

基于多尺度空洞卷积网络的运动想象脑电信号分类模型

张学军^{1,2}, 董宣¹

(1. 南京邮电大学电子与光学工程学院、柔性电子(未来技术)学院, 南京 210023; 2. 南京邮电大学射频集成与微组装技术国家地方联合工程实验室, 南京 210023)

摘要: 基于运动想象的脑机接口研究多采用单一尺度的特征提取方法, 依赖固定感受野的卷积或递归结构, 难以全面捕捉脑电信号的时序特征。针对以上问题, 本文提出了一种基于多尺度空洞卷积的运动想象脑电信号分类(Multi-scale dilated convolutional network, MSDCN)模型。该模型首先利用两层一维卷积提取时空特征, 然后通过多尺度空洞卷积增强对短时动态变化和长时依赖关系的建模能力, 并结合压缩与激励(Squeeze-and-excitation, SE)模块学习通道权重, 突出关键特征通道, 提高分类性能。在BCI Competition IV 2a数据集上的被试内实验准确率达到84.1%的准确率, 被试间实验准确率为69.1%; 在BCI Competition IV 2b数据集上的被试内实验准确率达89.8%。

关键词: 脑电信号; 运动想象; 多尺度; 空洞卷积

中图分类号: R318; TN911.7; TP18

文献标志码: A

引用格式: 张学军, 董宣. 基于多尺度空洞卷积网络的运动想象脑电信号分类模型[J]. 数据采集与处理, 2026, 41(4): 995-1009. ZHANG Xuejun, DONG Xuan. Motor imagery EEG signal classification based on multi-scale dilated convolutional network[J]. Journal of Data Acquisition and Processing, 2026, 41(4): 995-1009.

引言

作为一种连接人脑与外部设备的前沿技术, 脑机接口(Brain-computer interface, BCI)通过采集与解析脑电信号, 实现人脑与计算机或其他设备之间的直接交互^[1]。借助BCI技术, 用户脑电活动即可控制计算机、机器人、假肢和轮椅等外部设备, 为因伤病或先天性疾病导致运动障碍的患者提供新的交流与控制方式, 极大地改善了其生活质量。因此, BCI已成为当前人工智能与神经工程领域的重要研究方向之一。

在众多脑信号检测手段中, 脑电图(Electroencephalogram, EEG)因其低成本、高时间分辨率及便捷性, 成为BCI研究中最常用的信号来源^[2]。BCI技术中包含多种针对特定脑电信号的实验范式, 以满足不同的应用需求。这些实验范式主要包括运动想象(Motor imagery, MI)^[3]、P300^[4]、稳态视觉诱发电位(Steady-state visual evoked potential, SSVEP)^[5]以及情绪识别^[6]等。其中, 基于运动想象-脑电图(MI-EEG)的BCI因其良好的可控性与训练效果, 在机器人控制及辅助康复等领域广受关注^[7]。

在MI-EEG范式中, 当被试者想象左手或右手运动时, 大脑皮层会产生显著的事件相关去同步化(Event-related desynchronization, ERD)与事件相关同步化(Event-related synchronization, ERS)现象^[8]。这种大脑活动的变化为MI-BCI系统提供了关键的生理基础, 为MI-EEG的识别与分类奠定了重要的理论依据。

然而,受限于EEG信号的采集方式,实验过程中不可避免地会混入非目标信号的干扰,如肌电(Electromyogram, EMG)信号、环境电磁干扰及电极接触不良等噪声,导致信噪比较低,严重影响特征提取与分类效果^[9]。

传统方法多通过8~40 Hz的带通滤波提高信噪比^[10],或利用独立分量分析(Independent component analysis, ICA)^[11]、经验模态分解(Empirical mode decomposition, EMD)^[12]等技术去除伪迹;随后采用共同空间模式(Common spatial pattern, CSP)^[13]或滤波器组CSP(Filter bank CSP, FBCSP)^[14]提取特征,最后使用支持向量机(Support vector machine, SVM)^[15]或线性判别分析(Linear discriminant analysis, LDA)^[16]进行分类。尽管该类方法取得了一定效果,但其过于依赖人工先验知识,限制了模型的泛化能力和适应性。

近年来,深度学习技术得到了快速发展,其在自动特征提取和复杂模式学习方面展现出的强大优势有效弥补了传统特征提取的不足,显著提升了模型的特征表达能力和泛化性能。Lawhern等^[17]基于卷积神经网络(Convolutional neural network, CNN)提出的EEGNet采用深度分离卷积技术,高效提取EEG信号的时空特征,展现出优异的泛化能力与鲁棒性。Sakhavi等^[18]将传统的FBSCP算法与CNN网络相结合,利用FBSCP对EEG数据进行分频处理,然后通过CNN进一步提取特征,相较于传统的FBSCP与SVM方案,在同一数据集上提升了8%的准确率。为了有效捕捉时间序列中的关键特征,研究人员开始在其研究中引入长短期记忆(Long short-term memory, LSTM)网络^[19]。作为一种特殊形式的递归神经网络(Recursive neural network, RNN),LSTM网络通过其独特的记忆单元结构,能够有效地记录和利用前一层生成的历史信息,并将历史信息作为当前网络状态的重要参考,在处理时间序列数据时展现出显著优势。为了更好地获取时间序列中的特征信息,Bai等^[20]基于CNN提出了新的时间卷积网络(Temporal convolutional network, TCN),该网络以更简单的结构获得了更优于LSTM的性能。在该结构被提出不久后,Ingolfsson等^[21]就将TCN网络用于运动想象-脑电图(MI-EEG)信号的分类,提出了EEG-TCN,空洞卷积和因果卷积带来的更大的感受野使模型可以获得更多的时域信息。相比于传统的RNN和LSTM结构,该方法不仅在计算效率上占优,还在MI-EEG任务上展现出了更强的分类能力。

尽管已有研究在EEG信号分类中取得了一定进展,但大多数MI-EEG信号分类模型采用单一尺度的特征提取方式,依赖单一感受野的卷积或递归结构进行时序建模。然而,EEG信号的动态特性复杂,不同时间尺度上的特征可能蕴含不同的生理信息,单一尺度的建模方式难以全面捕捉EEG信号在不同时间维度上的关键特征。此外,传统的时序建模方法,如RNN和LSTM,在处理长时间序列时往往面临梯度消失和计算开销较大的问题。

针对以上问题,本文提出了一种基于多尺度空洞卷积与高效通道注意力的脑电信号分类(Multi-scale dilated convolutional network, MSDCN)模型,主要贡献如下:

- (1)设计多分支空洞卷积结构捕获多尺度时空特征,结合注意力机制实现特征自适应加权;
- (2)采用同类样本片段交换数据增强策略,缓解训练数据不足问题;
- (3)在2个数据集上将该模型与其他分类模型的分类结果进行比较,证明该方法在MI-EEG信号分类中的有效性。

