test data without using transformer



(c) 使用transformer的脑电信号测试数据分布
(c) Distribution of EEG signal test data using transformer

图 7 针对受试者 S07 提取的脑电特征经 t-SNE 降维后的二维分量的分布情况

Fig. 7 Distribution of two-dimensional components of EEG features extracted for subject S07 after t-SNE dimensionality reduction

4 结 论

本文针对运动想象脑电信号存在个体差异性强、信噪比低等特点而导致其识别困难以及 CNN 学习全局特征的局限性,提出了一种将 CNN 与 Transformer 相结合的脑电分类算法。采用简单的预处理方式后,通过时间和空间卷积层的卷积模块学习局部特征,并利用 Talking-Heads Attention 进一步关注长期依赖关系,从不同角度捕捉脑电信号的全局时间相关性。与现有的 CNN 模型相比,该模型能够学习到更多的特征表示,在四分类运动想象数据集上的分类精度更优,但模型的泛化能力不足,后续将通过迁移学习方法对网络模型进行优化来提高模型的泛化能力。

参考文献

- [1] 于鸿伟,谢俊,何柳诗,等. 基于卷积神经网络和支持向量机的事件相关电位识别方法[J]. 西安交通大学学报, 2021, 55(12): 47-54.
YU H W, XIE J, HE L SH, et al. Event-dependent potential recognition method based on convolutional neural network and support vector machine [J]. Journal of Xi'an Jiaotong University, 2021, 55(12): 47-54.
- [2] SINGH A, HUSSAIN A A, LAL S, et al. A comprehensive review on critical issues and possible solutions of motor imagery based electroencephalography brain computer interface[J]. Sensors, 2021, 21(6): 2173.
- [3] 李文平,徐光华,张凯,等. 一种高效运动想象脑电信号浅层卷积解码网络[J]. 西安交通大学学报, 2023, 57(10): 11-19.
LI W P, XU G H, ZHANG K, et al. An efficient shallow convolutional decoding network for motor

imagery EEG signals[J]. Journal of Xi'an Jiaotong University, 2023, 57(10): 11-19.

- [4] 蒋贵虎,陈万忠,马迪,等. 基于 ITD 和 PLV 的四类运动想象脑电分类方法研究[J]. 仪器仪表学报, 2019, 40(5): 195-202.
JIANG G H, CHEN W ZH, MA D, et al. Research on EEG classification method for four types of motor imagery based on ITD and PLV [J]. Chinese Journal of Scientific Instrument, 2019, 40(5): 195-202.
- [5] 王清杰,全海燕. 基于单形进化算法优化支持向量机的运动想象脑电分类研究[J]. 电子测量与仪器学报, 2021, 35(9): 157-163.
WANG Q J, QUAN H Y. A study on motor imagery EEG classification based on support vector machine optimized by monomorphic evolutionary algorithm[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(9): 157-163.
- [6] 李易,张本鑫,莫云,等. 特征可分性显式建模的跨数据库脑电解码方法[J]. 电子测量技术, 2024, 47(7): 95-105.
LI Y, ZHANG B X, MO Y, et al. A cross-database EEG decoding method for explicit modeling of feature separability[J]. Electronic Measurement Technology, 2024, 47(7): 95-105.
- [7] CHANDRA M A, BEDI S S. Survey on SVM and their application in image classification [J]. International Journal of Information Technology, 2021, 13(5): 1-11.
- [8] ZHAO S P, ZHANG B, YANG J, et al. Linear discriminant analysis [J]. Nature Reviews Methods Primers, 2024, 4(1): 70.
- [9] ZHANG S CH. Challenges in KNN classification[J]. IEEE Transactions on Knowledge and Data Engineering, 2021, 34(10): 4663-4675.
- [10] ZHANG J Y, LI K, YANG B H, et al. Local and global convolutional transformer-based motor imagery EEG classification [J]. Frontiers in Neuroscience, 2023, 17: 1219988.
- [11] 丛佩超,陈熙来,肖宜轩,等. 基于 SSVEP 信号的相频特性分类算法研究[J]. 电子测量技术, 2024, 47(5): 188-198.
CONG P CH, CHEN X L, XIAO Y X, et al. Research on phase-frequency characteristic classification algorithm based on SSVEP signal [J]. Electronic Measurement Technology, 2024, 47(5): 188-198.
- [12] SHERSTINSKY A. Fundamentals of recurrent neural network(RNN) and long short-term memory(LSTM)

