

微电子学与计算机  
*Microelectronics & Computer*  
ISSN 1000-7180, CN 61-1123/TN

## 《微电子学与计算机》网络首发论文

题目：基于 CNN 与 Transformer 的脑电解码研究  
作者：李响，艾尔肯·亥木都拉  
网络首发日期：2025-04-16  
引用格式：李响，艾尔肯·亥木都拉. 基于 CNN 与 Transformer 的脑电解码研究[J/OL]. 微电子学与计算机. <https://link.cnki.net/urlid/61.1123.tn.20250416.1456.006>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

# 基于 CNN 与 Transformer 的脑电解码研究

李响, 艾尔肯·亥木都拉

(新疆大学机械工程学院, 新疆乌鲁木齐 830017)

**摘要:** 针对现有深度学习模型在脑电信号解码能力的局限性, 本文提出了一种 RCA-Conformer 网络架构, 用于脑电图的运动想象分类任务。该架构通过多尺度时序卷积和空间卷积提取局部特征, 结合残差通道注意力进一步增强卷积特征表达, 并利用自注意力机制捕捉全局依赖性。此外, 通过引入残差连接, 将卷积模块与 Transformer 编码器模块提取的特征进行融合, 直接输入到分类器中进行分类, 从而显著提升模型的性能。在受试者依赖评估中, 该架构在 BCI 竞赛 IV-2a 数据集上达到了 80.29% 的分类准确率 ( $Kappa=0.7371$ ); 在 BCI 竞赛 IV-2b 数据集上达到了 85.74% 的分类准确率 ( $Kappa=0.7148$ )。通过与现有方法相比, RCA-Conformer 表现出了优异的性能, 验证了该方法的有效性, 为脑电图解码领域提供了新的解决方案。

**关键词:** 脑机接口; 脑电图; 多头注意力; 运动想象; 多尺度时序卷积

**中图分类号:** TP391 **文献标识码:** A

## Research on EEG decoding based on CNN and Transformer

LI Xiang, ARKIN Hamdulla

(School of Mechanical Engineering, Xinjiang University, Urumqi 830017, China)

**Abstract:** To address the limitations of existing deep learning models in EEG signal decoding capability, this paper proposes an RCA-Conformer network architecture for the motor imagery classification task of EEG. The architecture extracts local features through multi-scale temporal and spatial convolution, combines residual channel attention to further enhance the convolutional feature representation, and utilizes a self-attention mechanism to capture global dependencies. In addition, by introducing residual connectivity, the features extracted by the convolution module and the Transformer encoder module are fused and directly inputted into the classifier for classification, thus significantly improving the classification performance of the model. In the subject-dependent evaluation, the architecture achieved 80.29% classification accuracy ( $Kappa=0.7371$ ) on the BCI Competition IV-2a dataset; and 85.74% classification accuracy ( $Kappa=0.7148$ ) on the BCI Competition IV-2b dataset. By showing excellent performance compared with existing methods, RCA-Conformer validates the effectiveness of the method and provides a new solution for the field of EEG decoding.

**Keywords:** BCI; EEG; multi-head attention; motor imagery (MI); multi-scale temporal convolution

**基金项目:** 国家自然科学基金项目 (10971177)

## 1 引言

脑机接口 (Brain-Computer Interface, BCI) 技术通过解读电生理信号, 建立人类与外部设备之间的沟通桥梁, 从而实现对用户意图的识别<sup>[1]</sup>。脑电信号 (Electroencephalogram, EEG) 因其低成本、低风险、时间分辨率高及便携性等特点, 在医疗和非医疗领域中得到了广泛应用<sup>[2]</sup>。运动想象 (Motor Imagery, MI) 是指在没有实际执行身体动作的情况下, 仅通过大脑想象特定运动的过程<sup>[3]</sup>。基于运动想象的脑机接口 (MI-BCI) 系统已被证明可以作为瘫痪患者运动康复的辅助工具<sup>[4]</sup>。因为脑电信号易受噪声和其他生理信号干扰, 所以脑机接口技术面临的挑战是准确识别人类意图。因此, 构建能够有效提取复杂脑电信号特征的深度学习 MI-BCI 系统至关重要。

近年来, 深度学习的快速发展推动了运动想象脑电图 (MI-EEG) 领域的进步, 卷积神经网络 (Convolutional Neural Network, CNN) 通过卷积层提取局部特征, 已成为解码脑电信号的主流方法。EEGNet<sup>[5]</sup>在 CNN 基础上引入深度可分离卷积, 表现出了良好的泛化性能。Shallow ConvNet 和 DeepConvNet<sup>[6]</sup>通过时序卷积和时空卷积层从脑电信号中解码任务相关信息。ResNet<sup>[7]</sup>通过引入残差连接, 有效避免了梯度消失和爆炸问题, 从而显著提升了模型的性能。FBCSP<sup>[8]</sup>通过引入滤波器组对多个频带划分后进行特征提取, 从而提高脑电信号的识别能力。FBCNet<sup>[9]</sup>从多个频段提取特征并设计方差层提取时间特征, 从而降低特征维度。

注意力因其在捕获长期依赖性的优越能力被广泛应用于自然语言处理<sup>[10]</sup>和图像处理<sup>[11]</sup>。因此, 许多研究者尝试将其引入到深度学习模型之中, 以便探索其在 MI-EEG 解码中的潜力。Transformer<sup>[10]</sup>架构最初为解决自然语言处理 (natural language processing, NLP) 领域中