1 方 法

图1所示为本文提出模型MSDCN的整体框架,主要包括时空特征提取模块、多尺度空洞卷积模块、通道注意力模块、多尺度特征融合模块以及分类器。

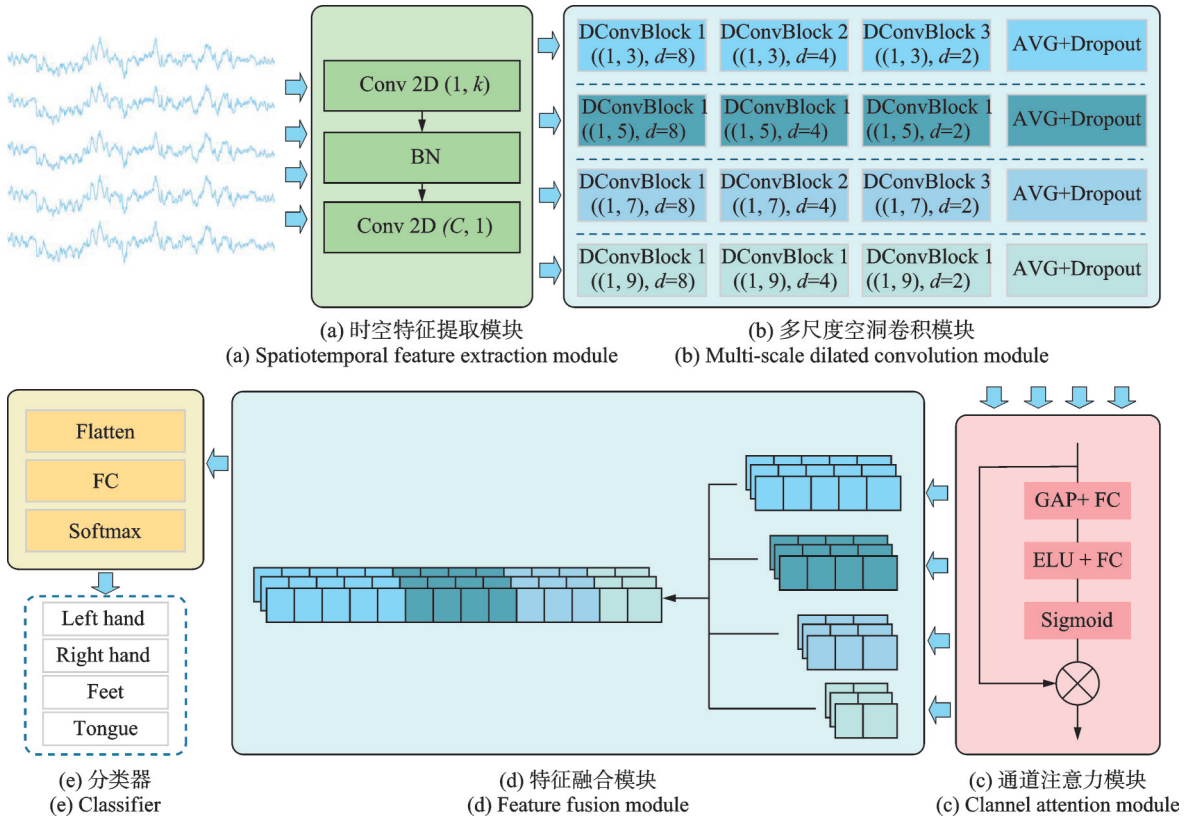


图1 MSDCN模型总体框架

Fig.1 Overall framework of MSDCN model

首先,原始EEG信号输入至时空特征提取模块,通过时间卷积与跨通道空间卷积联合建模,获得紧凑的时空特征表示;随后,多尺度空洞卷积模块采用不同尺度的空洞卷积并行结构,在保持时间分辨率的同时扩展感受野,以捕获EEG信号在不同时间尺度下的动态变化特征;在此基础上,引入通道注意力模块对不同特征通道进行自适应加权,突出与运动想象任务高度相关的判别性通道信息;接着,多分支特征在特征融合模块中进行拼接与整合,形成多尺度、多通道联合表征;最后,经全连接层与Softmax分类器完成运动想象任务的判别输出。

1.1 时空特征提取模块

受EEGNet的特征提取方式启发,本文设计了如图1(a)所示的时空特征提取模块,以获取原始脑电数据中的时空特征。

原始信号 $X \in \mathbb{R}^{B \times 1 \times \text{Ch} \times T}$ 通过形状为 $(1, k)$ 的一维卷积和在每个电极通道上独立进行卷积运算,以捕获EEG信号的时间特征。随后使用批量归一化(Batch normalization, BN)对时空特征进行标准化处理以加速收敛并稳定训练过程。最后通过形状为 $(\text{Ch}, 1)$ 的卷积核(其中Ch为电极数)提取空间特征,跨通道建模不同电极之间的空间相关性。

该模块实质上在空间维度上进行了降维处理,将原始特征维度由 $(B, 1, \text{Ch}, T)$ 压缩为 $(B, C, 1, T)$, 其中 C 为输出通道,时间维度 T 通过设置padding参数得以保持不变。该过程不仅有效融合了不同电极的空间信息,还降低了参数量和计算复杂度,为后续特征提取和分类提供了更加紧凑且具有判别性的时空特征表示。

1.2 多尺度空洞卷积编码模块

为了进一步获取时空特征中包含的信息,本文基于空洞卷积设计了如图1(b)所示的多尺度空洞卷积模块。其具体结构如图2所示。

该模块由4个结构相同的并行分支构成,每个分支包含3个相同卷积核的空洞卷积块和1个池化层。每个 DConvBlock 均包含1个卷积核大小为 $(1, k_n)$ 的二维空洞卷积(DConv 2D)、BN以及ELU激活函数,用于增强特征的非线性表达能力和稳定性。其中膨胀率(Dilation rate)设计为指数递减,分别为8、4、2,以此实现从全局到局部的特征提取策略,信息在不同感受野上逐层传递,避免由于较大空洞率导致的特征空洞现象,充分捕捉不同时间尺度下的特征信息。

第1个分支采用卷积核大小为 $(1, 3)$ 的空洞卷积,主要用于捕捉 EEG 信号中的短时局部特征。通过较小的卷积核配合不同的空洞率,能够在保证较小感受野的同时,扩展特征的时域感知范围,有效提取短时依赖。

第2个分支采用卷积核大小为 $(1, 5)$ 的空洞卷积,进一步扩大了卷积核的感受野。该分支在提取局部信息的基础上,能够捕捉到更多的中等时长特征,有助于补充第一分支在时序建模上的不足。

第3个分支采用卷积核大小为 $(1, 7)$ 的空洞卷积,主要用于提取 EEG 信号中的长时依赖特征。较大的卷积核和逐层递减的空洞设计,使得该分支在保持特征连续性的同时,更加关注长时间的变化规律和趋势信息。

第4个分支使用卷积核大小为 $(1, 9)$ 的空洞卷积,是感受野最大的一条分支。该分支重点捕捉 EEG 信号中的全局特征与长时间依赖特性,为模型提供更丰富的整体时序信息,进一步提升特征表达的全面性。

