

融合 Transformer 网络与卷积神经网络的 稳态运动视觉诱发电位解码方法

张焕卿, 谢俊, 杨翰林, 杜方钊, 金志伟, 陈昱洁

(西安交通大学机械工程学院, 710049, 西安)

摘要: 针对卷积神经网络(CNN)在感受野有限、缺乏对全局信息的有效感知,以及在处理短时稳态运动视觉诱发电位(SSMVEP)信号时分类效果欠佳的问题,提出了一种紧凑 EEGNet-Transformer(即 EEGNetformer)网络。EEGNetformer 网络融合了为脑电(EEG)信号识别任务而设计的通用的卷积神经网络 EEGNet 网络和 Transformer 网络的优势,有效地捕捉与处理脑电信号中的局部和全局信息,增强网络对 SSMVEP 特征的学习,进而实现良好的解码性能。EEGNet 网络用于提取 SSMVEP 的局部时间和空间特征,而 Transformer 网络用于捕捉脑电时间序列的全局信息。在基于 SSMVEP-BCI 范式采集的数据基础上,开展了实验以评估 EEGNetformer 网络的性能。实验结果显示,当在 2 s SSMVEP 数据条件下,EEGNetformer 网络在基于被试者内情况的平均准确率为 $88.9\% \pm 6.6\%$,在基于跨被试者情况的平均准确率为 $69.1\% \pm 4.3\%$ 。与传统的 CNN 算法相比,EEGNetformer 网络的分类性能提升了 $4.2\% \sim 17.4\%$ 。研究内容说明,EEGNetformer 网络在有效提高 SSMVEP-BCI 识别准确率方面具有显著优势,为进一步提升 SSMVEP-BCI 解码性能提供了新的研究思路。

关键词: 稳态运动视觉诱发电位;EEGNet 网络;Transformer 网络;EEGNetformer 网络

中图分类号: R318 **文献标志码:** A

DOI: 10.7652/xjtuxb202505005 **文章编号:** 0253-987X(2025)05-0044-10

Steady-State Motion Visual Evoked Potential Decoding Method Based on Integration of Transformer Network and Convolutional Neural Network

ZHANG Huanqing, XIE Jun, YANG Hanlin, DU Fangzhao, JIN Zhiwei, CHEN Yujie

(School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Addressing the issues of limited receptive fields and ineffective perception of global information in convolutional neural networks (CNNs), as well as their suboptimal classification performance when processing short-duration steady-state motion visual evoked potentials (SSMVEP) signals, a compact EEGNet-Transformer (i. e., EEGNetformer) network is proposed. EEGNetformer network integrates the strengths of the EEGNet network (a general-purpose CNN designed for EEG recognition tasks) and the Transformer network. It efficiently captures and processes both local and global information in the electroencephalogram (EEG) signals, enhancing the network's learning of the SSMVEP features, thereby achieving an excellent decoding performance. The EEGNet network is utilized to extract local temporal and spatial features of SSMVEP, while the Transformer

收稿日期: 2024-08-25。 作者简介: 张焕卿(1997—),男,博士生;谢俊(通信作者),男,副教授,博士生导师。 基金

项目: 国家自然科学基金资助项目(52365040)。

网络出版时间: 2025-01-13

网络出版地址: <https://link.cnki.net/urlid/61.1069.T.20250110.1520.009>

network is utilized to capture global information in EEG time series. Based on data collected using the SSMVEP-BCI paradigm, experiments were conducted to evaluate the performance of the EEGNetformer network. The results showed that under the condition of 2 s SSMVEP data, EEGNetformer achieved an average accuracy of $88.9\% \pm 6.6\%$ for within-subject conditions and $69.1\% \pm 4.3\%$ for cross-subject conditions. Compared with traditional CNN algorithms, the classification performance of the EEGNetformer network was improved by 4.2%—17.4%. These findings demonstrate the significant advantage of EEGNetformer network in enhancing the recognition accuracy of SSMVEP-BCI systems, providing a new research direction for further enhancing SSMVEP-BCI decoding performance.

Keywords: steady-state motion visual evoked potential (SSMVEP); EEGNet network; Transformer network; EEGNetformer network

脑机接口(BCI)是一种通过采集和解码大脑信号来实现大脑与外部设备进行直接交互的技术^[1-2]。近年来,随着神经科学和计算机技术的快速发展,BCI在医疗康复、神经科学和军事等领域展现出广泛的应用前景^[3]。其中,稳态视觉诱发电位(SSVEP)作为一种常见的BCI诱发电位信号,因其高信噪比、易于获取和稳定性好等特点,受到了广泛关注^[4-5]。

SSVEP是一种由外部周期性视觉刺激(如闪烁的光源)引起的大脑响应,其频率和外部刺激频率一致或成整数倍关系^[6-7]。通过解码SSVEP信号对用户视觉注意目标的识别,从而实现与外部设备交互。然而,使用传统SSVEP刺激容易导致用户视觉不适,降低了BCI系统的使用舒适度和性能^[8]。稳态运动视觉诱发电位(SSMVEP)在缓解用户视觉不适方面展现出了独特的优势和潜力^[9]。SSMVEP是由视觉运动刺激诱发,相比于静态的闪烁光源刺激,视觉运动刺激通过动态变化的视觉元素来诱发脑电响应^[10]。