的机器翻译问题而提出, 由于其在处理序列数据时具有的出色性能, 因此非常适合用于脑电信号处理。一些模型利用 Transformer 的自注意力机制来获取全局信息, 从而提升性能。然而, 此类模型往往忽视了学习局部特征的重要性, 但是局部特征对脑电信号的解码至关重要。

基于注意力的时序卷积网络 ATCNet<sup>[12]</sup>通过将多头注意力机制 (Multi-Head Attention, MHA) 与时空卷积结合, 提取最重要的特征。Conformer<sup>[13]</sup>将 Shallow ConvNet 与自注意力机制结合, 其中卷积模块负责提取局部特征, 而自注意力机制则捕捉局部时间特征中的全局依赖性。CTNet<sup>[14]</sup>利用 CNN 提取局部特征并通过 Transformer 编码器学习全局特征。TransNet<sup>[15]</sup>将 CNN 与自注意力机制相结合, 以融合多模态时间信息与全局依赖性。SE 注意力模块<sup>[16]</sup>通过建立通道间的相互关系, 自适应地调整通道特征响应。而 ECA 注意力模块<sup>[17]</sup>采用局部跨通道交互策略, 相比 SE 模块降低了模型复杂度与参数量。

为进一步提升运动想象脑电信号的分类性能, 本文提出了一种新型的 RCA-Conformer 架构, 该架构结合 CNN 和 Transformer, 通过多尺度时序卷积 (MSTCN) 和残差通道注意力 (RCA) 提取局部特征, 结合多头注意力机制强化全局特征。实验结果表明, 该方法在 BCI 竞赛 IV-2a 和 BCI 竞赛 IV-2b 数据集上表现出了卓越的性能。本文主要贡献如下:

- 1) 提出了一种残差通道注意力 (Residual Channel Attention, RCA) 模块, 能够自适应捕捉运动想象脑电信号中的跨通道交互, 以提升模型的空间 (通道) 特征提取能力, 有效突出关键特征并抑制冗余信息, 增强处理脑电信号等复杂数据的能力。

- 2) 采用多尺度时序卷积 (Multi-Scale Temporal Convolution, MSTCN), 通过四个不同大小的卷积核提取多尺度时间信息,

相较于单一时序卷积增强了时间局部特征的提取能力。

3) 利用残差连接使 CNN 和 Transformer 提取的特征能够直接传递给分类器,从而有效缓解梯度消失问题,增强模型的表达能力和分类精度。

## 2 研究方法

### 2.1 网络模型

本文提出了 RCA-Conformer 模型,如图 1 所示,该模型由卷积模块、Transformer 编码器模块和分类器组成。卷积模块结合了

多尺度时序卷积(MSTCN)、空间卷积和残差通道注意力(RCA)模块,分别提取多尺度时间特征、空间特征并增强空间信息选择性。卷积特征图经过重排列,输入至 Transformer 编码器模块中。Transformer 编码器模块通过多头注意力机制(MHA)捕捉全局上下文信息,并利用前馈网络(Feed-Forward Network,FFN)进一步处理特征。最后,分类器通过全局平均池化层和全连接层对特征向量进行分类并输出最终的预测结果。

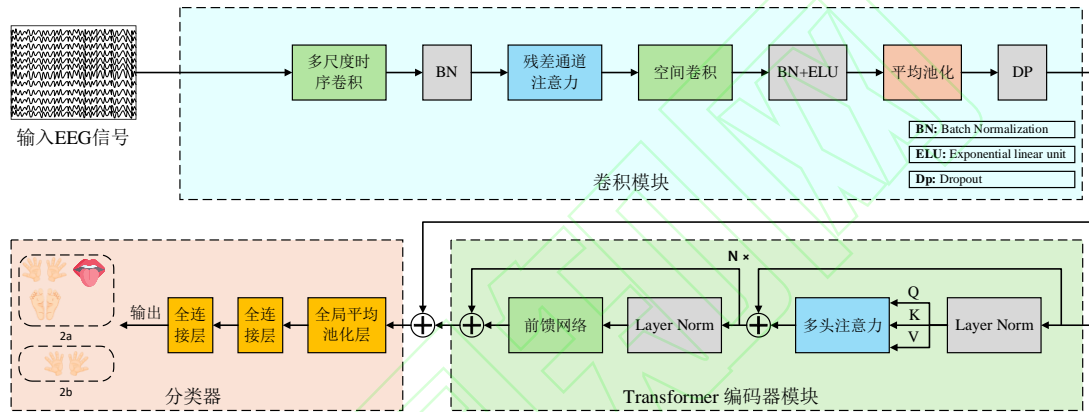


图 1 RCA-Conformer 模型结构

Fig. 1 RCA-Conformer model structure

### 2.2 卷积模块

卷积模块主要包括多尺度时序卷积(MSTCN)、残差通道注意力(RCA)和空间卷积,具体结构如表 1 所示。多尺度时序卷积用于在不同时间尺度上提取局部时间特征,以捕捉信号的时序变化;残差通道

注意力模块通过增强对 EEG 信号空间信息的选择性学习能力,建模通道间的交互关系,从而突出关键通道特征并抑制冗余信息;空间卷积则学习 EEG 信号的空间特征。

表 1 卷积模块的结构

Table1 Structure of the convolution module

Layer	In	Out	Kernel	Stride	Padding
Multi-Scale Temporal Conv	1	k/4	(1,15)	/	(0,7)
		k/4	(1,25)	/	(0,12)
		k/4	(1,51)	/	(0,25)
		k/4	(1,65)	/	(0,32)
Spatial Conv	k	k	(C,1)	(1,1)	/
Avg Pooling	k	k	(1,75)	(1,15)	/
Rearrange			$(k, 1, m) \rightarrow (m, k)$		