这种设计使模型能够同时获取短时动态变化和长时依赖信息,从而提高对 EEG 信号分类的时间特征建模能力。

多尺度的结构设计不仅丰富了特征的时域表达,还提升了模型对不同时间依赖特征的感知能力。通过4个不同卷积核大小的分支并行提取 EEG 信号的多尺度特征,能够有效缓解单一尺度特征提取带来的信息缺失问题。同时,各分支在不同尺度上进行特征提取,相互独立又相互补充,能够更全面地刻画 EEG 信号的多样性与复杂性。

1.3 通道注意力模块

为了评估不同输出通道在运动想象任务中的相对贡献程度,本文引用 SENet (Squeeze-and-excitation network)^[22]作为通道注意力模块,对各通道特征进行自适应加权。通过端到端训练,该模块会为不同通道分配权重,强化对分类关键的通道特征,弱化次要或干扰通道的影响,以此优化特征表示。SENet的结构如图3所示,分为3个主要步骤:压缩、激励和加权。

压缩是对输入信号 $u \in \mathbb{R}^{C \times H \times W}$ 的每个通道独立地应用全局平均池化(Global average pooling, GAP),将全局空间信息压缩到一个通道中,从而生成通道级别的特征描述,其计算公式为

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (1)$$

式中: z_c 表示第 c 个通道的全局特征, H 和 W 分别为特征图的高度和宽度。通过这种方式,SENet模块

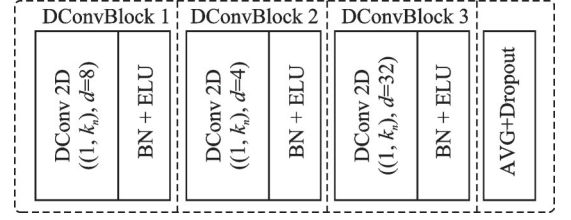


图2 空洞卷积模块结构图

Fig.2 Structure of the dilated convolutional module

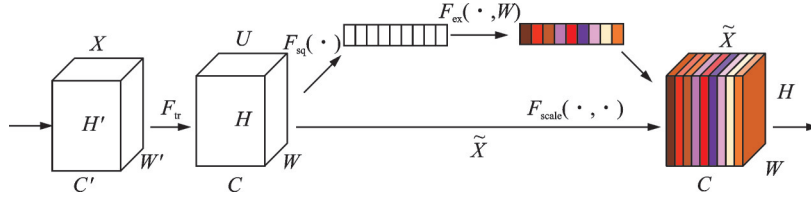


图3 SENet结构图

Fig.3 Structure diagram of SENet

能够捕获每个通道的整体响应,为后续的通道注意力分配提供全局信息支持。

激励机制的核心目标是充分捕获通道维度的依赖性。它具有足够的灵活性,能够学习通道之间的非线性关系,从而精确建模通道间的依赖性,并且具有非互斥性,可以确保多个重要通道能够被同时激活,其过程可表示为

$$s = F_{\text{ex}}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (2)$$

式中: $W_1 \in \mathbb{R}^{\frac{C}{r} \times \frac{C}{r}}$; $W_2 \in \mathbb{R}^{C \times \frac{C}{r}}$; δ 表示 ReLU 激活函数,在本文中将激活函数从原本的 ReLU 激活函数改为更适合处理脑电信号的 ELU 激活函数^[23]。

最终,将该权重对原始特征进行加权调整,从而增强关键通道的信息,同时抑制冗余或无关通道的信息。

多尺度时序建模在卷积网络中通常通过逐层扩展感受野的方式实现,即在网络深度方向上逐步引入更大的时间感受野,从而由局部时序模式过渡到长程依赖建模,且通常在多尺度特征融合阶段引入注意力机制。该类设计在时序卷积网络中被广泛采用,具有结构清晰的特点。

不同于上述逐层扩展感受野的建模范式,本文提出的 MSDCN 在同一特征提取阶段并行引入多种时间感受野,通过设置不同空洞率的卷积实现对长短期时序特征的同步建模。该设计在结构上体现为一种由大到小的感受野组织方式,即首先通过较大感受野捕获全局时序轮廓,再结合较小感受野补充局部细节信息,从而避免仅依赖深层堆叠才能获得全局依赖的问题。在此基础上,MSDCN 在每个尺度分支内独立引入通道注意力机制,使不同时间感受野下的通道特征能够根据其判别贡献进行自适应加权。多尺度特征在完成尺度内增强后再进行融合,实现了多尺度时序建模与通道建模的协同优化。

2 实验与结果分析

实验使用的硬件环境为 NVIDIA Geforce GTX 4060 GPU,软件环境采用基于 Python3.8 的 pytorch2.2.2 框架, CUDA 版本为 11.2。模型训练过程中使用 Adam 优化器和交叉熵损失函数,损失权重设置为 0.000 5,训练轮次为 2 000 个 epoch。

对比实验中,选取多种具有代表性的先进方法作为基线模型。对于对比方法,本文优先采用其作者公开代码或官方实现进行复现;若无代码则采用原文实验结果,以确保尽可能一致的实验设置。所有模型均在相同的硬件与软件环境下独立运行,并采用一致的评估指标(分类准确率)进行性能比较。

为了表现出 MSDCN 良好的分类能力,本文在四分类的 BCI Competition IV 2a 数据集^[24]和二分类的 BCI Competition IV 2b 数据集^[25]上进行分类实验,并与一些先进模型进行对比,其中训练集和测试集的划分均依照竞赛给定划分来进行。

2.1 数据集

BCI Competition IV 2a 数据集是一个用于 BCI 研究的公开数据集,包含 9 名受试者的 EEG 数据。每位受试者进行了 4 种想象任务:左手、右手、双脚和舌头。为了区分训练集和测试集,受试者被安排在

两个不同日期进行了两次数据采集。每次采集均为 288 次实验,其中四分类分布均匀即每种想象任务均有 72 次实验。如图 4 所示,每次实验持续时间约为 8 s。

第 0~2 s,为准备阶段,被试前的黑色屏幕中显示十字符号;第 2~3 s 屏幕上提示需要进行的想象任务,提示持续约 1.2 s,提示结束继续显示十字符号;第 3~6 s,被试根据屏幕提示进行想象,直到屏幕上的十字符号消失,然后进入休息阶段等待下一次实验开始。

该数据集共采用 25 个电极通道以 250 Hz 的频率采取数据,其中包括 22 个脑电通道和 3 个眼电通道,其中 3 个 EOG 通道的数据仅用来去除干扰不参与分类。

BCI Competition IV 2b 数据集是一组专门用于二分类任务(左手和右手)的 MI-EEG 数据集。该数据集包含 9 名被试,每名被试根据视觉提示想象左手或右手的运动。实验采用 3 个电极,分别为 C3、Cz、C4,并以 250 Hz 为采样频率,每次实验大约持续 8.5 s。在本文的实验中选取 2.5~7 s 的数据作为输入数据进行分类。