脑电信号解码算法影响着SSVEP-BCI系统的性能。早期Lin等提出了典型相关分析(CCA)^[11]方法。Chen等利用稳态视觉刺激能够同时诱发谐波频率的特性,提出了滤波器组典型相关性分析(FBCCA)^[12]方法。这些方法处理简单,不需要预训练数据,被广泛的运用于SSVEP的分类中。然而,由于脑电的复杂性,数据往往含有噪声,如自发脑电、眼电和电磁干扰等,严重污染了脑电数据。CCA方法只有在数据长度足够长的情况下才有满意的结果。此外这些特征提取方法是任务相关的,是在特定的先验知识下获得的,并且泛化能力有限^[13]。

深度学习模型具有强大的特征提取能力,可以直接应用于原始数据^[14]。目前,各种深度学习模型

已被用于不同的脑电分类任务,其中卷积神经网络(CNN)应用的最为广泛^[15]。EEGNet是一种紧凑的CNN模型,能够通过时空滤波实现有效的特征提取,具有明显优于传统方法的性能^[16]。复杂卷积神经网络(cCNN)利用SSVEP信号的复杂频谱特征作为CNN的输入,在被试者间和跨被试者训练情况下均实现优异的SSVEP分类性能^[17]。此外,基于时域的卷积神经网络(tCNN)用于解决较短时间刺激诱发的SSVEP在频域上的频率峰值不明显、频域信息作为输入会导致任务相关信息的丢失等问题,并获得了优异的识别性能^[18]。这些基于卷积神经网络的SSVEP解码算法,在处理短时SSVEP数据分类任务上相较于传统CCA和FBCCA算法取得了不错的分类性能。然而,由于CNN通过局部感受野学习特征,缺乏对全局信息的感知能力^[19],导致其对时间序列的处理效果不佳,尚未能充分提取SSVEP的有效特征。

近年来,Transformer网络成为最流行的模型网络之一,在自然语言和图像处理中广泛应用^[20-21]。Transformer网络可以捕获时间序列的全局特征,但会忽略含有细节的局部特征^[22]。由于Transformer网络对时间序列具有出色的特征提取能力,Transformer网络已经开始逐渐应用于脑电解码领域中^[23-24]。Guo等将Transformer网络应用于情绪脑电分类任务中^[25],Xie等将Transformer网络应用于运动想象脑电分类任务中^[23],均获得了优异的解码性能。这些研究说明Transformer网络在脑电信号分类算法中具有很好的研究前景。然而,Transformer网络忽略了对局部特征的学习,而学习局部特征对于脑电解码也是必要的。在这种情况下,需添加额外的提取局部特征的处理,以进行补偿^[26]。

由于真实环境中的噪声复杂,SSMVEP解码难度大,传统CNN分类算法的解码性能仍有较大的

提升空间。CNN 具有局部感知性强的优点,然而由于有限的感知域,CNN 仅提取脑电信号的局部时间和空间特征,缺乏对全局信息的感知力,对处理时间序列数据表现欠佳^[19],解码性能仍有较大提升空间。Transformer 网络在处理全局信息方面具有优势,但是对于局部信息处理能力相对较弱^[22]。因此针对上述问题,本研究将 EEGNet 网络与 Transformer 网络进行结合,旨在融合两者的优势,解决 CNN 算法在处理短时 SSMVEP 信号时分类效果欠佳,感受野有限,缺乏对全局信息的感知等问题。通过 EEGNet 网络对 SSMVEP 信号进行初步的特征提取处理,学习脑电信号局部时间和空间特征,通过 Transformer 网络充分学习 SSMVEP 全局时间序列关系,利用卷积操作和自注意力机制来有效地捕捉与处理脑电信号中的局部和全局信息,增强算法对 SSMVEP 特征的学习,进而实现良好的解码性能。

本文旨在着重探讨融合 EEGNet 网络和图像处理领域的视觉 Transformer 网络在解码短时 SSMVEP 信号的优势与潜力。为充分利用 EEGNet 网络和 Transformer 网络的优势,提出了一个 EEGNet-Transformer 网络,即 EEGNetformer 网络。该网络是在经典的 EEGNet 网络基础上引入 Transformer 网络模块进行改进的一种紧凑型网络结构。具体而言,EEGNetformer 网络结合了 EEGNet 网络的轻量化特性与 Transformer 网络的全局信息处理能力,以提升稳态运动视觉诱发电位解码的性能。通过消融实验表明,加入的 Transformer 网络对 EEGNetformer 网络分类准确率有显著的提升。通过与现有 CNN 解码算法对比实验,验证了 EEGNetformer 在被试者内和被试者外的分类性能均优于目前表现较好的 cCNN、tCNN 和 EEGNet 网络分类算法。

1 实验方法及研究方案

1.1 被试者

10 名视力或矫正视力正常的被试者参与了本次实验,编号为 S1~S10。在向所有被试者充分解释实验程序后,每个被试者都签署了书面知情同意书。

1.2 刺激范式

本研究使用了圆环收缩扩张形式的 SSMVEP 范式。如图 1 所示,该刺激范式由高对比亮度的背景与低对比亮度的圆环组成,低对比亮度圆环的主

要参数有圆环内径和外径,高对比度的背景确定了范式的最大直径,圆环与背景的面积比值 K 存在以下关系

$$K = \frac{S_1}{S - S_1} \quad (1)$$

式中: S_1 表示圆环总面积; S 表示背景的总面积。

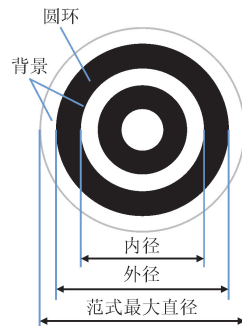


图 1 基于圆环收缩扩张形式的 SSMVEP 范式
Fig. 1 SSMVEP paradigm based on ring contraction-expansion form