### 2.2.1 多尺度时序卷积

受 TransNet<sup>[15]</sup>启发, 引入多尺度时序卷积 (Multi-Scale Temporal Convolution, MSTCN), 具体结构如图 2 所示。首先, 将原始脑电图数据沿时间维度展开为一维, 并输入到四个不同卷积核大小的二维时序卷积层中, 如表 1 所示, 其卷积核大小分别为 (1, 15)、(1, 25)、(1, 51) 和 (1, 65), 填充大小为 (0, 7)、(0, 12)、(0, 25) 和 (0, 32)。每个卷积层的输出通道数为  $k/4$ , 并且时间维度保持不变。其次, 将四个卷积层的输出沿通道维度拼接, 得到通道数为  $k$ 。最后, 经过批量归一化层 (Batch Normalization, BN) 进行归一化, 以调整输出特征尺度并稳定训练过程。

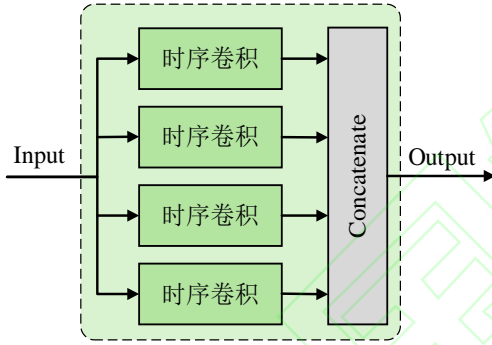


图 2 多尺度时序卷积

Fig. 2 Multi-scale temporal convolution

### 2.2.2 残差通道注意力

在 Conformer<sup>[13]</sup>和 TransNet<sup>[15]</sup>模型中, 卷积模块由时序卷积和空间卷积组成, 旨在有效提取脑电信号中的时间特征和空间特征。然而, 由于时序卷积主要侧重于时间特征的提取, 会导致其在提取空间信息时存在一定局限性, 可能引起空间信息的丢失。为解决该问题, 本文设计了残差通道注意力 (Residual Channel Attention, RCA) 模块, 提升模型空间特征提取的能力。

RCA 模块通过全局池化和一维卷积以捕获通道关系, 并引入残差连接强化模型稳定性, 避免过度依赖非线性层, 以提

升特征表达的多样性和鲁棒性, 从而加速收敛。具体过程如下。

1)如图 3 所示, RCA 模块的输入特征图维度为  $(N, H, W, C)$ , 其中  $C$  表示通道数。通过全局平均池化 (GAP) 操作将所输入的特征图维度变为  $(N, 1, 1, C)$ 。

2)如式 (1) 所示,  $C1D_k$  表示卷积核大小为  $k$  的一维卷积层, 以捕获局部跨通道交互的覆盖范围。在本文中, 采用 Sigmoid 函数将每个通道的权重限制在 0 到 1 之间, 从而实现特征的动态调整。

$$\omega = \sigma(C1D_k(y)) \quad (1)$$

3)卷积核大小  $k$  如式 (2) 所示,  $k$  通过  $C$  的映射自适应地确定, 且  $k$  与通道数  $C$  成比例。  $|t|_{odd}$  表示取最近的奇数, 并将  $y$  与  $b$  分别设置为 2 和 1。

$$k = \psi(C) = \left\lfloor \frac{\log_2 C}{y} + \frac{b}{y} \right\rfloor_{odd} \quad (2)$$

4) 残差连接将输入与输出相加, 从而加速模型的收敛速度, 此方法可在保留原始特征信息的同时增强特征表达的多样性。

### 2.2.3 空间卷积

空间卷积采用  $k$  个卷积核, 大小为  $(C, 1)$ , 步长为  $(1, 1)$ , 其中  $C$  表示 EEG 数据的电极通道数, 该操作旨在学习不同电极通道之间的空间特征。利用批量归一化 BN 并使用 ELU (Exponential Linear Units, ELU) 作为激活函数引入非线性变换, 如式 (3) 所示。最后, 采用卷积核大小为  $(1, 75)$ 、步长为  $(1, 15)$  的平均池化层以防止过拟合。其中, 将超参数  $k$  设置为 40, 并通过 dropout 层 ( $p=0.5$ ) 正则化以增强模型的泛化能力。

$$ELU = \begin{cases} \alpha(e^x - 1), & x \leq 0 \\ x, & x > 0 \end{cases} \quad (3)$$



