

DOI:10.19651/j.cnki.emt.2519476

DBTNet:一种用于运动想象脑电解码的融合网络^{*}

李响 艾尔肯·亥木都拉 王路路

(新疆大学智能制造现代产业学院(机械工程学院)乌鲁木齐 830017)

摘要: 针对运动想象脑电信号(MI-EEG)空间分辨率低及个体差异性大导致的模型解码性能受限问题,创新性地提出了一种基于注意力双分支卷积网络(ADBCN)和时序多尺度注意力机制(TMSA)的DBTNet模型。该模型通过双分支卷积网络多尺度提取时空特征,结合高效多尺度注意力机制(EMA)强化对EEG信号的空间特征提取能力。随后,采用时序多尺度注意力机制以不同感受野同时捕获局部特征与全局依赖关系,从而获得更丰富的特征信息。最终通过分类器进行特征融合,实现高效解码。在受试者依赖评估方法中,该模型在BCI竞赛IV-2a数据集上的4分类准确率为86.57%,Kappa系数为0.8210;在BCI竞赛IV-2b数据集上的二分类准确率为88.95%,Kappa系数为0.7790。实验结果表明,DBTNet模型具有优异的解码性能。

关键词: 脑机接口;脑电图;多头注意力;运动想象;双分支卷积网络

中图分类号: TN911.7 **文献标识码:** A **国家标准学科分类代码:** 310.61;520.2040

DBTNet: A fusion network for motor imagery EEG decoding

Li Xiang Arkin Hamdulla Wang Lulu

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

Abstract: To address the limitations in model decoding performance caused by the low spatial resolution and high inter-subject variability of motor imagery EEG signals, this paper proposes a novel DBTNet model based on an attention-enhanced dual-branch convolutional network and a temporal multi-scale attention mechanism. The model employs a dual-branch convolutional network to extract multi-scale spatiotemporal features and integrates an efficient multi-scale attention mechanism to enhance the extraction of spatial features from EEG signals. Subsequently, a temporal multi-scale attention mechanism is applied to capture both local features and global dependencies under different receptive fields, thereby obtaining more comprehensive feature representations. Finally, a classifier is used to fuse the extracted features for efficient decoding. In subject-dependent evaluations, the proposed model achieves a four-class classification accuracy of 86.57% with a Kappa coefficient of 0.8210 on the BCI Competition IV-2a dataset, and a two-class classification accuracy of 88.95% with a Kappa coefficient of 0.7790 on the BCI Competition IV-2b dataset. Experimental results demonstrate that the DBTNet model achieves superior model decoding performance.

Keywords: brain-computer interface; electroencephalogram; multi-head attention; motor imagery; dual-branch convolutional networks

0 引言

脑机接口(brain-computer interface, BCI)技术通过记录大脑中的神经反应来识别不同的大脑活动,以产生各种控制信号或命令,从而实现大脑与外部设备之间的直接通信和控制^[1]。脑电信号可以通过非侵入性的方法来捕获,其中脑电图(electroencephalogram, EEG)因其高分辨

率、低成本、便携性和操作简单性而被广泛使用^[2]。

运动想象(motor imagery, MI)是指人在不执行肢体运动的情况下,通过想象特定动作引发大脑皮层电活动的过程^[3-5]。运动想象与脑电图的结合(motor imagery electroencephalogram, MI-EEG)在神经康复、残障人士辅助以及非医疗领域应用中取得了显著成效^[6-7]。然而,脑电信号本身具有低信噪比、非平稳性以及个体差异等特点,从

收稿日期:2025-07-29

^{*} 基金项目:国家自然科学基金(10971177)项目资助

而显著增加了 MI-EEG 高效解码的难度^[8]。因此,开发具有高效解码性能模型成为推动运动康复和辅助技术等应用发展的关键。

基于运动想象的脑机接口技术通常涉及 5 个阶段:脑电数据采集、信号处理、特征提取、分类以及最终反馈^[9]。传统机器学习(machine learning, ML)方法在 MI-EEG 信号解码中受限于人工特征提取,且难以有效捕捉 EEG 信号中复杂的非线性特征^[10]。深度学习(deep learning, DL)凭借自动提取非线性特征的优势,已被广泛应用于 EEG 信号处理,并推动了 MI-EEG 解码技术的发展。

卷积神经网络(convolutional neural network, CNN)作为深度学习中的一种经典模型,因其在原始 EEG 信号中提取时间和空间特征的出色表现而被广泛应用^[11]。Schirrmeyer 等^[12]提出了基于改进 CNN 架构的 Deep ConvNet 和 Shallow ConvNet 用于 MI-EEG 信号解码,并在性能上优于传统方法。Lawhern 等^[13]提出了轻量化 CNN 架构 EEGNet,以较少参数实现了更优的解码性能。Mane 等^[14]设计了 FBCNet,通过多视图数据和空间滤波提取光谱空间判别特征以实现 MI 分类。Li 等^[15]提出了 MTFB-CNN 网络,采用并行多尺度时频块提取时频特征,并结合注意力机制与残差模块实现特征融合。Liu 等^[16]提出了多分支一维卷积网络 CMO-CNN,通过使用不同卷积核尺寸和分支深度提取多层次特征进行特征融合。Tang 等^[17]设计了多尺度混合 CNN 架构(MSHCNN),利用一维卷积提取高级时间特征,二维卷积提取时空特征。然而,上述基于改进 CNN 的方法受限于局部感受野和固定卷积核尺寸,难以捕捉时间序列中的全局依赖关系,导致其在复杂 MI-EEG 解码任务中主要侧重局部特征提取,限制了模型整体性能^[18]。

近年来,随着深度学习的快速发展,基于自注意力机制(self-attention, SA)的 Transformer^[19]网络架构在自然语言处理(natural language processing, NLP)和计算机视觉领域中取得了显著成功,激发了研究人员尝试将该模型用于 MI-EEG 解码任务。其中,自注意力机制的全局感受野能够有效捕获 EEG 信号的全局依赖关系,从而显著提升模型性能。Altafheri 等^[20]设计了基于注意力的时间卷积网络 EEG-ATCNet,取得了优异的解码性能。Zhao 等^[21]受 EEGNet 与 Transformer 的启发设计了 CTNet,通过结合

CNN 与 Transformer 编码器实现了较高的解码精度。Ma 等^[22]提出了 EEG-TransNet,采用多角度提取多模态时序特征,并利用自注意力机制捕获全局依赖关系,从而提升网络的泛化能力。Liang 等^[23]提出了 EISATC-Fusion,通过 CNN 与 cnnCos 注意力结合的方式,解决了注意力坍塌问题。然而,这些基于 Transformer 的方法在 MI-EEG 解码任务中对局部特征的提取能力有限,可能会遗漏关键局部细节,从而限制了模型解码性能的进一步提升。

为解决上述问题,本文提出了一种基于注意力双分支卷积网络(attentional dual-branch convolutional networks, ADBCN)和时序多尺度注意力机制(temporal multi-scale attention, TMSA)的新型融合网络 DBTNet 模型,用于 MI-EEG 解码任务。在两个公开数据集的结果表明,所提模型具有优异的解码性能。本文主要贡献为:

1)提出了一种新型的注意力双分支卷积网络,采用双分支卷积网络多尺度提取时空特征,并利用高效多尺度注意力机制(efficient multi-scale attention, EMA)增强对 MI-EEG 信号中局部和全局通道(空间)特征的捕获能力,显著提升模型解码性能。

2)采用时序多尺度注意力机制,通过利用局部和全局感受野同时捕获局部特征与全局依赖关系,从而提升特征信息的表征能力。

3)在 BCI 竞赛 IV-2a^[24]和 BCI 竞赛 IV-2b^[25]数据集上的实验结果表明,DBTNet 模型在受试者依赖评估方法中表现出了优异的性能和泛化能力,为 MI-EEG 信号解码领域做出了重要贡献。