范式初始时,设定每个圆环外径成等差数列分布,已知范式的最大直径 $d_{\max}$,则每个圆环的外径 d_1^i 和内径 d_2^i 与面积有如下的关系

$$d_1^i = (2i - 1)d_{\max}/2n \quad (2)$$

$$S^i = \pi(d_1^i)^2 - \pi(d_1^{i-1})^2 \quad (3)$$

$$S_1^i = \pi(d_1^i)^2 - \pi(d_2^i)^2 \quad (4)$$

式中: i 表示从内向外数第 i 个圆环; n 表示圆环的总数; S^i 表示第 i 个圆环外径与 $i - 1$ 个圆环外径所围成的面积; S_1^i 表示第 i 个圆环的面积。

为了保证范式在运动时总体的亮度不发生改变,圆环与背景的面积比值 K 应该时刻保持固定,因此圆环的内径需要通过计算获得,结合式(1)~式(4),可以计算出第 i 个圆环的内径 d_2^i

$$d_2^i = \sqrt{(1 - K)(d_1^i)^2 + K(d_1^{i-1})^2} \quad (5)$$

设圆环外径的变化区间在 $[(i + 1)d_{\max}/n, id_{\max}/n]$,圆环的外径运动曲线为正弦曲线,可以得到每个圆环的外径 d_1^i 随时间的变化如下

$$d_1^i = i + 1 - \cos(2\pi ft)d_{\max}/n \quad (6)$$

式中: f 表示刺激频率; t 表示刺激时间。

刺激范式的排布如图 2 所示,一共有 8 种不同刺激频率(3.00、3.25、3.50、3.75、4.00、4.25 和 4.75 Hz)的圆环范式。此外,为了验证本研究所提出的算法对空闲态和控制态的识别性能,设计了 7 个空闲位置用于采集空闲态脑电数据,保证被试者在注视虚线框内的时候处于系统空闲态并且不会输出控制指令。因此,实验共采集 9 类目标样本,包括

8 类刺激样本和 1 类空闲态样本。

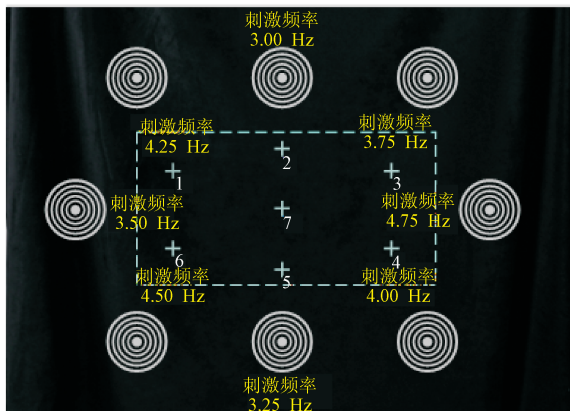


图 2 实验中 SSMVEP 刺激范式

Fig. 2 SSMVEP stimulation paradigm in the experiment

1.3 数据采集

本文通过 g. USBamp 系统采集脑电 (EEG)。SSMVEP 在大脑枕叶区域最为显著^[9], 因此选取位于枕叶区域的 O1、Oz、O2、PO3、POz 和 PO4 电极来采集 EEG。脑电采集电极按照国际 10-20 系统标准进行布置, 参考电极 A1 布置于被试者单侧耳垂, 地电极安放在前额叶处 Fpz 位置。EEG 以 1200 Hz 采样率进行采集。使用 4 阶 48~52 Hz 巴特沃斯陷波滤波器用于消除工频干扰, 8 阶 2~100 Hz 巴特沃

斯带通滤波器去除低频和高频干扰。

1.4 实验流程

实验中要求被试者依次注视 8 类不同刺激频率 (3.00、3.25、3.5、3.75、4.00、4.25、4.5 和 4.75 Hz) 的目标, 每类刺激目标进行了 20 次实验, 在 8 类刺激目标中还会穿插 7 种空闲位置的注视实验。每次实验前均有注视位置提示, 持续 5 s, 包括了提示时间 1 s, 范式刺激时间 3 s, 以及休息时间 1 s, 总共进行了 10 轮实验。每轮实验进行 16 次, 实验结束后, 被试者有 2 min 的休息时间。

1.5 EEGNetformer 结构设计

EEGNet 通过时空滤波器能够有效提取脑电信号的局部时空特征。Transformer 网络利用自注意力机制, 能够更好地捕捉长距离依赖关系, 提升模型的泛化能力和对输入数据的理解能力。因此, 基于 Transformer 网络和 EEGNet 网络, 构建了 EEGNetformer 网络。正如图 3 所示, EEGNetformer 网络由 4 个模块组成: EEGNet 网络中的特征滤波模块、特征混合模块、基于自注意力的特征泛化模块, 以及基于多层感知器 (MLP) 的特征分类模块。图 3 中, C 是脑电信号通道数, T 是脑电信号数据点数, F_1 是时间滤波器数, D 是深度乘数, N 是分类目标数目, $//$ 表示取整数。