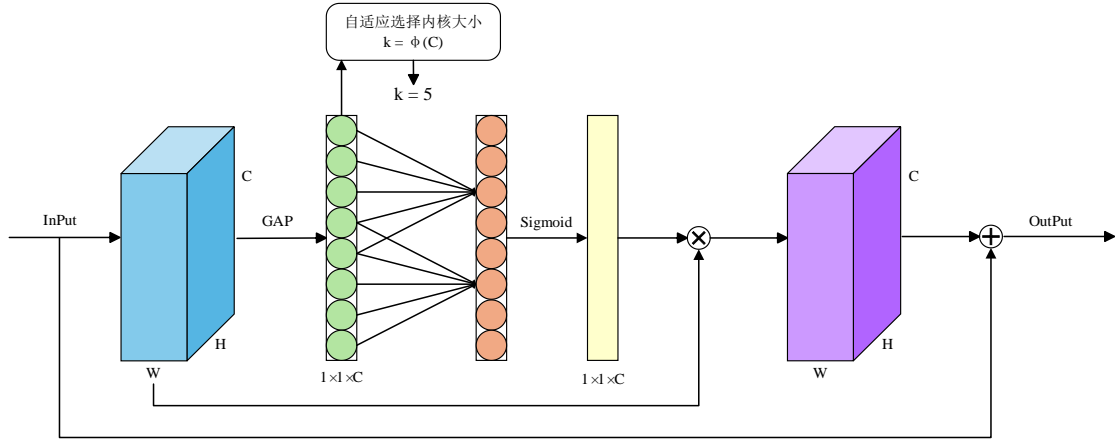


图 3 RCA 注意力机制

Fig. 3 RCAattention mechanism

### 2.3 Transformer 编码器模块

如图 1 所示，该模块主要包含多头注意力（Multi-Head Attention, MHA）与前馈网络（Feed-Forward Network, FFN）。采用自注意力机制学习脑电图特征的全局依赖关系，从而弥补卷积模块在感受野方面的局限性。LayerNorm 层与残差连接的结合可以提高模型的训练效率和鲁棒性。

卷积模块的输出经过线性变换转换为形状相同的查询(Q)、键(K)和值(V)。计算 Q 与所有 K 的点积得到注意力权重，并将每个 K 除以  $\sqrt{d_k}$  进行缩放，将结果通过 Softmax 函数得到权重矩阵，即注意力分数。然后，注意力分数与 V 进行点积操作进行加权。该过程可用公式（4）表示：

$$Attention(Q, K, V) = Soft\ max\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

其中  $d_k$  表示 K 的维度。

通过多头注意力机制进一步提高表示多样性。在 MHA 中，多个注意力头并行运行，每个头执行缩放点积注意力操作，注意力头数  $h$  包含查询 (Q)、键 (K) 和值 (V)，并将结果拼接起来传入线性层作为模块输出。其中，K 和 Q 的交互生成注意力分数，以突出 V 中重要的部分，如图 4 所示。该过程可用公式（5）和（6）表示：

$$MHA = (Q, K, V) Concat(H_1, H_2, \dots, H_h) W^O \quad (5)$$

$$H_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \quad (6)$$

其中  $H_i$  表示第  $i$  个 head 的输出， $W_i^Q \in \mathbb{R}^{d_m \times d_k}$ ,  $W_i^K \in \mathbb{R}^{d_m \times d_k}$ ,  $W_i^V \in \mathbb{R}^{d_m \times d_k}$ ，并且  $W^O \in \mathbb{R}^{h \cdot d_k \times d_m}$  表示权重矩阵。

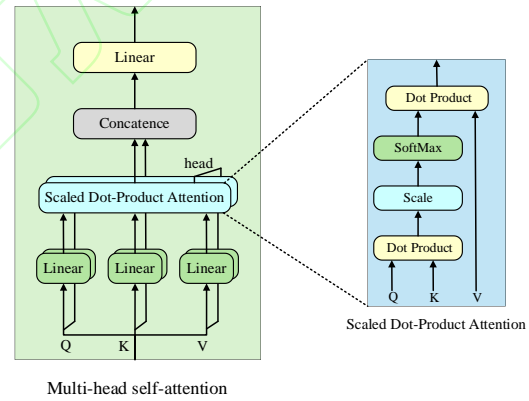


图 4 多头注意力

Fig. 4 Multi-head attention

### 2.4 分类器

将卷积模块和 Transformer 编码器模块提取的特征进行融合，使 CNN 提取的特征能够直接输入至分类器。分类器由一个全局平均池化层和两个全连接层构成，全局平均池化层将得到的特征向量序列累加求平均，然后输入至两个全连接层。利用 dropout (p=0.5) 层以减少过拟合，增强模型的泛化能力。最后，通过 Softmax

函数输出一个  $N$  维的向量。使用交叉熵作为整个模型的损失函数，该过程可用公式 (7) 表示：

$$L = -\frac{1}{M} \sum_{i=1}^M \sum_{c=1}^N y \log(\hat{y}) \quad (7)$$

其中  $M$  表示 EEG 的试验次数， $N$  表示 EEG 类别的数量， $y$  和  $\hat{y}$  分别表示真实标签和预测标签。

### 3 数据集与实验设置

#### 3.1 数据集介绍

本文在 BCI 竞赛 IV-2a 数据集<sup>[18]</sup>和 BCI 竞赛 IV-2b 数据集<sup>[19]</sup>上评估所提出的模型。

BCI 竞赛 IV-2a 数据集包含 9 名受试者 (A01-A09) 的脑电图数据，实验范式流程如图 5(a)所示。数据集记录了左手、右手、双脚和舌头的运动想象任务。使用 22 个 Ag/AgCl 电极并且以 250Hz 的采样率收集不同日期的两个 session。每个 session 包含 288 次脑电图试验，即每项任务 72 次试验。在每次试验中使用 [2, 6] 秒，并使用带通滤波器将 EEG 数据过滤为 [4, 40] Hz。第一次 session 用于训练，第二次 session 用于测试。

BCI 竞赛 IV-2b 数据集由 9 名受试者的脑电图数据组成，实验范式流程如图 5(b)所示。有两项运动想象任务，涵盖移动左手和右手的想象。使用三个双极电极 (C3、Cz 和 C4) 以 250 Hz 的采样率收集 5 个 session，每个 session 包含 120 次试验。在实验中使用了每次试验的 [3, 7] 秒，并在 [4, 40] Hz 之间进行带通滤波，以减少高频和低频噪声。前三次 session 用于训练集，后两次 session 用于测试集。