1 研究方法

1.1 网络模型

本文提出了一种用于 MI-EEG 解码的 DBTNet 模型,通过结合 CNN 与注意力机制实现高效解码。模型整体结构如图 1 所示,主要包含 3 个模块:注意力双分支卷积网络、时序多尺度注意力机制以及分类器。首先,ADBCN 模块多尺度提取原始 MI-EEG 信号的时空特征并增强模型对 EEG 信号的空间信息捕获能力。随后,利用位置编码(positional encoding, PE)引入时序位置信息,通过时序多尺度注意力机制捕获局部特征与全局依赖关系。最后,经过由展平层和全连接层组成的分类器输出最终的解码结果。

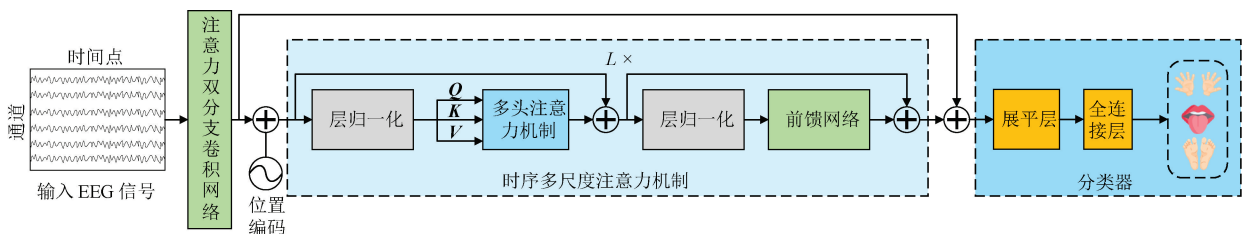


图 1 DBTNet 模型结构

Fig. 1 DBTNet model structure

1.2 注意力双分支卷积网络

ADBCN 模块主要包括双分支卷积网络与 EMA, 具体结构如图 2 所示。双分支卷积网络由时序卷积和双分支卷积结构组成, 其中时序卷积提取 MI-EEG 信号的时间特征, 双分支卷积结构多尺度地提取空间特征, 并通过相加融合

实现信息整合。EMA 进一步通过跨空间信息聚合方法增强通道(空间)特征的表达能, 从而提高 MI-EEG 信号的空间分辨率。最终, ADBCN 生成一个高级时间序列。该设计在增强 MI-EEG 信号时空特征提取能力的同时, 突出关键信息并减少过拟合, 从而提升模型解码精度和泛化能力。

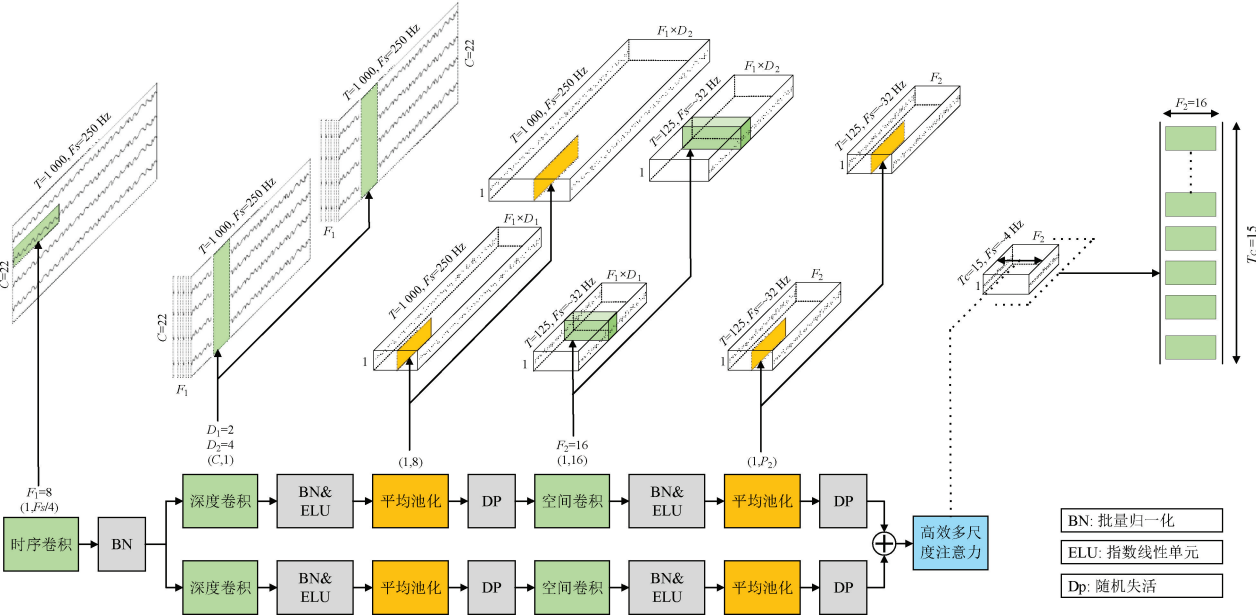


图 2 注意力双分支卷积网络

Fig. 2 Attentional dual-branch convolutional networks

双分支卷积网络的第 1 部分采用 F_1 个卷积核大小为 $(1, K_{c1})$ 的时序卷积层提取时间特征, 输出 F_1 个时序特征图。其中 K_{c1} 表示沿时间轴方向的卷积核长度, 并设置为采样率 F_s 的 1/4 (即 $K_{c1} = 64$), 从而使该层能够有效捕获频率高于 4 Hz 时间信息。第 2 部分采用双分支卷积结构, 各分支初始层均使用卷积核大小为 $(C, 1)$ 的深度卷积层提取空间特征, 分别输出 $F_1 \times D_1$ 和 $F_1 \times D_2$ 个空间特征图, 其中 D 表示各分支中每个时序特征图用于提取空间特征的卷积核数量, 并将 D_1 与 D_2 分别设置为 2 和 4。同时, C 表示 EEG 电极通道数 (BCI 竞赛 IV-2a 与 IV-2b 数据集的通道数 C 分别为 22 和 3)。在深度卷积层之后, 通过大小为 $(1, 8)$ 的平均池化层将 EEG 信号的采样率 F_s 和时间序列长度 T 降低 8 倍, 即 $F_s / 8 = \sim 32$ Hz 和 $T / 8 = 125$ 。随后, 采用 F_2 个卷积核大小为 $(1, K_{c2})$ 的空间卷积层提取空间特征。为确保输出一致性, 两个卷积分支的 K_{c2} 均设置^[6]为 16, 用于解码约 32 Hz 采样数据 500 ms 窗口内的运动想象 (MI) 活动。随后, 利用大小为 $(1, P_2)$ 的第 2 个平均池化层将采样率 F_s 和时间序列长度 T 进一步降至 $\sim 32 / P_2 = \sim 4$ Hz 和 $125 / P_2 = 15$, 其中 P_2 设置为 8。最后, 将两个卷积分支获得的时间序列进行相加融合。

在双分支卷积网络中, 各分支的深度卷积层与空间卷积层后均使用批量归一化 (batch normalization, BN) 提升训练效率, 并使用指数线性单元 (exponential linear units,

ELU) 作为激活函数引入非线性变换。在平均池化层之后, 均采用丢弃率为 0.5 的随机失活操作以缓解过拟合问题。ADBCN 模块最终会输出一个高级时间序列 $\mathbf{X} \in \mathbf{R}^{T_c \times d}$, 其中 T_c 表示时间序列长度。随后, 将各时间点的多特征通道组合成向量, 输入至时序多尺度注意力机制中进一步处理, 以捕捉高级时间序列中的局部特征与全局依赖关系。 T_c 计算过程如式 (1) 所示。

$$T_c = \frac{T}{8 \times P_2} \tag{1}$$

式中: T 表示原始 EEG 信号的时间序列长度, d 表示特征通道数, 即 $d = F_1 \times D_1 = F_2 = 16$ 。

Conformer^[26] 模型采用时序卷积层和空间卷积层提取时空特征, CTNet^[21] 在此基础上引入深度卷积层进一步提升了分类性能。然而, 此类架构在捕捉信号中复杂的空间依赖关系及通道间交互方面仍存在一定局限性, 可能导致空间特征的丢失。因此, 本文引入 EMA 强化模型对 MI-EEG 信号的空间特征提取能力。