2.2 数据预处理与增强

预处理在深度学习中是不可缺少的步骤。受限于信号采集设备与技术,在实际采集信号的过程中,往往无法精准地获得需求的信号。例如 EEG 信号容易受到各种噪声和伪影的影响,例如眼电、肌电以及环境噪声等。这些干扰不仅会降低信号的质量,还可能导致分类模型的性能下降。因此,在深度学习建模之前,必须对 EEG 信号进行预处理,以提取有效信息并提高模型的分类能力^[26]。

常见的 EEG 预处理方法包括滤波、归一化和伪影去除等。在本文中,为去除不同通道和不同被试幅值范围不同的影响,采用标准化来提高模型的稳定性和泛化能力。标准化过程可表示为

$$X'_C = \frac{X_C - \mu_C}{\sigma_C} \quad (3)$$

式中: X_C 为某电极通道的输入信号, μ_C 为对应通道的平均值, σ_C 为对应通道的标准差。

除了基本的信号清理,数据增强也是提升 EEG 任务性能的重要策略。由于 EEG 数据量有限,直接训练深度学习模型可能会导致过拟合。为了解决这一问题,本文使用片段切割交换的方法,将类别相同的数据两两组合,并将数据平均切割成 5 份,交换第 2 和第 4 片段,从而创造出新的训练数据。在保持时间序列信息的同时,增加数据多样性,从而提高模型的泛化能力。

这些预处理与增强手段可以最大程度地保留 EEG 信号的有效特征,为深度学习模型的训练提供高质量的数据输入。

2.3 评价指标

为体现模型的性能,本文选取准确率和 κ 系数作为模型的评价指标。此外,为了全面评估模型的分类能力,本文还计算了混淆矩阵(Confusion matrix),用于分析不同类别的识别情况。

准确率是最常用的分类性能指标之一,定义为模型正确分类的样本数占总样本数的比例,其计算公式为

$$Acc = \frac{TP + TN}{TP + FN + FP + TN} \quad (4)$$

式中:TP(True positive)为被正确分类为正例的样本数;TN(True negative)为被正确分类为负例的样

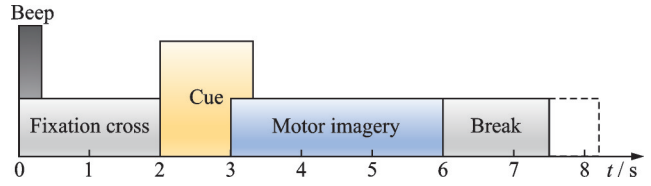


图4 范式的时序方案
Fig.4 Timing scheme of the paradigm

本数;FP(False positive)被错误分类为正例的负例样本数;FN(False negative)被错误分类为负例的正例样本数。

κ 系数是一种衡量分类模型一致性的指标,它考虑了模型预测的准确性,同时排除了随机猜测带来的影响,其表达式为

$$\kappa = \frac{P_o - P_e}{1 - P_e} \tag{5}$$

式中: P_o 为分类准确度, P_e 为随机猜测准确度。

2.4 对比试验

为了体现本文所提模型优异的分类能力,在两个数据集上与当前较为先进模型及部分经典模型进行了比较,结果如表1所示,其中,标注“*”的模型数据是本文基于原作者开源代码,在与本模型相同的软硬件环境下复现得到的结果;未作标注的模型数据则为直接引自原文的数据。所有复现实验均严格遵循原文的参数设置,同时在训练方法和数据预处理环节采用与本模型一致的实验流程,以确保对比实验的公平性。

表1 各模型在BCI-2a数据集上的被试内分类性能对比

Table 1 Comparison of within-subject classification performance of each model on the BCI-2a dataset

方法	准确率/%										κ 系数
	S01	S02	S03	S04	S05	S06	S07	S08	S09	AVG±Std	
EEGNet ^{[17]*}	84.0	66.0	90.6	66.3	78.5	64.9	91.7	85.1	84.4	79.1±10.7	0.72
EEGTCNet ^{[21]*}	82.3	62.8	93.1	75.7	78.8	64.6	90.3	85.8	86.8	80.0±10.7	0.73
DeepNet ^{[27]*}	84.7	67.0	88.2	83.3	67.4	62.2	93.4	77.1	84.4	78.6±10.8	0.72
MSVTNet ^{[28]*}	84.7	62.2	93.4	75.0	75.4	70.1	90.6	85.8	88.2	80.6±10.5	0.74
ACTNet ^{[29]*}	84.0	63.2	93.8	83.3	78.8	74.0	94.4	85.8	87.2	82.7±9.8	0.77
DMT-GAT ^[30]	88.3	66.7	96.7	73.3	81.7	66.7	91.7	86.7	88.3	82.2±10.9	—
MSRANet ^[31]	82.3	69.8	94.1	72.9	83.0	72.9	93.4	87.9	91.0	83.0±9.3	0.77
IMCTNet ^[10]	90.1	61.6	93.8	81.6	72.9	65.3	94.8	87.5	88.9	81.8±12.4	0.76
MSDCN	89.6	75.0	94.8	77.8	76.4	73.3	93.4	89.6	87.2	84.1±8.4	0.79

由表1可看出,本文提出的MSDCN在BCI-2a数据集的被试内实验上达到了较为优秀的分类准确率,在所有被试上的平均分类准确率达到了84.1%,高于所有对比模型。在数据质量较差的2号与6号被试上,准确率更是达到了75%和73.3%,说明本模型对比其他模型在劣质数据上的特征提取效果更好。与EEGNet相比,本模型的分类准确率提高了5%,相较于近年较为优秀的单尺度模型ACTNet,准确率提高了1.4%,对比同样为多尺度模型的MSVTNet和MSRANet,依旧有较为明显的提升。同时,MSDCN的 κ 系数亦优于其他对比模型,表明其分类一致性更强,泛化能力更优。

为了更直观地体现以上模型的分类效果,本文利用t-SNE工具包对特征分类进行了如图5所示的可视化处理。由图5可看出,所有被试的脑电信号被清晰地分为4个相对独立的簇,相同类别的数据明显地被划分在同一区域。这一现象在3号与7号被试上尤为显著。尽管2号和6号被试的分类准确率较低,但其数据分布仍显示出明显的类别边界,表明模型泛化性能良好。然而,部分相似类别之间仍存在一定程度的重叠,表明模型在区分相似信号方面仍有提升空间。总体而言,本文所提模型能够有效提取不同类别的特征并分类。