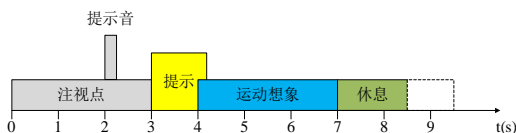
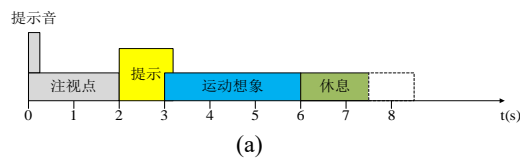


图 5 BCI 竞赛 IV-2a(a)和 2b 数据集(b)的实验范式流程

Fig.5 Experimental Paradigm Flow for BCI Competition IV-2a (a) and 2b datasets (b)

#### 3.2 数据集预处理

原始 EEG 试验的大小为  $C \times S$ ，其中  $C$  表示电极通道， $S$  表示时间样本。采用带通滤波来滤除无关的高频和低频噪声。采用 6 阶 Chebyshev 滤波器并进行标准化处理，以减少信号的波动性和非平稳性，从而增强数据的可比性，如公式 (8) 所示：

$$x_o = \frac{x_i - \mu}{\sqrt{\sigma^2}}, i = 1, 2, \dots, C \quad (8)$$

其中  $x_i$  和  $x_o$  分别表示带通滤波数据和标准化输出。 $\mu$  和  $\sigma^2$  表示均值和方差，由训练数据计算所得，并直接应用于测试数据。

#### 3.3 评估方法与评价指标

##### 3.3.1 数据增强

由于 EEG 采集耗时且数据集较小，容易导致模型过拟合。为缓解这一问题，一些方法通过数据增强扩充训练样本。然而，传统的高斯噪声或裁剪可能会破坏信号的原始特性。因此，本文采用时间域分割和重建(S&R)方法<sup>[20]</sup>，将同一类别的训练样本等分为  $n$  个片段，并在保持原始时间顺序的基础上，对这些片段进行随机拼接。该方法在每次迭代中生成与批次相等规模的增强数据，从而提升模型的泛化能力。

##### 3.3.2 评估方法

在 BCI 竞赛 IV-2a 数据集上进行实验，采用受试者依赖的评估方法，训练过程中使用第一次 session 的  $288 \times 9$  个运动想象 (MI) 实验数据，而测试则使用第二次 session 的数据。在 BCI 竞赛 IV-2b 数据集上，采用相同的评估方法，使用前第三次 session 的  $120 \times 9$  个 MI 实验数据作为训练集，后两次 session 作为测试集。

本文在两个公开数据集上均进行 10 次随机实验后取平均作为模型的实验结果, 以确保结果的准确性。

### 3.3.3 评价指标

本文使用分类准确率 (Accuracy) 和 Kappa 系数对所提模型性能进行评价, 具体过程如下:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

其中 TP 和 TN 分别表示由模型预测的正确阳性样本数和正确阴性样本数, FP 和 FN 分别表示由模型预测的假阳性样本数和假阴性样本数。

$$kappa = \frac{p_o - p_e}{1 - p_e} \quad (10)$$

其中  $p_o$  表示所有试验的精确度, 而  $p_e$  表示随机猜测的精确度。

### 3.3.4 实验环境与参数设置

本文的实验环境是 Window11 23 H2 操作系统, CPU 为 Intel i7-10700k 处理器, GPU 为 NVIDIA GeForce RTX 2080super 显卡, 显存为 8GB, 使用 Python 3.10, PyTorch 1.12.0。采用 Adam 优化器训练模型, 学习率设置为 0.0002,  $\beta_1$  和  $\beta_2$  分别为 0.5 和 0.999。将自注意力机制的执行次数 N 设置为 6, 头数 h 设置为 10, S&R 中的 n 设置为 8。在 BCI 竞赛 IV-2a 数据集上, epoch 为 2000, 批量大小为 72, 在 BCI 竞赛 IV-2b 数据集上, epoch 为 2000, 批量大小为 100。

## 4 实验结果与分析

### 4.1 多尺度时序卷积验证实验

本节旨在验证 RCA-Conformer 模型中所采用的多尺度时序卷积的有效性, 进行了如下实验。实验 1: 采用卷积核大小为 (1, 15), 填充大小为 (0, 7); 实验 2: 采用卷积核大小为 (1, 25), 填充大小为 (0, 12); 实验 3: 采用卷积核大小为 (1, 51),

填充大小为 (0, 25); 实验 4: 采用卷积核大小为 (1, 65), 填充大小为 (0, 32); 实验 5: 采用多尺度时序卷积。由表 2 可知, 实验 4 的分类准确率最低, 为 71.88%, Kappa 为 0.6250。而实验 5 的分类准确率最高, 达到了 80.29%, Kappa 达到了 0.7371。这表明, 相较于传统时序卷积, 多尺度时序卷积通过并行提取不同时间尺度的时间特征, 有效增强了时间特征的表达能力。此外, 由于脑电信号是时间序列数据, 单一的时序卷积难以有效捕获不同时间尺度上的变化, 而多尺度时序卷积能够适应脑电信号中的时间变化特征。因此, 证明了模型采用多尺度时序卷积在时间特征提取方面的优越性, 从而显著提高了模型的分类性能。

表 2 多尺度时序卷积验证实验

Table 2 Multi-scale temporal convolution validation experiments

Number	Accuracy/%	Kappa
1	73.38	0.6451
2	74.15	0.6554
3	72.88	0.6384
4	71.88	0.6250
5	80.29	0.7371