EMA^[27] 是一种无需通道降维的高效多尺度注意力机制, 具体结构如图 3 所示。EMA 将给定输入特征图 $\mathbf{X} \in \mathbf{R}^{C \times H \times W}$ (C 表示输入通道数, H 和 W 分别表示特征图的高和宽) 的通道维度划分为 G 组特征图, 通过 3 个并行分支提取每组特征图的注意力权重; 两个 1×1 分支分别沿水平和垂直方向采用一维全局平均池化操作对通道信息进

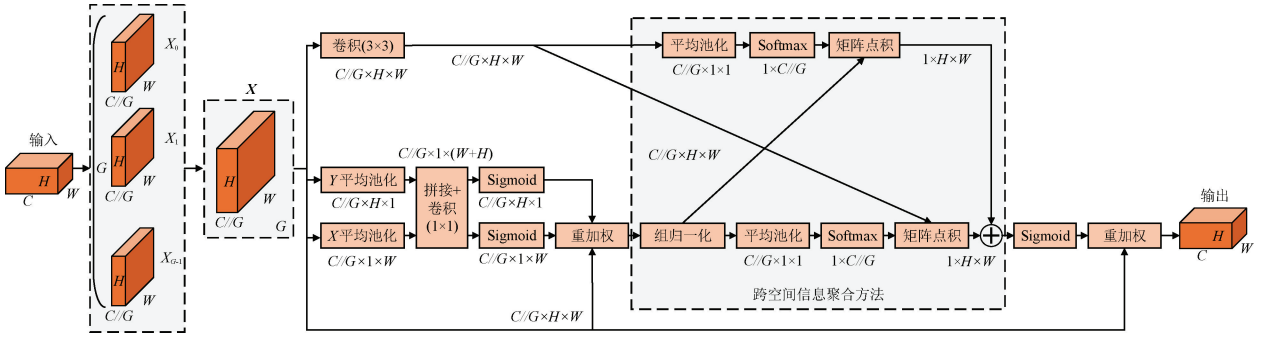


图3 高效多尺度注意力机制

Fig. 3 Efficient multi-scale attention mechanisms

行编码以捕捉全局信息,并沿高度方向进行拼接。利用共享的 1×1 卷积将输出分解为水平和垂直方向的向量后,分别采用 Sigmoid 函数获得权重,最后对特征图进行重加权操作以实现不同分支间的跨通道特征交互; 3×3 分支则用于捕获多尺度空间特征表示。

EMA 模块采用跨空间信息聚合方法,实现更丰富的特征融合。具体而言,在 1×1 和 3×3 分支中分别使用二维全局平均池化编码全局空间信息,并利用 Softmax 拟合线性变换,生成两个空间注意力图。最后,通过 Sigmoid 函数处理得到每组特征图的输出。具体过程如式(2)所示。

$$Z_c = \frac{1}{H \times W} \sum_j^H \sum_i^W x_c(i, j) \quad (2)$$

式中: x_c 表示第 c 个通道的输入特征; Z_c 则表示在第 c 个通道上的二维平均池化结果。

1.3 时序多尺度注意力机制

由于 EEG 信号是时间序列数据,缺乏时序位置信息可能导致模型难以有效捕捉信号随时间变化的动态特征,进而影响模型的性能。为此,通过引入位置编码为模型提供每个输入特征的时序位置信息,增强对时序依赖关系的学习能力^[19]。即采用基于正弦和余弦函数的位置编码方法,使模型能够捕捉时间序列数据中的位置关系,具体过程为:

$$PE_{(pos, i)} = \begin{cases} \sin\left(\frac{pos}{10000^{2i/d}}\right), & i = 0, 2, 4, \dots \\ \cos\left(\frac{pos}{10000^{2i/d}}\right), & i = 1, 3, 5, \dots \end{cases} \quad (3)$$

式中: pos 表示该数据在时间序列中的位置,即 $pos = 0, 1, 2, \dots, L-1$; i 表示偶数和奇数维度, d 则表示输入至多尺度注意力机制的特征维度。

为同时捕捉 MI-EEG 信号中的局部特征与全局依赖关系,受 TMSA-Net^[28]的启发,引入一种称为时序多尺度注意力机制的新型自注意力机制模块。该模块旨在通过不同感受野捕获不同时间段特征间的时序依赖关系,从而显著提升模型的解码性能。

TMSA 模块的核心部分包括多尺度一维卷积层与多头注意力机制(multi-head attention, MHA),具体结构如

图4所示。多尺度一维卷积层通过多个不同大小的卷积核来捕获不同时间尺度的特征。将各卷积分支的输出沿通道维度拼接获得多尺度的特征表示。具体而言,给定维度为 $X \in \mathbb{R}^{T \times C \times d}$ 的输入,采用卷积核大小为 $\{k_1, k_2, k_3, \dots, k_n\}$ 与对应填充为 $\{p_1, p_2, p_3, \dots, p_n\}$ 的一维卷积层,将输出沿通道维度进行拼接,具体过程为:

$$Concat(ConvlD(X, k_1, p_1), \dots, ConvlD(X, k_n, p_n)) \quad (4)$$

$$y_i = ConvlD(X, k_i, p_i) \quad (5)$$

$$k_{local} = Concat(y_1, y_2, \dots, y_n) \quad (6)$$

式中: k_i 表示卷积核大小, p_i 表示对应填充大小。

多头注意力机制包含线性层与缩放点积注意力机制(scaled dot-product attention),计算过程分为局部注意力(local attention)与全局注意力(global attention)两个分支,旨在通过不同感受野捕获局部特征与全局依赖关系。

在局部注意力中,首先将多尺度一维卷积层提取的局部特征 k_{local} 通过线性变换分别映射为查询(query, Q)、局部键(K_{local})和值(value, V)矩阵。随后对投影矩阵进行维度排列和转置以计算注意力得分,经过 softmax 归一化生成注意力权重并应用随机失活。最终,将注意力权重与值矩阵相乘输出局部注意力结果。具体过程为:

$$score_{local} = \frac{QK_{local}^T}{\sqrt{d_k}} \quad (7)$$

$$attention_{local} = dropout(\text{softmax}(score_{local})) \quad (8)$$

$$x_{local} = attention_{local}V \quad (9)$$

式中: d_k 表示键向量的维度。

在全局注意力中,直接输入原始序列 k 以类似方式输出全局注意力结果,具体过程为:

$$score_{globe} = \frac{QK_{globe}^T}{\sqrt{d_k}} \quad (10)$$

$$attention_{globe} = dropout(\text{softmax}(score_{globe})) \quad (11)$$

$$x_{globe} = attention_{globe}V \quad (12)$$

首先对局部注意力与全局注意力的输出特征进行相加融合,随后将融合特征沿维度拼接,最终通过线性层生成输出表示。具体过程为:

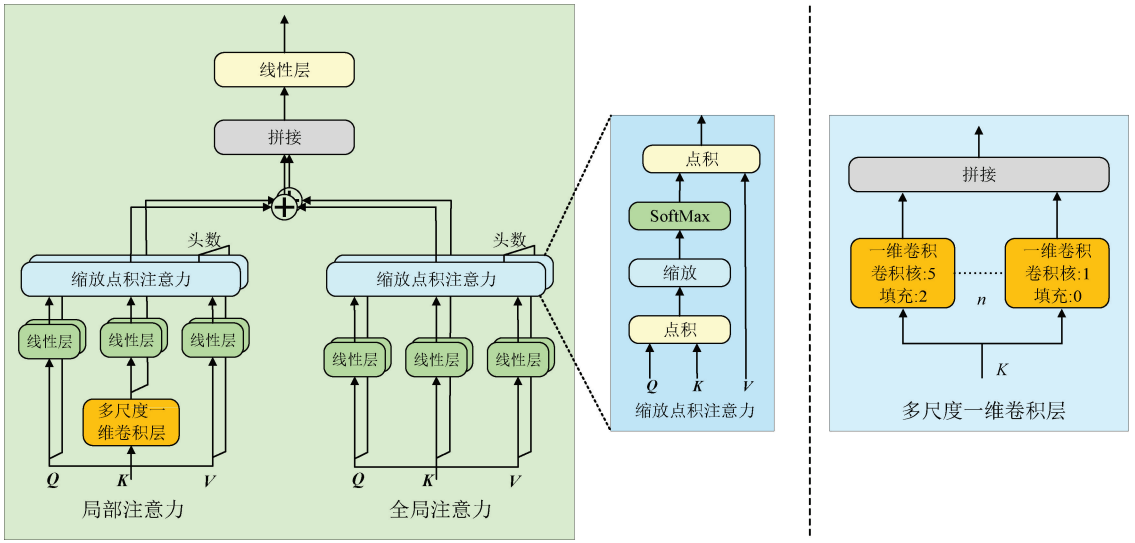


图 4 多头注意力机制

Fig. 4 Multi-head attention mechanisms

$$x = x_{local} + x_{globe} \quad (13)$$

$$MHA_{out} = linear(concat(x)) \quad (14)$$

$$x = x + MHA_{out}(LN(x)) \quad (15)$$

前馈网络(feed-forward network, FFN)由两个线性变换层组成,并结合高斯误差线性单元(Gaussian error linear unit, GELU)作为激活函数。具体过程为:

$$GELU(x) = \frac{x}{2} \left[1 + erf\left(\frac{x}{\sqrt{2}}\right) \right] \quad (16)$$

$$FFN(x) = GELU(xW_1 + b_1)W_2 + b_2 \quad (17)$$

$$Out = x + FFN(LN(x)) \quad (18)$$

式中: $erf(x)$ 表示误差函数, W_1 和 W_2 分别表示 FFN 中两个线性变换层的权重矩阵, b_1 和 b_2 分别表示对应的偏置项。

在时序多尺度注意力机制中,多头注意力机制与前馈网络前均应用层归一化(layer norm, LN)处理并引入残差连接^[29]。此种结合局部注意力与全局注意力的方法可同时捕获局部特征与全局依赖关系,从而显著提升模型解码性能。该过程在 TMSA 模块中重复 L 次,其中 L 为模块深度。

1.4 分类器

将 ADCBN 模块和时序多尺度注意力机制的输出特征相加融合后输入至展平层进行展平操作。该设计能够使 ADCBN 提取的局部特征直接输入至分类器中。然后,利用丢弃率为 0.5 的随机失活减少过拟合。最终,展平后的特征输入至一个包含 N 个单元的全连接层进行分类,其中 N 表示 MI 分类任务中的类别数。模型训练过程中使用交叉熵作为损失函数,具体过程为:

$$L = -\frac{1}{M} \sum_i^M \sum_j^N y_{ij} \log(\hat{y}_{ij}) \quad (19)$$

式中: M 表示 EEG 试验次数, y_{ij} 和 \hat{y}_{ij} 分别表示第 i 个样

本中第 j 个类别的真实标签和预测标签。

2 实验过程

2.1 数据集介绍

为验证 DBTNet 的分类性能及泛化性,在两个公开 MI-EEG 数据集上对该模型进行了评估,即 BCI 竞赛 IV-2a^[24] 和 IV-2b 数据集^[25]。

BCI 竞赛 IV-2a 数据集共包含 9 名受试者的脑电数据,包括 4 个运动想象任务:左手、右手、双脚和舌头。数据采集使用 22 个 Ag/AgCl 电极,信号采样频率为 250 Hz。实验使用 [2 s, 6 s] 作为分析时段,对信号进行 [0.5 Hz, 100 Hz] 的带通滤波处理,同时使用 50 Hz 的陷波滤波器去除噪声。每个受试者均在不同的两天分别参与一次实验,使用第 1 次的实验数据作为训练集,第 2 次的实验数据作为测试集。每次实验包含 288 次运动想象任务,每个任务类别包括 72 次实验。

BCI 竞赛 IV-2b 数据集共包含 9 名受试者的脑电数据,包括左手和右手的运动想象任务。通过 3 个双极电极通道 (C_3 , C_2 , C_4) 进行数据采集,采样频率为 250 Hz。实验数据分析时段为 [3 s, 7 s],脑电信号经过 [0.5 Hz, 100 Hz] 的带通滤波处理,并使用 50 Hz 的陷波滤波器去除噪声。每名受试者参与五次实验,使用前第 3 的实验数据作为训练集,后两次实验数据作为测试集。训练集包含 400 次实验,测试集包含 320 次实验。

本研究在两个数据集上均采用受试者依赖的评估方法,即对于每个受试者的数据独立进行模型训练与测试。模型训练时,每个训练集的 30% 被划分为验证集,用于评估模型性能。为确保结果的有效性,对两个数据集上的结果均进行 10 次随机重复实验,取实验结果的均值作为模型的最终性能结果。

2.2 数据集预处理

在预处理阶段,仅去除 EEG 信号中的 3 个心电图 (EOG)通道,不进行伪影去除。随后,通过 Z 分数标准化进行预处理,以确保数据尺度一致性并提升模型训练效率。每次原始试验 $x_i \in \mathbf{R}^{C \times T}$ 包含 C 个通道的电极信号和 T 个时间点的样本。标准化处理过程为:

$$x'_i = \frac{x_i - \mu}{\sqrt{\sigma^2}}, i = 1, 2, 3, \dots \quad (20)$$

式中: $x'_i \in \mathbf{R}^{C \times T}$ 表示标准化处理后的结果; μ 和 σ^2 表示均值和方差。

该模型旨在将标准化处理后的运动想象 (MI) 试验输入 x'_i 与其对应的类别 y_i 关联。即使用包含 m 个标记的 MI 实验样本的集合 $S = \{x'_i, y_i\}_{i=1}^m$, 其中 $y_i = \{1, 2, 3, \dots, N\}$ 表示第 i 个 MI 实验对应的类别标签, N 表示类别总数。在 BCI 竞赛 IV-2a 数据集中,时间序列长度 T 为 1 000,包含 5 184 个 MI 试次;在 BCI 竞赛 IV-2b 数据集中,时间序列长度 T 为 1 000,包含 6 520 个 MI 试次。

2.3 数据增强

为解决运动想象 (MI) 任务中 EEG 数据集规模有限的问题,本文引入时间域分割与重建 (segmentation and recombination, S&R)^[30] 数据增强方法,优化深度学习模型训练过程,降低过拟合风险。

如图 5 所示,对于给定维度为 $x'_i \in \mathbf{R}^{C \times T}$ 的输入,构建 M 个 EEG 试验集合 $\Gamma \in \{x'_i\}_{i=1}^M$ 。将每个 EEG 试验 x'_i 按原始时间维度分割为 L 个连续片段 $x''_l \in \mathbf{R}^{C \times T/L}$, 其中 $l \in [1, L]$ 。然后,随机选取 L 个片段进行重组以生成新的人工试验 $x'''_i = [x''_{r_1}, x''_{r_2}, x''_{r_3}, \dots, x''_{r_L}] \in \mathbf{R}^{C \times T}$, 其中 $r_L \in [1, M]$, $[A, B, C]$ 表示将片段 A、B、C 的样本依次拼接。该方法可以有效保留时间顺序,并增强训练数据的多样性,从而提升模型的泛化能力。