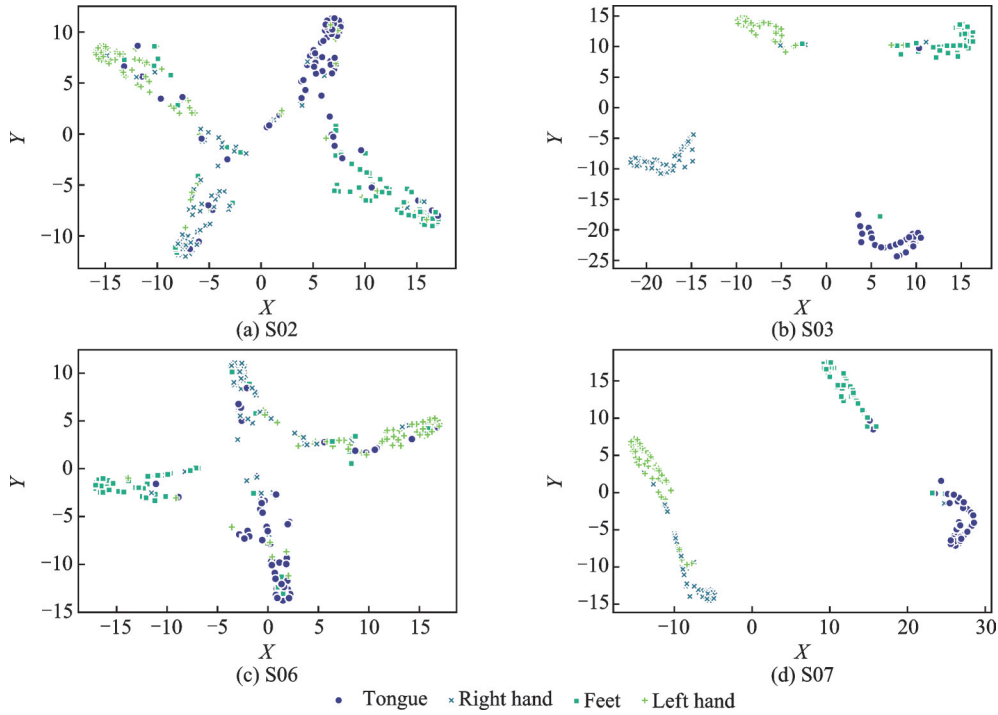


图5 部分被试在BCI-2a数据集上的特征可视化

Fig.5 Feature visualization of partial subjects on the BCI-2a dataset

混淆矩阵能够直观地反映模型在各类别间的分类性能,图6展示了3号与7号被试的混淆矩阵。矩阵的对角线元素反映了模型对各类别的正确分类情况。在BCI-2a数据集中,每个被试共进行288次实验(每类72次),因此理想情况下对角线数值越接近72,表明该类别的分类准确率越高。如图6所示,本模型在四类运动想象任务上的分类性能表现均衡,未出现对特定类别信号的偏好,这表明模型在不同MI任务上的特征提取较为稳定,能够较好地适应不同类别的数据,证实模型具备可靠的类间区分能力。

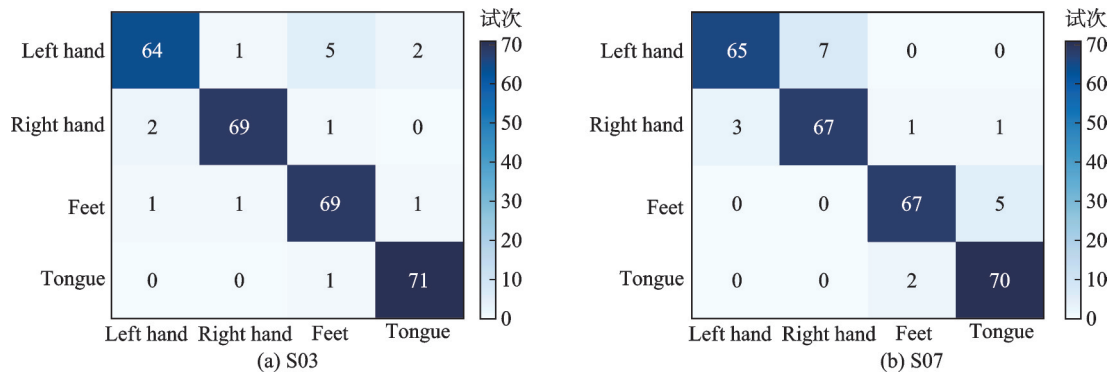


图6 混淆矩阵

Fig.6 Confusion matrix

为了进一步体现本模型优秀的泛化能力,本文在BCI Competition IV 2b数据集上进行了二分类的被试内实验,结果如表2所示。由表2可看出,本模型在所有对比模型中取得了最高的分类准确率,相较于EEGNet,提升了28.6%,并显著优于近年提出的ACTNet和IMCTNet等先进模型。0.8的 κ 系数

表明本模型在二分类任务中比其他对比模型具有更好的分类一致性,对不同分类任务具有较强的适应能力。

表 2 各模型在 BCI-2b 数据集上的被试内分类性能对比

Table 2 Comparison of within-subject classification performance of each model on the BCI-2b dataset

方法	(AVG±Std)/%	κ 系数	方法	(AVG±Std)/%	κ 系数
EEGNet*	61.2±2.1	0.22	ACTNet*	85.3±9.7	0.71
EEGTCNet*	85.7±9.8	0.71	IMCTNet	86.5±11.9	0.73
DeepNet*	84.8±10.3	0.70	MSDCN	89.8±7.4	0.80
MSVTNet*	80.0±12.7	0.60			

除了被试内实验,被试间实验也是衡量模型性能好坏的一项重要指标。为验证本模型在被试间实验上的性能,本文在 BCI-2a 数据集上采用了留一法(Leave-one-out, LOO)进行被试间训练,即将其中一名被试的数据作为测试集,而其余被试的所有数据作为训练集。这种实验设计能够有效评估模型在跨被试条件下的泛化能力,模拟真实应用场景中的适应性。结果如表 3 所示。

表 3 各模型在 BCI-2a 数据集上的被试间分类性能对比

Table 3 Comparison of cross-subject classification performance of each model on the BCI-2a dataset

方法	准确率/%										κ 系数
	S01	S02	S03	S04	S05	S06	S07	S08	S09	AVG±Std	
EEGNet*	71.4	58.5	80.4	60.1	62.2	59.7	73.3	76.2	73.8	68.4±8.3	0.58
EEGTCNet*	72.2	55.9	82.1	61.5	63.0	58.2	76.7	76.9	73.8	68.9±9.4	0.59
DeepNet*	71.5	54.7	80.0	56.8	56.8	57.6	77.4	73.8	64.6	65.9±9.9	0.55
MSVTNet*	71.7	53.0	76.6	58.3	59.4	47.4	76.0	74.0	67.5	64.9±10.7	0.53
ACTNet*	73.3	55.9	77.1	60.9	69.8	55.0	73.1	74.3	68.1	67.5±8.2	0.57
MSDCN	76.0	56.3	83.7	60.6	60.1	55.0	74.5	80.7	74.7	69.1±11.0	0.59