### 4.2 分类结果对比与分析

本节使用 BCI 竞赛 IV-2a 数据集与 BCI 竞赛 IV - 2b 数据集评估 RCA-Conformer 模型的性能, 选用了一些经典与先进的方法进行比较, 包括 Conformer<sup>[13]</sup>、EEGNet<sup>[5]</sup>、ConvNet<sup>[6]</sup>。表 3 对比了本文方法与其他方法在 BCI 竞赛 IV-2a 数据集上的分类性能, 其对比结果与参数均来源于相关文献。结果表明, RCA-Conformer 模型表现出了优秀的性能, 准确率和 Kappa 均优于其他模型, 其准确率达到 80.29%, Kappa 达到了 0.7371。其中, 所提出的模型仅在受试者 S05、S07 和 S08 的分类准确率未达到最高。本文的方法相比基于 CNN 的 EEGNet 和 ConvNet 方法分别提高了 5.79% 和 7.76%。尽管



CNN 具有强大的特征表示能力,但是基于 CNN 的方法只关注局部特征而忽略了全局相关性,以至于会影响脑电信号的解码性能。Conformer 是将 CNN 与 Transformer 相结合来封装局部和全局依赖性,所以其表现优于 EEGNet 和 ConvNet,从而表明与 Transformer 相结合弥补了感知域的局

限性。本文的方法相较于 Conformer 进一步提升了 CNN 的空间特征提取能力,并通过残差连接以丰富所提取的特征信息。因此,其性能优于 Conformer,从而进一步验证了该模型在分类稳定性和一致性方面的优越性。

表 3 BCI 竞赛 IV-2a 数据集与其他方法的性能对比

Table3 Performance of the BCI Competition IV-2a dataset compared to other methods

Methods	S01	S02	S03	S04	S05	S06	S07	S08	S09	Average	Kappa
ConvNet <sup>[6]</sup>	76.39	55.21	89.24	74.65	<b>56.94</b>	54.17	<b>92.71</b>	77.08	76.39	72.53	0.6337
EEGNet <sup>[5]</sup>	85.76	61.46	88.54	67.01	55.90	52.08	89.58	83.33	86.81	74.50	0.6600
Conformer <sup>[13]</sup>	88.19	61.46	93.40	78.13	52.08	65.28	92.36	<b>88.19</b>	88.89	78.66	0.7155
<b>Proposed</b>	<b>90.28</b>	<b>63.19</b>	<b>95.14</b>	<b>83.68</b>	55.21	<b>65.63</b>	91.32	87.85	<b>90.28</b>	<b>80.29</b>	<b>0.7371</b>

为了验证 RCA-Conformer 模型的解码能力和泛化性,本文在 BCI 竞赛 IV-2b 数据集进行二分类任务实验,以测试该模型的泛化性能。表 4 中展示了该模型与其他方法的性能比较。结果表明,二分类结果与 BCI 竞赛 IV-2a 数据集表现出相似的趋势。RCA-Conformer 显著提升了整体性能,准确率与 Kappa 分别达到了 85.74% 和 0.7148。相比于 EEGNet 和 ConvNet 方法分别提高了 5.26%和 6.37%。同时,本文方法的准确率和 kappa 均优于 Conformer,其中,在受试者 1 上的准确率提高了 2.08%。进一步验证了该模型的有效性与泛化性。

混淆矩阵是一种用于评估分类模型性能的重要评估方法。图 6(a)和(b)分别展示了本文方法与 Conformer 模型在 BCI 竞赛 IV-2a 数据集与 BCI 竞赛 IV-2b 数据集上的混淆矩阵。该混淆矩阵中的横轴为预测标签,纵轴为真实标签,右侧颜色条通过颜色深浅反映分类准确率的大小。矩阵主对角线上的数据表示各类别的正确分类准确率,而副对角线则表示某个类别被错误分类到其他类别的情况。结果表明,该模型在两个公开数据集上的分类性能相较于 Conformer 均有所提升。

表 4BCI 竞赛 IV-2b 数据集与其他方法的性能对比

Table4 Performance of the BCI Competition IV-2b dataset compared to other methods

Methods	S01	S02	S03	S04	S05	S06	S07	S08	S09	Average	Kappa
ConvNet <sup>[6]</sup>	76.56	50.00	51.56	96.88	93.13	85.31	83.75	91.56	85.62	79.37	0.5874
EEGNet <sup>[5]</sup>	75.94	57.64	58.43	98.13	81.25	88.75	84.06	93.44	89.69	80.48	0.6096
Conformer <sup>[13]</sup>	82.50	<b>65.71</b>	63.75	98.44	<b>86.56</b>	90.31	87.81	94.38	92.19	84.63	0.6926
<b>Proposed</b>	<b>84.58</b>	64.58	<b>66.25</b>	<b>99.17</b>	85.83	<b>90.83</b>	<b>90.00</b>	<b>96.25</b>	<b>94.17</b>	<b>85.74</b>	<b>0.7148</b>