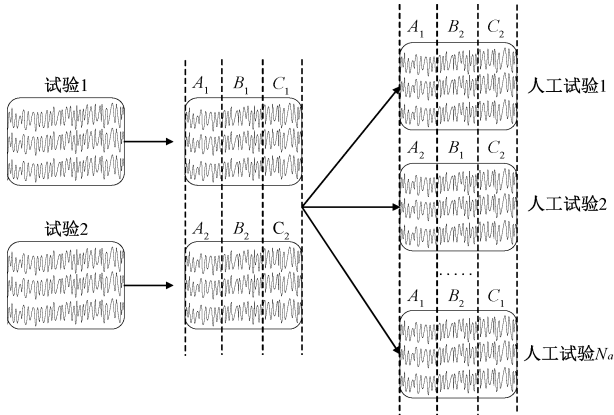


图 5 数据增强

Fig. 5 Data augmentation

2.4 实验环境

本研究的实验环境是 Windows 11 23 H2 操作系统,配置为 Intel i7-10700k 处理器与 NVIDIA GeForce RTX

2080 Super 显卡 (8 GB 显存)。模型采用 Python 3.10 及 PyTorch 1.13.1 框架,使用 Adam 优化器训练模型,学习率设置为 0.001, β_1 和 β_2 分别为 0.5 与 0.999。注意力机制的头数 h 设置为 2,深度 L 设置为 6。在训练过程中,训练轮次为 2 000,批量大小为 288,丢弃率为 0.5。在两个数据集上的所有实验均执行相同的训练配置。

2.5 评价指标

本研究采用准确率 (Accuracy) 和 Kappa 系数来评估模型的解码性能。具体过程为:

$$Accuracy(Acc) = \frac{TP + TN}{TP + TN + FP + FN} \quad (21)$$

式中: TP 和 TN 分别表示模型正确预测为正类和负类的样本数, FP 和 FN 则分别表示错误预测为正类和负类的样本数。

$$Kappa = \frac{p_o - p_e}{1 - p_e} \quad (22)$$

式中: p_o 表示观测准确度,而 p_e 表示期望准确度。

3 实验结果分析

3.1 分类结果对比

本节旨在将 DBTNet 与几种先进的方法进行比较,对比方法包括 Conformer^[26]、CTNet^[21]、EISATC-Fusion^[23] 以及 EEGNet^[13]。为确保实验公平性,所有实验均采用相同的预处理方法和数据集划分方式。同时,在复现对比模型时,相关配置参数均使用各自文献中的设定。

表 1 比较了本文方法与现有先进方法在 BCI 竞赛 IV-2a 数据集上的性能对比。实验结果表明,DBTNet 分类准确率达到了 86.57%, Kappa 系数为 0.821 0,标准差为 7.95%,模型分类性能优于其他方法。与基于 CNN 的 EEGNet 相比,DBTNet 在分类准确率上提升了 9.56%。这一结果表明,该模型通过采用 ADBCN 与 TMSA 的设计,有效克服了传统 CNN 方法在捕捉全局依赖关系方面的不足,从而提升了模型的解码性能。Conformer、CTNet 和 EISATC-Fusion 作为 CNN 与 Transformer 结合的融合模型,兼顾了局部和全局特征,但在关键局部细节与全局依赖关系的交互作用上仍存在不足。相比之下,DBTNet 在 CNN 框架中结合了双分支卷积网络与 EMA,能够多尺度地有效提取局部时空特征,并利用 TMSA 进一步关注局部特征与全局依赖关系,从而显著提升了模型性能。实验结果表明,DBTNet 分类准确率较 Conformer、CTNet 和 EISATC-Fusion 分别提高了 7.52%、3.82% 和 3.16%,充分验证了该模型的有效性。

为验证 DBTNet 的泛化性,在 BCI 竞赛 IV-2b 数据集上与其他方法的分类性能进行了对比。如表 2 所示,结果与在 BCI 竞赛 IV-2a 数据集上的结果相似,并且在所有深度学习方法中取得了最佳分类性能。具体而言,DBTNet 的分类准确率达到 88.95%, Kappa 系数为 0.779 0,标准差为

表 1 BCI 竞赛 IV-2a 数据集与其他模型的性能对比

Table 1 Performance comparison with other models on the BCI competition IV-2a dataset

受试者	DBTNet		Conformer ^[26]		CTNet ^[21]		EISATC-Fusion ^[23]		EEGNet ^[13]	
	Acc/%	Kappa	Acc/%	Kappa	Acc/%	Kappa	Acc/%	Kappa	Acc/%	Kappa
S01	93.06	0.907 4	85.07	0.800 9	88.89	0.851 9	87.50	0.833 3	86.11	0.814 8
S02	75.69	0.675 9	56.60	0.421 3	72.92	0.638 9	71.18	0.615 7	64.58	0.527 8
S03	95.49	0.939 8	92.71	0.902 8	92.71	0.902 8	95.14	0.935 2	93.40	0.912 0
S04	87.15	0.828 7	77.78	0.703 7	81.25	0.750 0	84.03	0.787 0	62.50	0.500 0
S05	79.86	0.731 5	75.00	0.666 7	75.69	0.675 9	75.35	0.671 3	71.88	0.625 0
S06	74.31	0.657 4	65.97	0.546 3	66.32	0.550 9	72.22	0.629 6	58.33	0.444 4
S07	93.06	0.907 4	87.85	0.838 0	90.97	0.879 6	90.97	0.879 6	85.42	0.805 6
S08	89.93	0.865 7	86.11	0.814 8	88.54	0.847 2	86.46	0.819 4	85.07	0.800 9
S09	90.63	0.875 0	84.38	0.791 7	87.50	0.833 3	87.85	0.838 0	85.76	0.810 2
Avg	86.57	0.821 0	79.05	0.720 7	82.75	0.770 1	83.41	0.778 8	77.01	0.693 4
Std	7.950	0.106 0	11.57	0.154 3	9.210	0.122 9	8.520	0.113 6	12.77	0.170 2

表 2 BCI 竞赛 IV-2b 数据集与其他模型的性能对比

Table 2 Performance comparison with other models on the BCI competition IV-2b dataset

受试者	DBTNet		Conformer ^[26]		CTNet ^[21]		EISATC-Fusion ^[23]		EEGNet ^[13]	
	Acc/%	Kappa	Acc/%	Kappa	Acc/%	Kappa	Acc/%	Kappa	Acc/%	Kappa
S01	80.31	0.606 3	72.19	0.443 8	76.25	0.525 0	75.94	0.518 8	79.06	0.581 3
S02	69.29	0.385 7	70.00	0.400 0	71.79	0.435 7	66.43	0.328 6	68.57	0.371 4
S03	85.94	0.718 8	80.31	0.606 3	86.88	0.737 5	82.81	0.656 3	85.31	0.706 3
S04	98.44	0.968 8	97.50	0.950 0	98.13	0.962 5	97.19	0.943 8	97.81	0.956 3
S05	98.75	0.975 0	96.25	0.925 0	95.63	0.912 5	95.31	0.906 3	95.00	0.900 0
S06	87.19	0.743 8	81.25	0.625 0	86.56	0.731 3	84.06	0.681 3	86.88	0.737 5
S07	95.31	0.906 3	90.31	0.806 3	93.75	0.875 0	92.81	0.856 3	93.75	0.875 0
S08	95.00	0.900 0	93.75	0.875 0	94.69	0.893 8	94.69	0.893 8	93.13	0.862 5
S09	90.31	0.806 3	88.75	0.775 0	89.38	0.787 5	89.38	0.787 5	90.63	0.812 5
Avg	88.95	0.779 0	85.59	0.711 8	88.12	0.762 3	86.51	0.730 3	87.79	0.755 9
Std	9.610	0.192 2	10.15	0.202 9	8.980	0.179 5	10.25	0.205 1	9.200	0.184 0