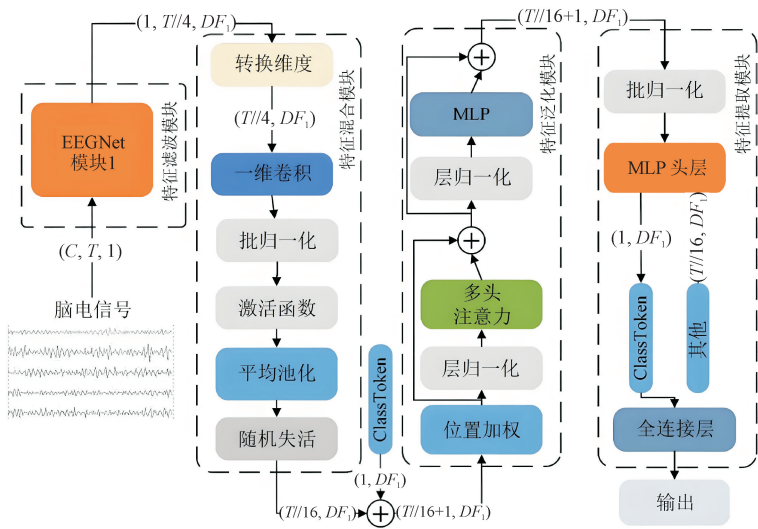


图 3 EEGNetformer 结构示意图

Fig. 3 EEGNetformer structure diagram

特征滤波模块: 特征滤波模块与 EEGNet 网络的第 1 个模块参数一致, 通过时间滤波器和空间滤波器初步提取脑电信号的时空特征。该模块依次执行两个卷积步骤, 首先拟合 8 个尺寸为 (1, 64) 二维

卷积核, 步长为 (1, 1), 并且卷积操作后保持维度不变。对输出结果批标准化后, 使用 16 个尺寸为 (6, 1) 的深度方向卷积核学习空间滤波器以提取空间特征, 并再次进行批标准化, 以促进训练过程并减少过

拟合风险。使用高斯误差线性单元(GELU)作为非线性激活函数。随后,应用尺寸为(1, 4)且步长为(1, 4)的平均池化层平滑时间特征,从而进一步减少过拟合风险并降低计算复杂度。最后,添加一个随机失活(Dropout)层,以 50%的概率舍弃训练参数。具体细节可参考文献^[16]。

特征混合模块:本模块通过一维卷积层对空间滤波后的数据进行特征混合,以进一步提取有效特征,同时减少特征维度和冗余信息。首先,移除维度为 1 的通道,以获得适用于一维卷积的输入。接着,使用 16 个尺寸为 16 的一维卷积核进行卷积操作,步长为 1,并保持卷积后的维度不变。随后,依照特征滤波模块中的方法,对数据进行批标准化,并使用 GELU 作为非线性激活函数。之后,应用尺寸为(1, 4)且步长为(1, 4)的平均池化层。最后,添加一个随机失活层,以 50%的概率舍弃部分训练参数。

特征泛化模块:Transformer 网络能够有效提升模型的泛化能力。在此模块中,利用自注意力机制学习 SSMVEP 特征的全局时间依赖性,以补偿卷积模块在全局特征捕捉能力上的局限性。首先,在输入特征序列的最前端添加一个专门用于分类的向量,称为类别令牌(Class Token)。接着,通过位置嵌入层对包含类标记的特征进行位置信息编码。随后,将编码后的特征输入归一化层进行归一化处理,并通过自注意力机制层进行特征编码。

注意力机制首先将输入向量通过 3 个可训练的变换矩阵 W^Q 、 W^K 、 W^V 进行相应的矩阵变换,得到查询向量 Q 、键值向量 K 和值向量 V 。随后在 Q 和 K 上使用点积来评估不同类别令牌之间的相关性,并通过归一化指数(Softmax)函数对结果进行归一化,得到加权矩阵,即注意力得分。然后,将注意力得分与 V 进行点积运算,从而实现加权特征的组合。这一过程可以表述为

$$A(Q,K,V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{7}$$

式中: d_k 表示 K 向量的维度; A 是经过自注意力机制计算后所得的矩阵。

在自注意力机制层之后,加入了一个多层感知器层,用于进一步对自注意力层输出的特征进行变换和融合。该层首先将自注意力层的输出通过全连接方式输入到隐藏层,然后通过线性整流激活函数生成新的特征表示,最后将结果传递至输出层。

特征分类模块:在本模块中,首先对数据进行批标准化,然后通过多层感知器头层提取类标记。多层感知器头层将输出的特征向量映射到特定的表示空间,从而对特征进行更高级的抽象和表示,以更好地适应具体任务的需求。这种处理方式能够提取出更具表达力的特征表示,从而有助于提高模型的性能。最后,通过全连接层输出分类结果。

具体的 EEGNetformer 网络参数如表 1 所示,其中 DF_1 是空间滤波器数。

表 1 EEGNetformer 详细参数
Table 1 Detailed parameters of EEGNetformer

模块	层	滤波器数	尺寸	输出形状	参数选择
输入	输入层			$(C,T,1)$	
特征滤波	二维卷积层	F_1	(1,64)	(C,T,F_1)	输出尺寸与输入尺寸一致
	批归一化层			(C,T,F_1)	
	二维深度可分离卷积层	DF_1	(1,16)	$(1,T,DF_1)$	输出尺寸与输入尺寸一致
	批归一化层			$(1,T,DF_1)$	
	激活函数			$(1,T,DF_1)$	GELU 激活函数
	平均池化层		(1,4)	$(1,T//4,DF_1)$	步长为(1,4)
	随机失活			$(1,T//4,DF_1)$	随机失活概率 50%
特征混合	转换维度			$(T//4,DF_1)$	
	一维卷积层	DF_1	16	$(T//4,DF_1)$	输出尺寸与输入尺寸一致
	批归一化层			$(T//4,DF_1)$	
	激活函数			$(T//4,DF_1)$	GELU 激活函数
	平均池化层		(1,4)	$(T//4,DF_1)$	步长为(1,4)
	随机失活			$(T//4,DF_1)$	随机失活概率 50%