- network[J]. *Physica D: Nonlinear Phenomena*, 2020, 404: 132306.
- [13] 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.
- [14] SAKHAVI S, GUAN C T, YAN SH CH. Learning temporal information for brain-computer interface using convolutional neural networks [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, 29(11): 5619-5629.
- [15] SUN B, ZHAO X, ZHANG H, et al. EEG motor imagery classification with sparse spectrotemporal decomposition and deep learning [J]. *IEEE Transactions on Automation Science and Engineering*, 2020, 18(2): 541-551.
- [16] 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.
- [17] LUO J, WANG Y J, XIA S X, et al. A shallow mirror transformer for subject-independent motor imagery BCI[J]. *Computers in Biology and Medicine*, 2023, 164: 107254.
- [18] TORTORA S, GHIDONI S, CHISARI C, et al. Deep learning-based BCI for gait decoding from EEG with LSTM recurrent neural network[J]. *Journal of Neural Engineering*, 2020, 17(4): 046011.
- [19] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. *Advances in Neural Information Processing Systems*, 2017, 30, DOI: 10.48550/arXiv.1706.03762.
- [20] SUN J Y, XIE J, ZHOU H H. EEG classification with transformer-based models [C]. 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (Lifetech). IEEE, 2021: 92-93.
- [21] LIU X L, SHI R Y, HUI Q X, et al. TCACNet: Temporal and channel attention convolutional network for motor imagery classification of EEG-based BCI[J]. *Information Processing & Management*, 2022, 59(5): 103001.
- [22] 宋耀莲,殷喜喆,杨俊.基于时空特征学习 Transformer 的运动想象脑电解码方法[J]. *南京大学学报(自然科学)*, 2023, 59(2): 313-321.
- SONG Y L, YIN X ZH, YANG J. An EEG decoding method for motor imagery based on spatio-temporal feature learning Transformer[J]. *Journal of Nanjing University(Natural Science)*, 2023, 59(2): 313-321.
- [23] ZHANG Y, NAM C S, ZHOU G X, et al. Temporally constrained sparse group spatial patterns for motor imagery BCI[J]. *IEEE Transactions on Cybernetics*, 2018, 49(9): 3322-3332.
- [24] LOTTE F. Signal processing approaches to minimize or suppress calibration time in oscillatory activity-based brain-computer interfaces [J]. *IEEE*, 2015, 103(6): 871-890.
- [25] FANG H, JIN J, DALY I, et al. Feature extraction method based on filter banks and Riemannian tangent space in motor-imagery BCI [J]. *IEEE Journal of Biomedical and Health Informatics*, 2022, 26(6): 2504-2514.
- [26] MANE R, CHEW E, CHUA K, et al. FBCNet: A multi-view convolutional neural network for braincomputer interface [J]. *ArXiv preprint arXiv: 2104.01233*, 2021.
- [27] ZHAO X F, LIU D, MA L, et al. Deep CNN model based on serial-parallel structure optimization for four-class motor imagery EEG classification[J]. *Biomedical Signal Processing and Control*, 2022, 72: 103338.

作者简介

胡博, 硕士, 主要研究方向为脑机接口与康复应用。

E-mail: q1092597661@163.com

谢俊(通信作者), 教授, 博士生导师, 主要研究方向为脑机接口技术与脑控机器人康复应用。

E-mail: xiejun@xju.edu.cn