9.61%。相比于 Conformer、CTNet、EISATC-Fusion 和 EEGNet, DBTNet 的分类准确率分别提高了 3.36%、0.83%、2.44%和 1.16%。尽管相较于 CTNet 和 EEGNet 的整体分类性能提升幅度较小,但在大多数受试者中, DBTNet 仍表现出了更高的分类准确率。

综合上述两个公开数据集的实验结果,共同验证了所提模型在 MI-EEG 解码任务中的可行性和优越性。

3.2 混淆矩阵

混淆矩阵能够直观反映真实类别与预测类别之间的对应关系。通过对混淆矩阵的分析,可以全面了解模型在 MI 分类任务中的效果。图 6 展示了不同模型(DBTNet、Conformer、CTNet、EISATC-Fusion 及 EEGNet)在 BCI 竞赛 IV-2a 和 IV-2b 数据集上的混淆矩阵结果,用于评估分类性能。图中横轴表示模型预测类别,纵轴表示真实类别,主对角线元素表示各自 MI 分类任务被正确预测的

比例。

在 BCI 竞赛 IV-2a 数据集中, DBTNet 在四分类任务中均获得了最高的类别正确率;在 IV-2b 数据集的二分类任务中,模型在左手 MI 分类任务中表现出较高的类别正确率,同时在右手 MI 分类任务上取得了最高的类别正确率,验证了所提模型的有效性。

3.3 参数选取

为评估模型中几个关键参数对解码性能的影响,在 BCI 竞赛 IV-2a 和 IV-2b 数据集上进行了实验,主要参数包括:深度卷积层输出通道数、时序多尺度注意力机制深度和注意力头数。

1)输出通道数选取:分析双分支卷积网络中,第 2 分支($D_2=4$)的深度卷积层输出通道数变化对 DBTNet 模型分类准确率的影响。

深度卷积层的输出通道数 F_{out} 由式(23)确定,其中 F_1

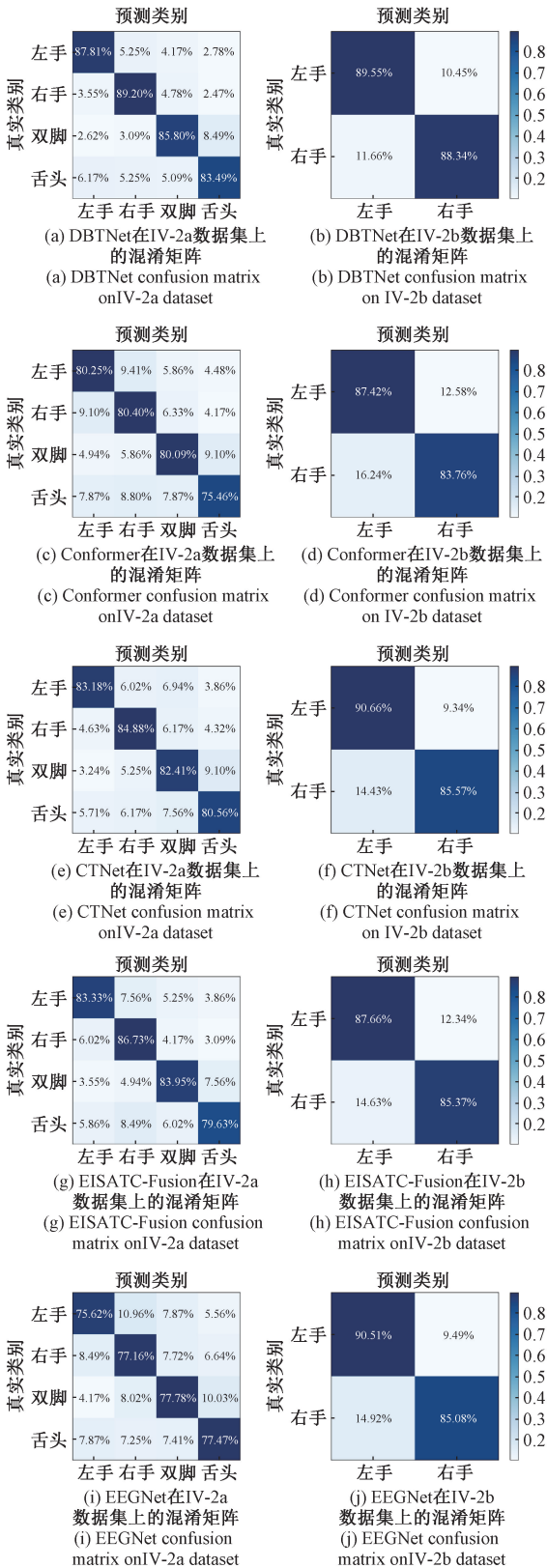


图 6 不同模型在 2 个数据集上的混淆矩阵

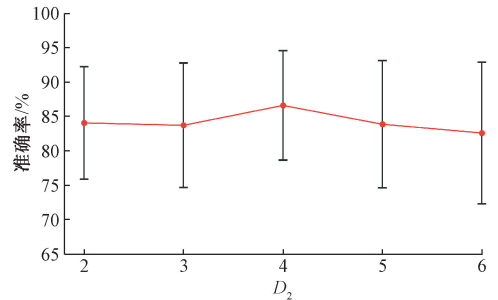
Fig. 6 Confusion matrices of different models on two datasets

表示输入通道数, D_i 表示第 i 个分支中用于提取空间特征

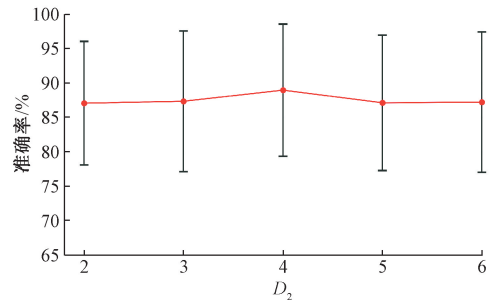
的卷积核数量。其中,两个卷积分支中深度卷积层的输入通道数 F_1 均设置为 8,将第 1 分支的参数 D_1 固定^[6,21] 为 2,以探讨第 2 分支中参数 D_2 的变化对模型解码性能的影响。其中, D_2 的取值范围为 2~6 的整数。图 7 展示了在固定参数 D_1 的情况下,随着 D_2 从 2 增加到 6,模型分类准确率的变化趋势。

$$F_{outi} = F_1 \times D_i, i = 1, 2 \quad (23)$$

实验结果表明,随着参数 D_2 增加,模型在两个数据集上的准确率呈现相似的变化趋势。当 D_2 为 4 时,模型解码性能达到最高,表明此时双分支卷积网络的空间特征提取能力处于最佳状态,进而有效促进了模型解码性能的提升。



(a) BCI竞赛IV-2a数据集上参数 D_2 对模型性能的影响
(a) Effect of parameter D_2 on model performance in the BCI competition IV-2a dataset



(b) BCI竞赛IV-2b数据集上参数 D_2 对模型性能的影响
(b) Effect of parameter D_2 on model performance in the BCI competition IV-2b dataset

图 7 输出通道数对模型的影响

Fig. 7 Effect of the number of output channels on the model

2)深度选取:时序多尺度注意力机制深度 L 是影响模型拟合能力的关键因素之一,并在很大程度上决定模型的参数规模。现针对不同深度对模型解码性能的影响进行分析,其中深度 L 的取值范围为 1~10 的整数。

如图 8 所示,DBTNet 在两个数据集上的分类准确率总体呈现出先上升后下降的变化趋势,均在深度为 6 时达到峰值。同时,从参数曲线可以看出,参数量随着深度的增加呈线性增长。然而,尽管模型的参数量不断增加,但分类准确率并未呈现出相应的提升,表明此时模型可能出现了过拟合现象,且该情况在两个数据集上均有体现。