续表

模块	层	滤波器数	尺寸	输出形状	参数选择
特征泛化	类别令牌层			$(T//4+1,DF_1)$	随机失活概率 20%
	位置嵌入			$(T//4+1,DF_1)$	
	层归一化			$(T//4+1,DF_1)$	
	自注意力			$(T//4+1,DF_1)$	
	层归一化			$(T//4+1,DF_1)$	
	多层感知器层			$(T//4+1,DF_1)$	
特征分类	批归一化层			$(T//4+1,DF_1)$	
	多层感知器头层			$(1,DF_1)$	
	展平层			DF_1	
输出	输出层			N	

1.6 数据扩充与训练

为了对采集到的数据进行训练,首先将原始数据的采样频率降低到 240 Hz,以减少网络的输入数据量,从而提高网络的推理速度。接下来,通过滑动窗口的方式对数据进行扩充,具体来说,先将 9 个目标样本按照 7 : 3 的比例均匀地划分为训练样本和测试样本,然后设定窗口长度,对数据进行滑动窗口处理,重叠率设为 0.8。这种方法确保了训练集和测试集之间没有数据泄露,从而避免因数据泄露而导致模型的测试准确率虚高。最后,通过滑动窗口扩充得到的数据集合并进行拼接,得到最终的训练样本和测试样本,使用 TensorFlow 和 Keras 工具搭建、调试并验证上述算法结构。模型训练采用自适应矩估计(Adam)优化器,学习率设为 0.001,使用交叉熵作为损失函数,最大训练轮数为 100。

1.7 模型评价指标

在深度学习中,真阳性、真阴性、假阳性和假阴性是评估分类模型性能的重要术语。真阳性表示模型将正样本正确预测为正样本的数量,真阴性表示模型将负样本正确预测为负样本的数量,假阳性表示模型将负样本错误预测为正样本的数量,假阴性表示模型将正样本错误预测为负样本的数量。

基于这些指标,深度学习模型的常见评价指标包括准确率 A 、精确率 P 、召回率 R 和 F1 分数 F 。准确率是最直观的性能指标,表示在所有预测中,正确预测的比例。精确率表示在所有预测为正类别的情况下,实际为正类别的比例。召回率表示在所有实际为正类别的情况下,正确预测为正类别的比例。F1 分数是精确率和召回率的调和平均数,用于在精

确率和召回率之间达到平衡。F1 分数值越高,模型的性能通常被认为越好。各指标具体计算式如下

$$A = \frac{T_P + T_N}{T_P + T_N + F_P + F_N}$$

(8)

$$P = \frac{T_P}{T_P + F_P}$$

(9)

$$R = \frac{T_P}{T_P + F_N}$$

(10)

$$F = \frac{2PR}{P + R}$$

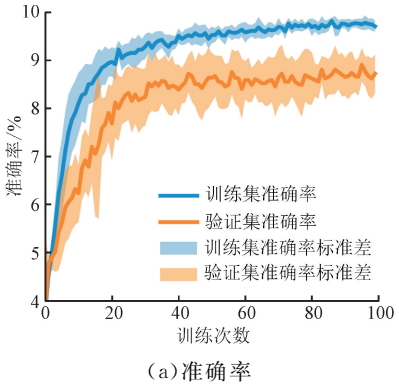
(11)

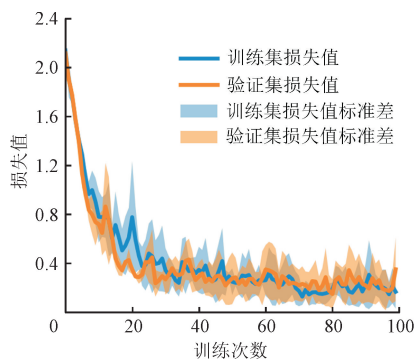
式中: T_P 是真阳性值; T_N 是真阴性值; F_P 是假阳性值; F_N 是假阴性值。

2 结果与讨论

2.1 模型训练收敛性验证

图 4 显示了 EEGNetformer 模型训练过程中训练集和验证集的准确率和损失变化情况。随着训练次数的增加,模型的分类准确率逐渐上升,而损失函数逐步下降。在训练 50 次后,验证集的准确率和损失值基本趋于稳定。这表明本研究提出的 EEGNetformer 模型在解码 SSMVEP 时表现出良好的收敛性,验证了算法结构的合理性。





(b) 损失值

图 4 EEGNetformer 网络训练集和验证集准确率和损失值的动态变化

Fig. 4 Dynamics of accuracy and loss at training set and validation set in EEGNetformer network training

2.2 模块消融实验验证

为了探究 EEGNetformer 中各模块对模型性能贡献,进一步进行了消融实验。具体地,分别依次去除特征滤波模块、特征混合模块和特征泛化模块,并将剩余模块拼接成新的网络架构,在 2 s 的数据上进行训练和测试。实验结果通过配对样本 t 检验进行显著性分析。

如图 5 所示,移除特征滤波模块后,除被试者 S3 外,其他所有被试者的识别准确率均有所下降。其中,被试者 S8 的下降幅度最大,为 6.3%,而被试者 S3 的准确率则上升了 0.7%。总体来看,所有被试者的平均准确率相比完整的 EEGNetformer 模型下降了 2.2%。在移除特征混合模块后,所有被试者的识别准确率均出现下降。其中,被试者 S10 的准确率下降最为显著,为 7.6%,被试者 S3 的下降幅度最小,为 0.7%。总体上,所有被试者的平均准确率显著下降了 2.8% ($p < 0.001$)。移除特征泛化模块后,所有被试者的识别准确率均出现下降,其中被试者 S6 下降最多,为 13.1%,被试者 S4 的下降幅度最小,为 2.7%。所有被试者的平均准确率显