由表 3 可知,本文模型在四分类跨被试任务中的平均准确率达到 69.1%,其中 S03 最高,S02 最低。这一显著差异反映了典型的跨被试问题,即不同被试之间存在明显的个体差异(Inter-subject variability)。该差异主要体现在脑区激活位置、节律响应强度以及信号信噪比等方面,使得模型在不同被试上的特征分布存在偏移,从而增加分类难度。其中 S02 与 S06 的性能较低,可能由于其 EEG 信号中噪声干扰较强或运动想象诱发的脑电模式不稳定,导致判别特征不明显。从对比模型来看,MSVTNet 与 ACTNet 在被试内任务中表现优于 EEGNet 与 EEGTCNet,但在跨被试任务中性能下降更为明显,说明其更依赖于个体特有的特征模式,泛化能力相对有限。相比之下,MSDCN 在被试内与被试间任务中均保持较为稳定的性能,表明其具备更强的跨被试泛化能力。这一优势主要来源于模型结构设计:首先,多尺度空洞卷积通过引入不同时间感受野,使模型能够同时建模短时动态特征与长时依赖关系,从而学习到具有更强时间不变性的判别特征,减弱由个体差异引起的时序分布偏移;其次,通道注意力机制能够对不同电极通道的重要性进行自适应加权,使模型在不同被试间自动关注更具判别力的脑区,从而缓解空间分布差异带来的影响。此外,由图 7 可

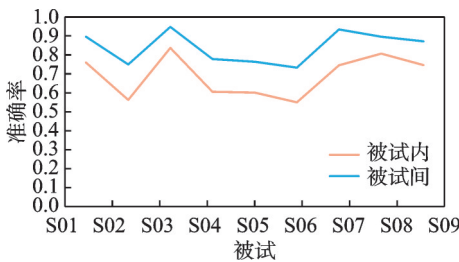


图 7 MSDCN 在 BCI-2a 数据集上不同实验方法对比图

Fig.7 Comparison of different experimental methods of MSDCN on BCI-2a dataset

观察到,尽管被试间实验整体性能较被试内下降约10%,但各被试的性能变化趋势基本一致,说明模型在不同被试间具有相对稳定的判别行为。这一现象进一步表明,MSDCN并非简单依赖个体特征进行分类,而是能够提取跨被试共享的判别性特征,从而提升模型在复杂EEG场景下的鲁棒性与泛化能力。

综合被试内和被试间实验的结果,MSDCN不仅在个体特征差异较小的被试内任务中展现了较强的特征提取能力,而且在存在显著个体差异的被试间任务中同样具备较好的泛化能力和鲁棒性。在面对低质量EEG数据时,MSDCN依然能够保持较高的分类准确率。

2.5 消融试验

为了选取合适的并行分支数,本文从单分支开始依次添加分支数,每次添加分支的卷积核大小递增。各个分支的卷积核大小依次为(1,3)、(1,5)、(1,7)、(1,9)、(1,11),并BCI-2a数据集上进行被试内实验,结果如图8所示。由图8可看出,前4次实验模型的分类准确率随着分支数量的提升而逐渐递增,说明随着更广感受野的加入,模型可以提取更加全面的特征;但是分支数增加到5时,准确率开始下降,说明过多的分支反而会使模型学到冗余或无效特征,导致模型的特征提取能力出现下降。从 κ 系数的变化趋势来看,模型的 κ 值同样随着分支数的增加在前4次实验中逐步提升,并在四分支结构下达到最高值0.7881,表明该结构在不同类别之间具有更高的一致性和更稳定的判别能力;而当分支数增加至五分支时, κ 系数下降至0.7737,进一步验证了过度扩展多尺度结构可能导致模型判别稳定性下降的问题。

上述实验主要验证了分支数量对模型性能的影响,但卷积核尺寸的具体组合方式同样对多尺度特征建模具有重要作用。为保证多尺度特征覆盖的连续性与稳定性,本文在卷积核设计中采用了奇数递增的尺寸组合(1,3/5/7/9),主要考虑如下。

(1) 中心对齐原则:奇数尺寸卷积核能够保证时间维度上的对称感受野,使当前时刻始终位于感受野中心,从而避免偶数卷积核可能引入的时序偏移问题。

(2) 尺度均匀覆盖:相邻卷积核尺寸间隔为2,使不同分支之间形成较为均匀的感受野跨度,避免尺度分布过于密集或稀疏,从而提升多尺度特征的覆盖完整性。

(3) 冗余抑制:若采用更大卷积核(如(1,11)及以上),其感受野与大膨胀率卷积存在重叠,容易引入冗余信息。

图8的实验结果也表明,当分支数扩展至包含更大卷积核时,模型性能反而下降,验证了过大尺度组合的有效性有限。

综合考虑模型的分类准确率和判别稳定性,最终本文选取了4个并行分支作为MSDCN的最佳结构设计,在保证模型具备较强特征提取能力的同时,避免了因分支过多带来的性能退化与潜在计算冗余。

为验证空洞卷积膨胀率设计的合理性,本文对不同膨胀率组合进行了对比实验。具体选取3种典型策略:递减结构(4,2,1)、递增结构(1,2,4)以及大尺度

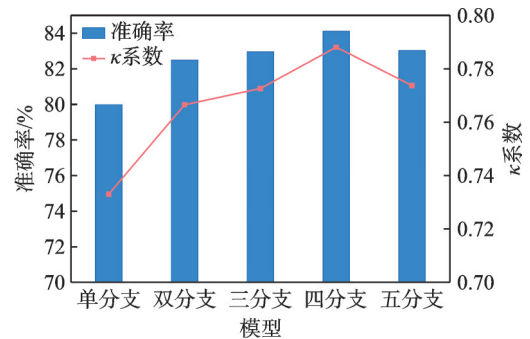


图8 不同分支模型准确率对比图

Fig.8 Accuracy comparison of different branch models

表4 不同空洞因子组合实验对比

Table 4 Comparison of different dilation rate combinations

组合	AVG/%	κ 系数
(1, 2, 4)	83.3	0.78
(4, 2, 1)	82.8	0.77
(2, 4, 8)	82.6	0.77
(8, 4, 2)	84.1	0.79

主导结构(2,4,8),并与本文采用的(8,4,2)进行比较,实验结果如表4所示。

从结果可以看出,本文所采用的(8,4,2)膨胀率组合在分类准确率与 κ 系数上均取得最优性能,平均准确率达到84.1%,优于其他3种对比方案。其中,(1,2,4)虽然在局部特征建模方面具有一定优势,但由于整体感受野较小,对长时依赖信息的捕获能力有限;而(4,2,1)采用由大到小的递减策略,虽然在一定程度上兼顾了多尺度信息,但其初始感受野不足,难以充分捕捉全局时序特征。(8,4,2)与(2,4,8)在理论感受野大小上相同,但其特征建模方式存在显著差异。(2,4,8)采用由小到大的膨胀策略,使得高层特征依赖大空洞率卷积进行建模,容易引入采样稀疏问题,从而影响特征连续性;而(8,4,2)通过先利用大感受野捕获全局信息,再逐步细化局部特征,实现由全局到局部的渐进式建模过程,有效缓解了空洞效应带来的信息损失问题。因此,本文最终选取(8,4,2)作为多尺度空洞卷积的膨胀率设置。