3)头数选取:注意力头数 h 是时序多尺度注意力机制中的关键参数,直接影响模型的特征学习能力。现就不同注意力头数变化对模型解码性能的影响展开探讨。

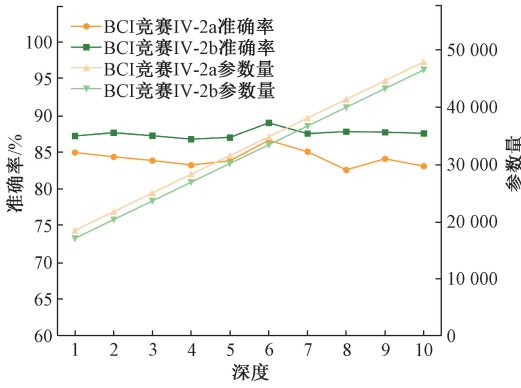


图 8 深度对模型的影响

Fig. 8 Effect of model depth on the model

图 9 展示了在两个公开数据集上,不同受试者在不同注意力头数设置下,模型分类准确率的变化情况。实验结果表明,模型在 BCI 竞赛 IV-2a 数据集上,部分受试者的准确率波动明显大于 IV-2b 数据集,表现出对头数变化更强的敏感性。当注意力头数 h 为 2 时,模型在两个数据集上的分类准确率达到最佳。这表明注意力头数的变化对模型解码性能具有显著影响。

3.4 特征分布可视化

为进一步探究 TMSA 模块与数据增强的作用,采用 t-分布随机邻域嵌入 (t-distributed stochastic neighbor embedding, t-SNE) 方法^[31]对不同学习到的特征进行可视化。图 10 展示了在 BCI 竞赛 IV-2a 数据集上受试者 S01 的 t-SNE 可视化结果。由图 10(b)和(c)可知,在未采用数

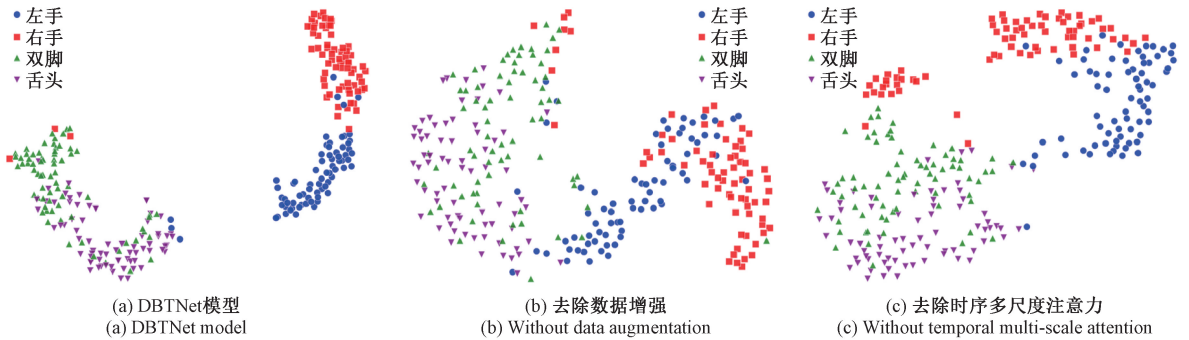


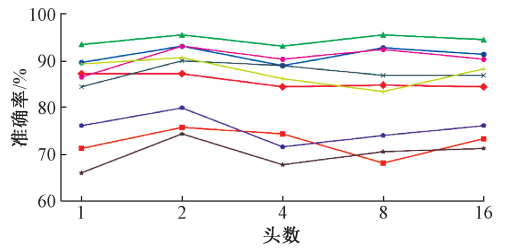
图 10 针对受试者 S01 的 t-SNE 可视化结果

Fig. 10 t-SNE visualization results for subject S01

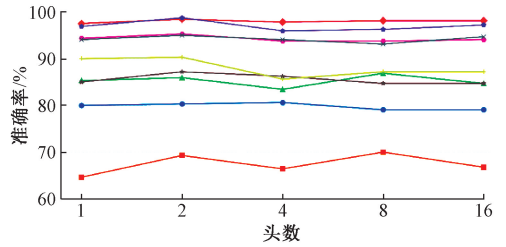
3.5 收敛分析

为了更直观地展示所提模型的训练过程,本研究分别在两个数据集上各选取一名受试者进行分析。图 11(a)与(b)分别展示了 DBTNet 在 BCI 竞赛 IV-2a 与 IV-2b 数据集中,受试者 S03 与 S04 对应的训练损失与训练准确率变化曲线。

由图 11 可知,随着训练轮次的增加,训练损失率逐渐下降,训练准确率逐渐提升,整体训练过程较为稳定。在 BCI 竞赛 IV-2a 数据集中,模型在约 400 个训练轮次后趋



(a) BCI竞赛IV-2a数据集上头数对模型性能的影响
(a) Effect of head number on model performance in the BCI competition IV-2a dataset



(b) BCI竞赛IV-2b数据集上头数对模型性能的影响
(b) Effect of head number on model performance in the BCI competition IV-2b dataset

— S01 — S03 — S05 — S07 — S09
— S02 — S04 — S06 — S08

图 9 头数对模型的影响

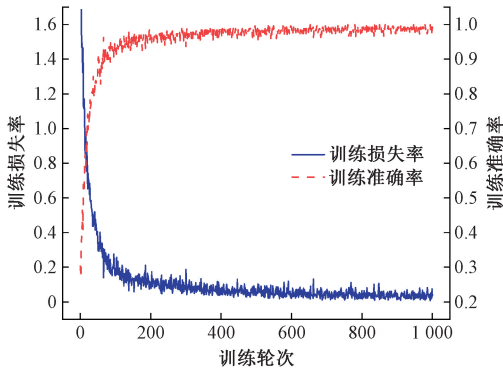
Fig. 9 Effect of the number of heads on the model

据增强和时序多尺度注意力时,特征在投影空间中分布较为分散;而图 10(a)中 DBTNet 所学习到的特征在每个类别中均呈现出更紧密的聚类状态。综上所述,时序多尺度注意力与数据增强均能提升所学习特征的判别性,进而有助于提高模型的解码性能。

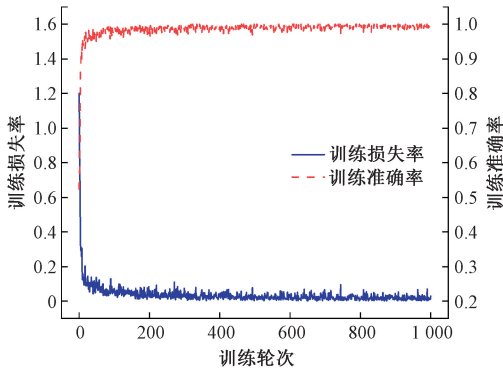
于收敛;而在 IV-2b 数据集中,模型在约 200 个训练轮次后训练损失率趋于稳定。实验结果表明,所提模型在两个数据集上均表现出良好的稳定性与收敛性。

3.6 消融实验

在 DBTNet 模型中,注意力双分支卷积网络模块 (ADBCN)和时序多尺度注意力机制 (TMSA)在提取 EEG 信号的关键信息与全局依赖关系方面发挥了重要作用。此外,数据增强 (Aug) 也对模型的解码性能起到了积极作用。本次消融实验在 BCI 竞赛 IV-2a 数据集上进行,主要



(a) 受试者S03在IV-2a数据集上的训练曲线
(a) Training curve of subject S03 on the IV-2a dataset



(b) 受试者S04在IV-2b数据集上的训练曲线
(b) Training curve of subject S04 on the IV-2b dataset

图 11 模型在两个数据集上特定受试者的训练曲线

Fig. 11 Model training curves for specific subjects on two datasets

评估了EMA模块、TMSA模块、分支2($D_2=4$)以及数据增强对模型解码性能的影响,结果如表3所示。

表 3 模块消融实验

Table 3 Module ablation experiment

去除的模块	Accuracy/%	Kappa
不去除	86.57	0.8210
EMA	84.88	0.7984
分支二	81.75	0.7567
EMA+分支二	83.49	0.7798
TMSA	83.41	0.7788
Aug	81.06	0.7474
TMSA+Aug	77.85	0.7047