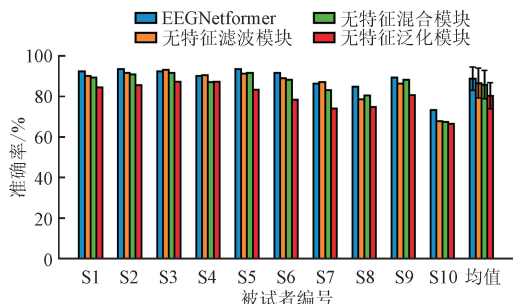
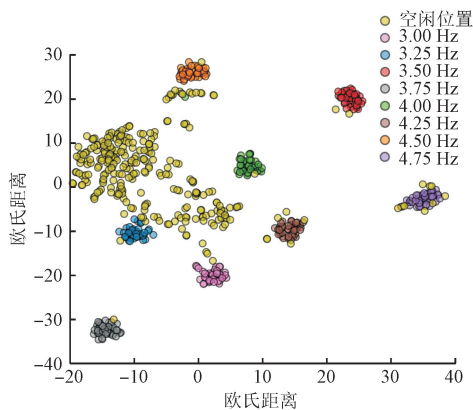


图 5 EEGNetformer 网络消融实验结果

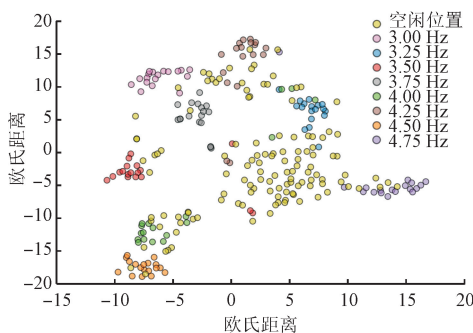
Fig. 5 Results of the EEGNetformer network ablation experiment

著下降了 8.5% ($p < 0.001$)。

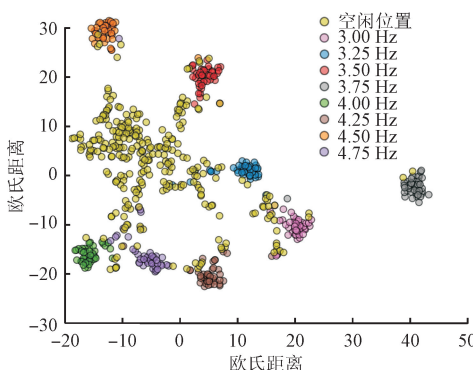
在消融实验中,消融后的网络准确率相较于完整的 EEGNetformer 网络均有所下降,其中移除特征泛化模块后的准确率下降幅度最大。为进一步探究特征泛化模块在 EEGNetformer 网络中的具体作用,采用 t -分布随机近邻嵌入(t -SNE)方法,对典型被试者 S6 的数据进行了分析。 t -SNE 是一种基于概率的技术,可以将数据的高维结构映射到低维空间中,保留数据之间的欧氏相似性关系。 t -SNE 根据高维数据点之间的相似度计算它们之间的欧氏距离并生成一组新的坐标,使得每个数据点在新的低维空间中都能够更清晰地显示其在原始高维空间中的关系。图 6 对比了不同刺激频率下 EEGNetformer 网络和 EEGNetformer 移除特征泛化模块后的网络在训练集和测试集上提取的特征聚类差异。



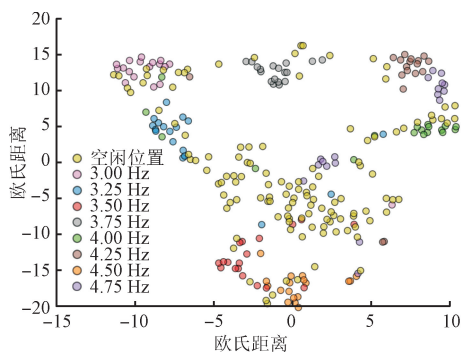
(a) EEGNetformer 网络在训练集



(b) EEGNetformer 网络在测试集



(c) 移除特征泛化模块的网络在训练集



(d)移除特征泛化模块的网络在测试集

图 6 被试者 S6 在不同网络训练集和测试集上的聚类结果
Fig. 6 Clustering results of subject S6 in different network on the test set and the training set

从图 6(a)和图 6(c)可以看出,移除特征泛化模块后,提取的特征边界较为模糊,部分类别之间出现了混淆现象。加入特征泛化模块后,类间距离显著增大,类内距离减小,各类别之间的区分度明显提高。从图 6(b)和图 6(d)可以看出,使用 EEGNetformer 网络提取的特征虽然类内距较小,但部分类别的类间边界仍存在模糊现象,且少量样本存在混淆。相比之下,移除特征泛化模块后的特征提取结果在某些类别上出现了较大程度的混淆,并且类间距减小,类间区分度明显下降。

综上所述,加入特征泛化模块后,经 *t*-SNE 降维处理的特征表现出更小的类内距和更大的类间距,具备更明显的特征边界。此外,分类准确率上与未添加特征泛化模块的模型相比差异显著。这表明特征泛化模块能够有效提升模型对数据的理解和区分能力。

2.3 EEGNetformer 网络被试者内性能分析

通过评估和对比不同时长下 EEGNetformer 网络算法与其他 4 种性能优异的对比算法(即 cCNN、tCNN、EEGNet-SSVEP (EEGNetS) 和 EEGNet),进一步探讨了 EEGNetformer 网络算法在解码 SSMVEP 中的性能表现。图 7 展示了 5 种算法在不同时长的数据上的识别准确率。