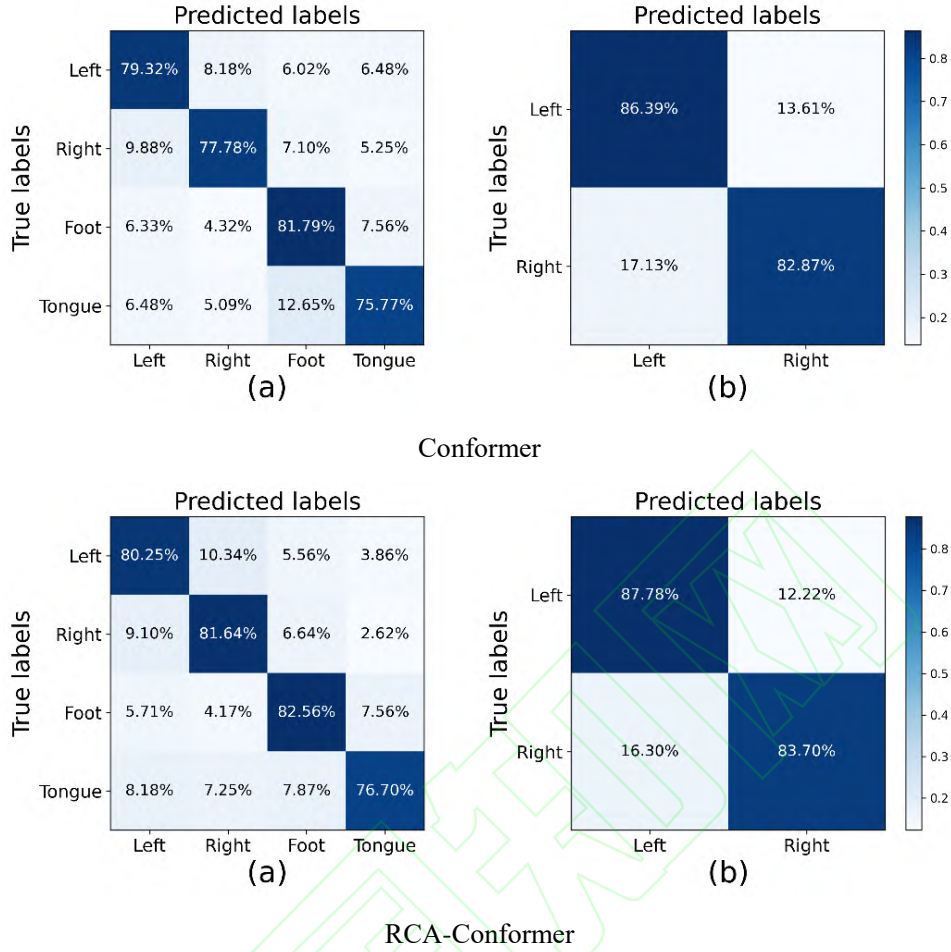
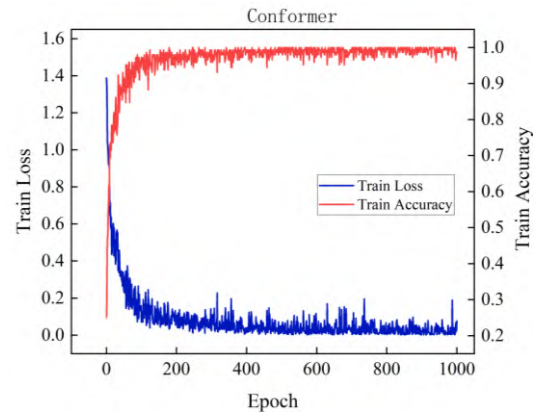


图 6 本文方法与 Conformer 模型在 BCI 竞赛 IV-2a(a)和 2b(b)数据集上的混淆矩阵

Fig. 6 Confusion matrix between the method in this paper and the Conformer model on the BCI Competition IV-2a(a) and 2b(b) datasets

为了分析本文方法与现有方法在特定受试者中的运动想象脑电信号分类的性能，进行了相关实验。图 7 展示了本文方法与 Conformer 模型在 BCI 竞赛 IV-2a 数据集上的受试者 S03 的训练损失率和训练准确率曲线。由图 7 可知，随着 epoch 的逐渐递增，Conformer 模型的训练损失率在 250 个 epoch 附近开始快速收敛，而 RCA-Conformer 的训练损失率则是在 200 个 epoch 附近显著降低，并且训练准确率迅速增加，即模型在此时快速收敛，证明了本文方法的有效性。同时，当 RCA-Conformer 模型训练至 400 个 epoch 附近时，训练损失与训练准确率逐渐趋于稳定，表现出了较高的鲁棒性。



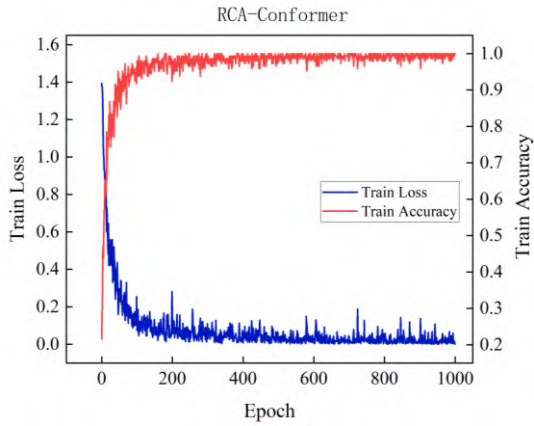


图7 本文方法与 Conformer 模型在 BCI 竞赛 IV-2a 数据集中受试者 S03 的训练曲线

Fig. 7 Training curves of this paper's method with the Conformer model for subject S03 in the BCI Race IV-2a dataset

### 4.3 消融实验

针对本文提出的 RCA-Conformer 模型在 BCI 竞赛 IV-2a 数据集上进行了消融实验,主要包括一下内容:去除 transformer 模块、去除数据增强模块、同时去除 transformer 模块和数据增强模块、去除 RCA 模块。如图 8 所示,去除 RCA 模块后,模型的平均分类准确率下降了 0.66%,受试者 5 的分类准确率下降幅度最大,达到了 2.78%。这表明 RCA 模块可以显著提升模型分类准确率。RCA 模块通过通道注意力机制动态调整各通道的权重,帮助模型更好地捕获空间特征,弥补了多尺度时序卷积在空间信息提取上的局限性。因此,当去除 RCA 模块后,模型的性能出现下降,进一步验证了其在空间(通道)特征提取中具有积极作用。在同时去除 transformer 模块和数据增强模块时,模型的整体分类性能下降幅度最大,达到了 14.43%。其中,受试者 4 的分类准确率下降幅度最为显著,为 26.74%,受试者 3 的分类准确率下降幅度最小,达到了 6.25%。这表明 transformer 模块和数据增强模块对提高模型分类性能具有重要作用。去除 transformer 模块时,受试者 4 的分类准确率下降幅度显著,为 11.11%,而受试者 8