为了验证本实验各个模块的有效性,分别对数据增强、空洞卷积和SE模块在BCI-2a数据集上进行了消融实验,结果如图9所示,其中Model 1为未进行数据增强的模型;Model 2为不含SE模块的模型;Model 3为不使用空洞卷积的等效感受野模型。由图9可知,本文的所有工作均对模型分类能力的提升有效。经过数据增强后,模型的准确率提高了2%,同时 κ 系数亦呈现同步提升趋势。这一现象表明,数据增强不仅提高了整体分类精度,也增强了模型在不同类别间判别的一致性,能够有效缓解EEG训练数据不足的问题,使模型更充分地学习脑电信号的关键特征,从而提升分类性能。此外,数据增强的效果也验证了模型对多样性数据的适应能力,进一步增强了其泛化性。在引入SE通道注意力模块后,模型准确率提升约1.9%,对应的 κ 系数亦明显提高,表明通过对通道进行自适应加权,模型能够更加聚焦于与运动想象任务高度相关的关键通道,从而在保证整体精度提升的同时,提高不同类别预测结果的一致性。

此外,空洞卷积模块的引入在有效扩展感受野的同时,使模型能够更加充分地捕捉EEG信号中的长短期时序依赖关系。如表5所示,与采用普通卷积结构的Model 3相比,所提出的MSDCN在分类准确率提升约1.3%, κ 系数同步提高的情况下,模型参数量由 220.52×10^3 降低至 95.46×10^3 ,计算复杂度由 5.19×10^9 FLOPs降低至 2.77×10^9 FLOPs。上述结果表明,多尺度空洞卷积不仅提升了模型的判别性能,还在显著降低模型规模和计算开销的前提下,实现了更加高效的时序特征建模。

从整体结构和功能分工角度分析,本文引入的数据增强、多尺度空洞卷积和通道注意力模块分别作用于模型的不同阶段与不同建模维度,三者之间具有明确的层级划分和互补关系。数据增强主要在数据层面提升训练样本的多样性,缓解MI-EEG数据样本有限及被试差异带来的过拟合问题;多尺度空洞卷积作用于特征提取阶段,通过引入不同时间感受野增强模型对长短期时序依赖关系的建模能力;通道注意力模块则在各尺度分支内对通道特征进行自适应重加权,以突出对分类判别更为关键的通道信息。

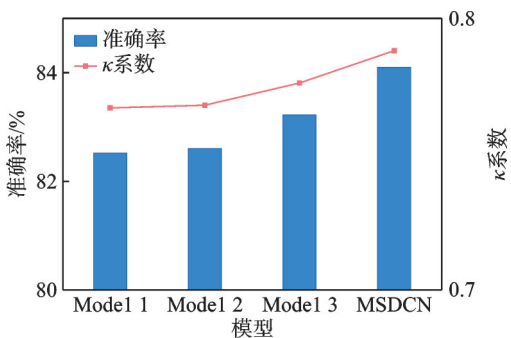


图9 消融实验对比

Fig.9 Comparison of ablation experiments

表5 等效感受野下的模型参数与计算量对比
Table 5 Comparison of model parameters and computation under equivalent receptive field

模型	Params/ 10^3	FLOP/ 10^9
Model 3	220.52	5.19
MSDCN	95.46	2.77

由于上述模块分别关注数据分布、多尺度时序建模以及空间通道判别性建模,其作用对象和优化目标不同,因此在组合使用时不会对同一信息进行重复建模。相反,多尺度空洞卷积所提取的判别性时序特征为通道注意力提供了更有效的特征基础,而通道注意力进一步强化了不同时间尺度下关键通道的响应,从而形成协同增强效应。该分析结果与消融实验中各模块逐步叠加后分类性能持续提升的实验现象保持一致。

3 结束语

本文提出了一种基于多尺度空洞卷积的运动想象脑电信号分类模型(MSDCN),通过多分支空洞卷积结构对不同时间尺度下的时序特征进行联合建模,并结合通道注意力机制实现对关键判别通道的自适应加权,从而提升模型对MI-EEG信号时序与空间信息的综合表征能力。实验结果表明,在BCI Competition IV 2a数据集的被试内实验中,所提模型取得了84.1%的平均分类准确率和0.788的 κ 系数;在被试间实验设置下,模型在不同被试间表现出较好的稳定性和泛化能力。在BCI Competition IV 2b数据集上,模型同样取得了较高的分类性能,验证了其在不同数据集和任务设置下的有效性。

消融实验进一步从分类准确率和 κ 系数两个维度验证了模型关键模块的作用。其中,多尺度空洞卷积结构在扩展感受野的同时,有效提升了模型对EEG信号长短期时序依赖关系的建模能力,在分类性能提升约1.3%的同时,又大大降低了计算参数,在性能与计算效率之间实现了良好的平衡。此外,通道注意力模块与数据增强策略在提升模型判别一致性和结果稳定性方面同样发挥了积极作用。

多尺度卷积方法在EEG研究中展现出了强大的特征提取能力,尤其适用于非平稳、低信噪比的MI信号。当前,跨受试者建模仍是当前EEG研究的瓶颈,未来将重点设计更精细的感受野调整策略,进一步优化对不同时间尺度特征的提取,提升分类精度,利用多尺度卷积提取多样的特征表示,从而提高模型在被试间实验的准确率。

参考文献:

- [1] 刘珈汐,高威.脑机接口研究综述[J].人工智能与机器人研究,2023,12(1):17-21.
LIU Jiaxi, GAO Wei. Review of brain-computer interface technology[J]. Artificial Intelligence and Robotics Research, 2023, 12(1): 17-21.
- [2] LI X, YANG Z, TU X, et al. MFRC-Net: Multi-scale feature residual convolutional neural network for motor imagery decoding[J]. IEEE Journal of Biomedical and Health Informatics, 2025, 29(1): 224-234.
- [3] MOUFASSIH M, TARAHI O, HAMOU S, et al. Boosting motor imagery brain-computer interface classification using multiband and hybrid feature extraction[J]. Multimedia Tools and Applications, 2024, 83: 49441-49472.
- [4] KIM J, CHO Y S, KIM S P. Task-relevant stimulus design improves P300-based brain-computer interfaces[J]. Journal of Neural Engineering, 2024, 21(6): 066046.
- [5] ZHANG X, WEI W, QIU S, et al. Enhancing SSVEP-based BCI performance via consensus information transfer among subjects[J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 36(8): 13833-13847.
- [6] 李存波,杨蕾,陈昭瑾,等.基于脑电网络图特征的情绪识别研究[J].数据采集与处理,2023,38(4):815-823.
LI Cunbo, YANG Lei, CHEN Zhaojin, et al. Emotion recognition based on graph features extracted from EEG networks[J]. Journal of Data Acquisition and Processing, 2023, 38(4): 815-823.
- [7] RAKSHIT A, KONAR A, NAGAR A K. A hybrid brain-computer interface for closed-loop position control of a robot arm[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 7(5): 1344-1360.
- [8] RAO Y, ZHANG L, JING R, et al. An optimized EEGNet decoder for decoding motor image of four class fingers flexion[J]. Brain Research, 2024, 1841: 149085.
- [9] 陈贺,张昊,柴一帆,等.低密度脑电自适应去噪方法[J].数据采集与处理,2023,38(4):824-836.