实验结果表明,当去除分支2时,分类准确率下降了4.82%;去除EMA时,分类准确率下降了1.69%,进一步证明EMA与分支2模块在ADBCN中的关键作用。当同时去除EMA和分支2时,准确率下降了3.08%,这表明EMA与分支2的协同作用显著增强了模型的解码性能。此外,单独去除TMSA模块或数据增强时,模型解码性能分别下降了3.16%和5.51%;当二者同时去除时,分类准确率下降了8.72%,这表明TMSA模块与数据增强对提

升模型解码性能具有重要影响。

4 结 论

本研究提出了一种新型的DBTNet模型用于运动想象脑电信号解码领域。该模型通过注意力双分支卷积网络模块(ADBCN)多尺度提取MI-EEG信号的时空特征并捕捉关键信息,而时序多尺度注意力机制(TMSA)则用于捕捉局部特征和全局依赖关系。在BCI竞赛IV-2a与IV-2b数据集上的实验结果表明,DBTNet的分类准确率提升显著,且优于其他先进方法。在未来的工作中,将继续优化模型架构,探索新方法以进一步提升模型的解码性能。

参考文献

- [1] 廉小亲,刘春权,高超,等. 基于RSCM与黎曼空间的运动想象脑电分类研究[J]. 电子测量技术, 2025, 48(9):84-93.
LIAN X Q, LIU CH Q, GAO CH, et al. Research on motor imagery EEG classification based on RSCM and Riemann space [J]. Electronic Measurement Technology, 2025, 48(9):84-93.
- [2] 胡博,谢俊,刘俊杰,等. 用于运动想象脑电信号分类的深度学习网络[J]. 电子测量技术, 2025, 48(1):46-54.
HU B, XIE J, LIU J J, et al. Deep learning networks for motor imagery EEG signal classification [J]. Electronic Measurement Technology, 2025, 48(1):46-54.
- [3] ZHONG X C, WANG Q S, LIU D, et al. EEG-DG: A multi-source domain generalization framework for motor imagery EEG classification[J]. IEEE Journal of Biomedical and Health Informatics, 2024, 29(4):2484-2495.
- [4] MOAVENINEJAD S, D'ONOFRIO V, TECCHIO F, et al. Fractal dimension as a discriminative feature for high accuracy classification in motor imagery EEG-based brain-computer interface[J]. Computer Methods and Programs in Biomedicine, 2024, 244:107944.
- [5] 张家琦,漆石钰. 一种基于相位与残差信息的运动想象分类方法研究[J]. 电子测量与仪器学报, 2024, 38(12):155-162.
ZHANG J Q, QI SH Y. Research on a motion imagery classification method based on phase and residual information [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(12):155-162.
- [6] ZHOU K, HAIMUDULA A, TANG W Y. Dual-branch convolution network with efficient channel attention for EEG-based motor imagery classification [J]. IEEE Access, 2024, 12:74930-74943.
- [7] 莫云,李易,张本鑫,等. 基于权重融合特征重标定网络的运动想象脑电分类[J]. 电子测量与仪器学报, 2025, 39(1):70-79.
MO Y, LI Y, ZHANG B X, et al. Weight fusion-based feature recalibration network for motor imagery EEG classification[J]. Journal of Electronic Measurement and

- Instrumentation, 2025, 39(1):70-79.
- [8] LUO T J. Selective multi-view time-frequency decomposed spatial feature matrix for motor imagery EEG classification [J]. Expert Systems with Applications, 2024, 247:123239.
- [9] NICOLAS-ALONSO L F, GOMEZ-GIL J. Brain computer interfaces, a review [J]. Sensors, 2012, 12(2):1211-1279.
- [10] SHE Q SH, CHEN T, FANG F, et al. Improved domain adaptation network based on Wasserstein distance for motor imagery EEG classification [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2023, 31:1137-1148.
- [11] ZHAO W, ZHAO W B, WANG W F, et al. A novel deep neural network for robust detection of seizures using EEG signals [J]. Computational and Mathematical Methods in Medicine, 2020, 2020:9689821.
- [12] 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.
- [13] 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.
- [14] MANE R, CHEW E, CHUA K, et al. FBCNet: A multi-view convolutional neural network for brain-computer interface [J]. ArXiv preprint arXiv:2104.01233, 2021.
- [15] LI H L, CHEN H Y, JIA Z Y, et al. A parallel multi-scale time-frequency block convolutional neural network based on channel attention module for motor imagery classification [J]. Biomedical Signal Processing and Control, 2023, 79:104066.
- [16] LIU X G, XIONG SH CH, WANG X D, et al. A compact multi-branch 1D convolutional neural network for EEG-based motor imagery classification [J]. Biomedical Signal Processing and Control, 2023, 81:104456.
- [17] TANG X L, YANG C Q, SUN X, et al. Motor imagery EEG decoding based on multi-scale hybrid networks and feature enhancement [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2023, 31:1208-1218.
- [18] LIU K, YANG T, YU ZH L, et al. Msvtnet: Multi-scale vision transformer neural network for eeg-based motor imagery decoding [J]. IEEE Journal of Biomedical and Health Informatics, 2024, 28(12):7126-7137.
- [19] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. ArXiv preprint arXiv:1706.03762, 2017.
- [20] ALTAHERI H, MUHAMMAD G, ALSULAIMAN M. Physics-informed attention temporal convolutional network for EEG-based motor imagery classification [J]. IEEE Transactions on Industrial Informatics, 2023, 19(2):2249-2258.
- [21] ZHAO W, JIANG X L, ZHANG B C, et al. CTNet: A convolutional transformer network for EEG-based motor imagery classification [J]. Scientific Reports, 2024, 14(1):20237.
- [22] MA X ZH, CHEN W H, PEI ZH C, et al. Attention-based convolutional neural network with multi-modal temporal information fusion for motor imagery EEG decoding [J]. Computers in Biology and Medicine, 2024, 175:108504.
- [23] LIANG G J, CAO D G, WANG J Q, et al. EISATC-Fusion: Inception self-attention temporal convolutional network fusion for motor imagery EEG decoding [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2024, 32:1535-1545.
- [24] BRUNNER C, LEEB R, MÜLLER-PUTZ G, et al. BCI Competition 2008-Graz data set A [J]. IEEE Dataport, 2024, DOI:10.21227/katb-zv89.
- [25] LEEB R, BRUNNER C, MÜLLER-PUTZ G, et al. BCI Competition 2008-Graz data set B [J]. Graz University of Technology, Austria, 2008, 16:1-6.
- [26] SONG Y H, ZHENG Q Q, LIU B CH, et al. EEG conformer: Convolutional transformer for EEG decoding and visualization [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022, 31:710-719.
- [27] OUYANG D L, HE S, ZHANG G ZH, et al. Efficient multi-scale attention module with cross-spatial learning [C]. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2023:1-5.
- [28] ZHAO Q, ZHU W N. TMSA-Net: A novel attention mechanism for improved motor imagery EEG signal processing [J]. Biomedical Signal Processing and Control, 2025, 102:107189.
- [29] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016:770-778.
- [30] 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.
- [31] MAATEN L, HINTON G. Visualizing data using t-SNE [J]. Journal of Machine Learning Research, 2008, 9:2579-2605.

作者简介

李响, 硕士研究生, 主要研究方向为脑机接口、运动想象脑电信号处理。

E-mail: 1436996720@qq.com

艾尔肯·亥木都拉(通信作者), 硕士, 副教授, 硕士生导师, 主要研究方向为脑机接口、信号处理和高端农牧装备与特种机器人。

E-mail: arkin@xju.edu.cn

王路路, 硕士研究生, 主要研究方向为脑机接口、运动想象脑电信号处理。

E-mail: wanglululyl@163.com