的分类准确率下降幅度最小,为 3.47%。同时,模型的性能下降了 7.30%。进一步验证了与传统 CNN 方法相比,transformer 模块增强了模型对全局特征的学习能力。此外,当去除数据增强模块时,模型的平均分类准确率下降了 7.41%。其中,受试者 4 的分类准确率下降幅度最大,达到了 10.07%,受试者 3 的分类准确率下降幅度最小,为 4.51%。充分表明了数据增强可以显著提高模型的性能,具有积极影响。

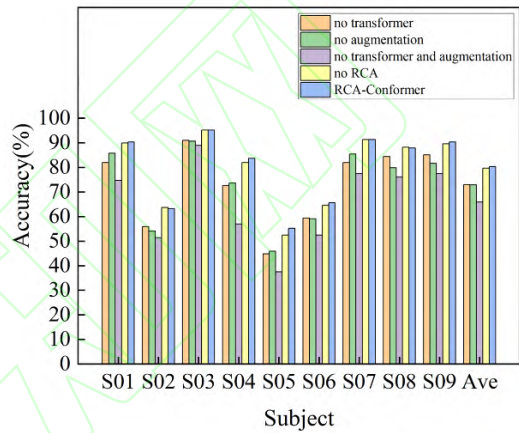


图8 模块消融实验

Fig. 8 Module ablation experiment

### 5 结束语

针对运动想象脑电信号解码能力受限问题,提出了一种将 CNN 和 Transformer 相结合的 RCA-Conformer 的网络模型。通过多尺度时序卷积(MSTCN)提取局部时间信息,残差通道注意力(RCA)捕获数据中的通道间交互,增强空间(通道)特征的学习能力。空间卷积学习电极通道之间的空间特征,自注意力机制学习脑电图特征的全局依赖性,以此弥补卷积模块中有限的感受野。最后,将卷积模块和 Transformer 编码器模块的输出特征直接传输至分类器,并采用全局平均池化层和全连接层,最终经过 Softmax 进行分类。具体结论如下:

1) 在 BCI 竞赛 IV-2a 数据集上的消融实验证明 RCA-Conformer 模型中添加

的模块有效提升了模型的整体性能。与其他方法相比,本文方法的受试者依赖准确率为80.29%, Kappa为0.7371, 优于其他方法, 证明该模型能够有效分类运动想象脑电信号信号, 不同受试者和不同类别中提取运动想象特征的能力。

2) 为了验证 RCA-Conformer 模型的泛化性能, 在 BCI 竞赛 IV-2b 数据集上进行了二分类实验。结果表明, 该模型的分分类准确率为85.74%, Kappa为0.7148, 优于该数据集上的其他方法, 证明了其在运动想象脑电信号解码和分类方面的优越性能。在未来的工作中, 将进一步探索和优化模型框架, 以提升网络模型的分分类性能和泛化能力。

### 参考文献:

- [1] Wang X, Liesaputra V, Liu Z, et al. An in-depth survey on Deep Learning-based Motor Imagery Electroencephalogram (EEG) classification[J]. Artificial intelligence in medicine, 2024, 147: 102738.
- [2] Shin J, Chung W. Multi-band CNN with band-dependent kernels and amalgamated cross entropy loss for motor imagery classification[J]. IEEE journal of biomedical and health informatics, 2023, 27(9): 4466-4477.
- [3] 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.
- [4] Ma Z Z, Wu J J, Cao Z, et al. Motor imagery-based brain-computer interface rehabilitation programs enhance upper extremity performance and cortical activation in stroke patients[J]. Journal of neuroengineering and rehabilitation, 2024, 21(1): 91.
- [5] 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.
- [6] Schirrneister 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.
- [7] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [8] 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.
- [9] 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.
- [10] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
- [11] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arxiv preprint arxiv:2010.11929, 2020.
- [12] Altafheri 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.
- [13] Song Y, Zheng Q, Liu B, et al. EEG conformer: Convolutional transformer for EEG decoding and visualization[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2023, 31: 710-719.
- [14] Zhao W, Jiang X, Zhang B, et al. CTNet: a convolutional transformer network for EEG-based motor imagery classification[J].

Scientific Reports, 2024, 14(1): 20237.

- [15] Ma X, Chen W, Pei Z, 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.
- [16] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
- [17] Wang Q, Wu B, Zhu P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2020: 11534-11542.
- [18] Brunner C, Leeb R, Müller-Putz G, et al. BCI Competition 2008–Graz data set A[J]. Institute for knowledge discovery (laboratory of brain-computer interfaces), Graz University of Technology, 2008, 16: 1-6.
- [19] 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.
- [20] Lotte F. Signal processing approaches to minimize or suppress calibration time in oscillatory activity-based brain-computer interfaces[J]. Proceedings of the IEEE, 2015, 103(6): 871-890.

#### 作者简介:

李响 硕士研究生, 1436996720@qq.com

艾尔肯·亥木都拉(通讯作者) 硕士, 硕士生导师, 副教授, arkin@xju.edu.cn