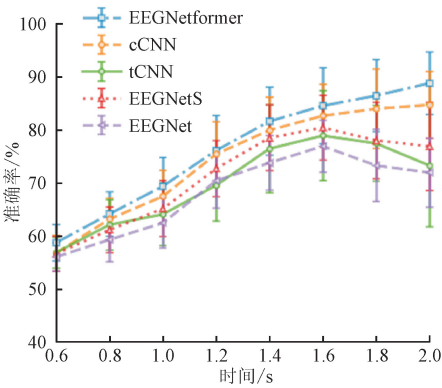


图 7 5 种算法对不同时长 SSMVEP 数据的识别准确率对比
Fig. 7 Comparison of recognition accuracies of five algorithms for SSMVEP data with different time-window lengths

如图 7 所示,EEGNetformer 网络在各个时长的识别准确率均优于其他算法。在所有时长上,EEGNetformer 网络与 tCNN、EEGNetS 和 EEGNet 的识别准确率均存在显著性差异($p<0.05$)。特别是在 2 s 时,EEGNetformer 网络与其他算法的准确率差异最大,其准确率达到 $88.9\% \pm 6.6\%$,相比最优的对比算法 cCNN 提高了 4.2% ($p<0.001$),相比表现最差的 EEGNet 提高了 17.4% ($p<0.001$)。这些结果进一步证明了 EEGNetformer 网络在解码 SSMVEP 中的显著优势。

2.4 EEGNetformer 网络跨被试者性能分析

进一步分析了 EEGNetformer 网络在跨被试者情况下的解码性能。在跨被试者实验中,以一名被试者的数据作为训练集,其余 9 名被试者的数据作为测试集,依次进行 10 次测试,并将这 10 次测试结果的平均值作为最终结果。在 2 s 数据条件下,使用 5 种算法进行训练,并通过准确率和 F1 分数衡量实验结果,具体结果如表 2 所示。使用配对样本 *t* 检验对结果进行显著性检验,* 代表 $p<0.05$,** 代表 $p<0.001$,加粗的数据表示同一指标下不同算法中的最高值。

表 2 2 s 数据长度下 5 种算法在跨被试者情况下的准确率和 F1 分数结果

Table 2 Accuracy and F1 score results of the five algorithms in the inter-subject case for 2 s data

指标	数值				
	EEGNet former	cCNN	tCNN	EEGNetS	EEGNet
准确率	0.691±0.043	0.631±0.054**	0.444±0.091**	0.594±0.067**	0.543±0.043**
F1 分数	0.673±0.061	0.574±0.082**	0.416±0.163**	0.502±0.097**	0.477±0.088**

在跨被试者条件下,EEGNetformer 网络的准确率和 F1 分数均显著优于 4 种对比算法($p<0.001$)。

与 cCNN、tCNN、EEGNetS 和 EEGNet 算法相比,EEGNetformer 网络的跨被试者识别准确率分别提升

了6%、24.7%、9.7%和14.8%,F1分数则分别提升了9.9%、25.7%、17.1%和19.6%。这表明,EEGNetformer网络在提取单一被试者的特征时具有较高的稳定性,能够较好地适应不同被试者的差异。同时,也说明了EEGNetformer网络在提取和学习目标特征方面的能力较强,优于其他4种算法。

3 结 论

本文针对现有SSMVEP-BCI范式中CNN分类算法在解码微弱特征信号能力上的不足,提出了一种结合Transformer网络和EEGNet网络的EEGNetformer网络。EEGNetformer网络通过时空滤波实现局部特征的提取,然后基于Transformer网络进一步捕捉全局信息,从而增强模型的泛化能力和特征表达能力。实验结果表明,EEGNetformer网络在解码SSMVEP时表现出显著的有效性。通过消融实验,以准确率作为评价指标,结果显示引入Transformer网络显著提高了模型分类准确率。此外,对输出特征的聚类分析表明,添加Transformer网络能够有效减小类内距,并增大类间距。进一步在被试者内和跨被试者的实验中对EEGNetformer网络的性能进行了评估,结果表明EEGNetformer网络在性能上优于cCNN、tCNN、EEGNetS和EEGNet等传统的卷积神经网络算法。总之,EEGNetformer网络在SSMVEP解码方面展现了良好的性能,证明了其在脑电信号解码与脑机接口领域的应用潜力。EEGNetformer网络的设计将有助于推动CNN在SSVEP-BCI解码方面的研究。通过结合EEGNet网络和Transformer网络的优势,该方法克服CNN缺乏对全局信息的感知力的局限性。促进了对脑电信号局部与全局特征的全面理解,提高了对短时SSMVEP信号的解码性能。这种创新方法有助于推动CNN在SSMVEP-BCI解码领域的发展,为未来开发更高效、更精确的脑电信号解码算法提供了新思路和技术支撑。

参考文献:

- [1] WOLPAW J R. Brain-computer interfaces [M]. Amsterdam, Holland: Elsevier, 2013: 67-74.
- [2] 姜耿, 赵春临. 基于EEG的脑机接口发展综述 [J]. 计算机测量与控制, 2022, 30(7): 1-8.
JIANG Geng, ZHAO Chunlin. A review of the development of brain-computer interface based on EEG [J]. Computer Measurement & Control, 2022, 30(7): 1-8.
- [3] VÄRBÜ K, MUHAMMAD N, MUHAMMAD Y. Past, present, and future of EEG-based BCI applications [J]. Sensors, 2022, 22(9): 3331.
- [4] BONDRE C, KAPGATE D. Brain computer interfaces using SSVEP: an overview [J]. International Journal of Engineering & Technology, 2014, 1: 9-17.
- [5] 邱爽, 杨帮华, 陈小刚, 等. 非侵入式脑-机接口编解码技术研究进展 [J]. 中国图象图形学报, 2023, 28(6): 1543-1566.
QIU Shuang, YANG Banghua, CHEN Xiaogang, et al. A survey on encoding and decoding technology of non-invasive brain-computer interface [J]. Journal of Image and Graphics, 2023, 28(6): 1543-1566.
- [6] KAPGATE D, KALBANDE D. A review on visual brain computer interface [C]//Advancements of Medical Electronics. New Delhi: Springer India, 2015: 193-206.
- [7] 梁栗炎, 孔姝懿, 张倩, 等. 脑机接口的范式、算法与编解码概念探讨与研究 [J]. 信息通信技术与政策, 2024, 50(5): 61-70.
LIANG Liyan, KONG Shuyi, ZHANG Qian, et al. Exploration and research on paradigm, algorithm, as well as encoding and decoding concepts of brain-computer interface [J]. Information and Communications Technology and Policy, 2024, 50(5): 61-70.
- [8] 李丽, 陈泉宇, 随力. 稳态运动视觉诱发电位的诱发及在脑机接口中的应用进展 [J]. 上海理工大学学报, 2022, 44(1): 27-33.
LI Li, CHEN Xiaoyu, SUI Li. Induction of steady-state motion visual evoked potential and its application in brain-computer interface [J]. Journal of University of Shanghai for Science and Technology, 2022, 44(1): 27-33.
- [9] XIE Jun, XU Guanghua, WANG Jing, et al. Steady-state motion visual evoked potentials produced by oscillating Newton's rings: implications for brain-computer interfaces [J]. PLoS One, 2012, 7(6): e39707.
- [10] 刘贵彤, 张志敏, 柴晓珂, 等. 基于SSMVEP的脑-机接口视觉刺激探究 [J]. 中国医疗器械杂志, 2018, 42(5): 313-316.
LIU Guitong, ZHANG Zhimin, CHAI Xiaoke, et al. Study of steady state motion visual evoked potential-based visual stimulation of BCI system [J]. Chinese Journal of Medical Instrumentation, 2018, 42(5): 313-316.
- [11] LIN Zhonglin, ZHANG Changshui, WU Wei, et al.

- Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs [J]. IEEE Transactions on Biomedical Engineering, 2006, 53(12): 2610-2614.
- [12] CHEN Xiaogang, WANG Yijun, GAO Shangkai, et al. Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain-computer interface [J]. Journal of Neural Engineering, 2015, 12(4): 046008.
- [13] ZHANG Yue, XIE S Q, WANG He, et al. Data analytics in steady-state visual evoked potential-based brain-computer interface: a review [J]. IEEE Sensors Journal, 2021, 21(2): 1124-1138.
- [14] CRAIK A, HE Yongtian, CONTRERAS-VIDAL J L. Deep learning for electroencephalogram (EEG) classification tasks: a review [J]. Journal of Neural Engineering, 2019, 16(3): 031001.
- [15] XU Dongcen, TANG Fengzhen, LI Yiping, et al. An analysis of deep learning models in SSVEP-based BCI: a survey [J]. Brain Sciences, 2023, 13(3): 483.
- [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] RAVI A, BENI N H, MANUEL J, et al. Comparing user-dependent and user-independent training of CNN for SSVEP BCI [J]. Journal of Neural Engineering, 2020, 17(2): 026028.
- [18] DING Wenlong, SHAN Jianhua, FANG Bin, et al. Filter bank convolutional neural network for short time-window steady-state visual evoked potential classification [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2021, 29: 2615-2624.
- [19] PENG Zhiliang, HUANG Wei, GU Shanzhi, et al. Conformer: local features coupling global representations for visual recognition [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2021: 357-366.
- [20] HAN Kai, WANG Yunhe, CHEN Hanting, et al. A survey on vision transformer [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(1): 87-110.
- [21] HAN Kai, XIAO An, WU Enhua, et al. Transformer in transformer [C]//Advances in Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates, Inc, 2021: 15908-15919.
- [22] MAURÍCIO J, DOMINGUES I, BERNARDINO J. Comparing vision transformers and convolutional neural networks for image classification: a literature review [J]. Applied Sciences, 2023, 13(9): 5521.
- [23] XIE Jin, ZHANG Jie, SUN Jiayao, et al. A transformer-based approach combining deep learning network and spatial-temporal information for raw EEG classification [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022, 30: 2126-2136.
- [24] SONG Yonghao, JIA Xueyu, YANG Lie, et al. Transformer-based spatial-temporal feature learning for EEG decoding [EB/OL]. (2021-06-11) [2024-08-02]. <https://arxiv.org/abs/2106.11170>.
- [25] GUO Jiayi, CAI Qing, AN Jianpeng, et al. A Transformer based neural network for emotion recognition and visualizations of crucial EEG channels [J]. Physica: A Statistical Mechanics and its Applications, 2022, 603: 127700.
- [26] 戴洋毅, 何康, 瑚琦, 等. CNN-Transformer 混合模型在计算机视觉领域的研究综述 [J]. 建模与仿真, 2023, 12(4): 3657-3672.
- DAI Yangyi, HE Kang, HU Qi, et al. Review of CNN-Transformer hybrid model in computer vision [J]. Modeling and Simulation, 2023, 12(4): 3657-3672.

(编辑 杜秀杰)