- CHEN He, ZHANG Hao, CHAI Yifan, et al. An adaptive denoising algorithm for few-channel EEG[J]. *Journal of Data Acquisition and Processing*, 2023, 38(4): 824-836.
- [10] ZHU L, WANG Y, HUANG A, et al. An improved multi-scale convolution and transformer network for EEG-based motor imagery decoding[J]. *International Journal of Machine Learning and Cybernetics*, 2025, 16(7): 4997-5012.
- [11] MARTINEZ-PEON D, GARCIA-HERNANDEZ N V, BENAVIDES-BRAVO F G, et al. Characterization and classification of kinesthetic motor imagery levels[J]. *Journal of Neural Engineering*, 2024, 21(4): 046024.
- [12] YANG Z, LUO S, ZHONG P, et al. An EMD and IMF energy entropy-based optimized feature extraction and classification scheme for single trial EEG signal[J]. *Journal of Mechanics in Medicine and Biology*, 2023, 23(8): 2340063.
- [13] MAHMOOD A, ZAINAB R, AHMAD R B, et al. Classification of multi-class motor imagery EEG using four band common spatial pattern[C]//*Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. [S.l.]: IEEE, 2017.
- [14] 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: 39.
- [15] MWATA-VELU T Y, NAVARRO RODRÍGUEZ A, MFUNI-TSHIMANGA Y, et al. EEG-BCI features discrimination between executed and imagined movements based on FastICA, Hjorth parameters, and SVM[J]. *Mathematics*, 2023, 11(21): 4409.
- [16] YANG B, LI D, MA B, et al. Motor imagery EEG classification method based on adaptive decision surface of LDA classifier [C]//*Proceedings of the 2021 11th International Conference on Bioscience, Biochemistry and Bioinformatics*. Tokyo, Japan: Association for Computing Machinery, 2021: 37-41.
- [17] 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.
- [18] SAKHAVI S, GUAN C, YAN S. 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.
- [19] LI H, DING M, ZHANG R, et al. Motor imagery EEG classification algorithm based on CNN-LSTM feature fusion network [J]. *Biomedical Signal Processing and Control*, 2022, 72: 103342.
- [20] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[EB/OL]. (2018-04-19). <https://doi.org/10.48550/arXiv.1803.01271>.
- [21] INGOLFSSON T M, HERSCHE M, WANG X, et al. EEG-TCNet: An accurate temporal convolutional network for embedded motor-imagery brain-machine interfaces[C]//*Proceedings of the 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. [S.l.]: IEEE, 2020.
- [22] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//*Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.]: IEEE, 2018.
- [23] GAO D, YANG W, LI P, et al. A multiscale feature fusion network based on attention mechanism for motor imagery EEG decoding[J]. *Applied Soft Computing*, 2024, 151: 111129.
- [24] BRUNNER C, LEEB R, MÜLLER-PUTZ G, et al. BCI competition 2008-Graz data set A[R]. Graz: Graz University of Technology, 2008.
- [25] LEEB R, BRUNNER C, MÜLLER-PUTZ G, et al. BCI Competition 2008—Graz data set B[R]. Graz: Graz University of Technology, 2008.
- [26] 董雷, 孙建起, 龚聪, 等. 基于运动想象的脑电信号处理研究综述[J]. *物联网技术*, 2024, 14(7): 104-106.
- [27] SCHIRMEISTER 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.
- [28] LIU K, YANG T, YU Z, 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.
- [29] 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.
- [30] SHI H, PEI Z, TANG Z, et al. A dense multi-scale temporal feature fusion and spatial information embedding model for electroencephalogram-based motor imagery classification[C]//*Proceedings of the 2024 IEEE 19th Conference on Industrial Electronics and Applications (ICIEA)*. [S.l.]: IEEE, 2024.
- [31] LI J, LI M, ZHANG S, et al. Decoding of motor imagery based on multi-scale residual network and multi-head attention block [C]//*Proceedings of the 2024 43rd Chinese Control Conference (CCC)*. Kunming, China: IEEE, 2024: 5417-5422.

作者简介:

张学军(1969-),男,博士,教授,硕士生导师,研究方向:智能信息处理、脑机接口技术、无线射频识别技术等, E-mail: xjzhang@njupt.edu.cn。



董宣(2001-),通信作者,男,硕士研究生,研究方向:智能信息处理、脑电信号处理技术, E-mail: 2373516450@qq.com。

(编辑:王静)

Motor Imagery EEG Signal Classification Based on Multi-Scale Dilated Convolutional Network

ZHANG Xuejun^{1,2}, DONG Xuan^{1*}

(1. School of Electronic and Optical Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210023, China;

2. Nation-Local Joint Project Engineering Lab of RF Integration & Micropackage, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: Most motor imagery-based brain-computer interface studies rely on single-scale feature extraction methods, which use convolutional or recurrent structures with fixed receptive fields and struggle to comprehensively capture the temporal characteristics of electroencephalogram (EEG) signals. Aiming at the above problem, we propose a multi-scale dilated convolutional network (MSDCN) model for motor imagery EEG signal classification. The proposed model first extracts spatiotemporal features through two layers of one-dimensional convolution, then enhances its ability to model both short-term dynamic variations and long-term dependencies using multi-scale dilated convolutions. Additionally, a squeeze-and-excitation (SE) module is integrated to learn channel-wise feature importance, highlighting key feature channels and improving classification performance. Experimental results show that on the BCI Competition IV 2a dataset, the model achieves an accuracy of 84.1% for within-subject experiments and 69.1% for cross-subject experiments, and on the BCI Competition IV 2b dataset, the within-subject accuracy reaches 89.8%.

Highlights:

1. A novel multi-scale dilated convolutional network (MSDCN) is proposed for motor imagery EEG classification, which adopts four parallel branches with different convolution kernel sizes (1×3 , 1×5 , 1×7 , 1×9) to synchronously capture short-term local temporal features and long-range global dependencies of EEG signals.
2. A descending dilation rate strategy ($d=8, 4, 2$) is designed inside each dilated convolution block, which models global temporal information first and then supplements local details, effectively alleviating the sparse feature loss problem caused by large dilation rates.
3. The squeeze-and-excitation (SE) channel attention module is embedded in each multi-scale branch to adaptively assign weights to feature channels, strengthening discriminative EEG channels related to motor imagery tasks and suppressing irrelevant interference features.

Key words: electroencephalogram (EEG) signal; motor imagery (MI); multi-scale; dilated convolution

Foundation item: Postgraduate Research & Practice Innovation Program of Jiangsu Province (No.KYCX24_1162).

Received: 2025-05-12; **Revised:** 2026-04-26

***Corresponding author, E-mail:** 2373516450@qq.com.