

# 基于 SSVEP 信号的相频特性分类算法研究<sup>\*</sup>

丛佩超 陈熙来 肖宜轩 李文彬 刘俊杰 张欣

(广西科技大学机械与汽车工程学院 柳州 545616)

**摘要:** 目前基于稳态视觉诱发电位(SSVEP)的脑-机接口在人机协作中受到广泛关注,现有面向 SSVEP 信号的相位与频率信息的深度学习分类方法,仍存在由于信息利用不充分导致的 SSVEP 信号分类效果较差等问题。而目前已出现多种分类算法用于解决上述问题。本文基于迁移学习思想提出一种用于 SSVEP 信号分类的深度神经网络模型,将快速傅里叶变换后的复向量作为输入,对各个导联的实、虚部向量进行卷积,学习对应的相频特性。该模型分为两部分:第一部分利用所有被试者之间的统计共性获得相位和频率信息的全局相频特征模块;第二部分利用训练好的全局相频特征模块对局部相频特征模块进行初始化,通过局部相频特征模块的进一步强化学习对训练参数进行微调,以减少每个被试者之间的个体差异。在公开数据集 BETA 上进行测试,在时窗长度为 1.5 s 时,平均准确率和平均信息传输率分别为 89.98% 和 71.80 bit/min。实验结果表明,与其他方法相比,本文的分类算法模型取得了较为不错的分类效果,所设计的全局、局部相频特征模块能够改善个体差异因素对分类结果的影响,为深入挖掘、利用 SSVEP 信号中的相位和频率信息提供了全新思路。

**关键词:** 稳态视觉诱发电位;迁移学习;深度神经网络;相频特性

**中图分类号:** TN911.7    **文献标识码:** A    **国家标准学科分类代码:** 520.2060

## Research on phase-frequency characteristic classification algorithm based on SSVEP signals

Cong Peichao Chen Xilai Xiao Yixuan Li Wenbin Liu Junjie Zhang Xin

(School of Mechanical and Automotive Engineering, Guangxi University of Science and Technology, Liuzhou 545616, China)

**Abstract:** Currently, the brain-machine interface based on steady-state visual evoked potential (SSVEP) has received wide attention in human-computer collaboration, and the existing deep learning classification methods oriented to the phase and frequency information of SSVEP signals still have problems such as poor classification of SSVEP signals due to insufficient utilization of the information. And a variety of classification algorithms have appeared for solving the above problems. In this paper, a deep neural network model for SSVEP signal classification is proposed based on the idea of migration learning, which takes the complex vectors after the fast Fourier transform as inputs, and convolves the real and imaginary part vectors of each lead to learn the corresponding phase frequency characteristics. The model is divided into two parts: the first part uses the statistical commonality among all subjects to obtain the global phase-frequency feature module for phase and frequency information; the second part uses the trained global phase-frequency feature module to initialize the local phase-frequency feature module, and fine-tunes the training parameters through further reinforcement learning of the local phase-frequency feature module in order to reduce the individual differences between each subject. Tested on the public dataset BETA, the average accuracy and average information transfer rate are 89.98% and 71.80 bit/min, respectively, when the time window length is 1.5 s. The experimental results show that the classification algorithm model in this paper achieves a relatively good classification effect compared with other methods, and the designed global and local phase-frequency feature modules are able to improve the effect of individual differences on the classification results. The designed global and local phase-frequency feature module can improve the influence of individual differences on the classification results, which provides a brand new idea for the in-depth mining and utilization of phase and frequency information in SSVEP signals.

**Keywords:** steady-state visual evoked potential (SSVEP); transfer learning; deep neural network (DNN); phase-frequency characteristics

## 0 引言

脑-机接口(brain-computer interface, BCI)是一种连接

人脑和外部设备的通信方法。BCI 对大脑活动产生的脑电信号(electroencephalogram, EEG)进行采样和处理,以解码人类思维并了解大脑的工作机制。根据诱发 EEG 信号

类型, BCI 系统可分为三类:P300-BCI、运动想象(motor imagery, MI)-BCI 和稳态视觉诱发电位(steady-state visual evoked potential, SSVEP)-BCI。其中基于 SSVEP 信号的 BCI 具有训练时间短、高信噪比等优点<sup>[1]</sup>, 因此, 该类 BCI 在学术研究和相关领域得到广泛关注<sup>[2]</sup>。SSVEP 信号中不同频率的视觉刺激会引发与之对应的近似正弦信号。通过事先将不同频率闪烁的视觉刺激与相应的字符关联起来, 被试者可以通过观察目标视觉刺激来诱发相应的 SSVEP 信号。通过算法分析这些信号, 并将其与事先确定的频率进行比较, 可以确定被诱发 SSVEP 信号所对应的字符或语义。然而, 要实现高性能的 BCI, 系统需要能够在短时间内通过对脑电信号的分析, 快速准确地识别目标。因此, 基于 SSVEP 信号的 BCI 系统性能优劣很大程度上取决于对 SSVEP 信号快速、准确的分类能力<sup>[3]</sup>。然而, 作为整个 BCI 系统中最关键的部分, 如何进一步提高 SSVEP 信号分类准确率仍然面临着巨大的挑战。

当前用于 SSVEP 信号分类任务的主要方法是基于典型相关分析算法(canonical correlation analysis, CCA)。2015 年 Chen 等<sup>[4]</sup>通过构建多个 CCA 子模型并将它们集成, 提出扩展典型相关分析算法(extended CCA, eCCA)。eCCA 作为一种扩展的 CCA 方法, 仍然基于对信号之间线性关系的假设, 使其在 SSVEP 信号中的非线性频率信息利用上不足。2018 年, 基于任务相关分析算法(task-related component analysis, TRCA)具有抑制非目标信号、强大抗噪性以及训练速度快等特点, Nakanishi 等<sup>[5]</sup>在其基础上提出集成任务相关成分分析算法(ensemble task-related component analysis, eTRCA), 该算法引入参考信号和自适应信号选择策略, 可以进一步增强目标信号的辨别度, 更好地减少非目标信号对目标信号的干扰。然而, 该算法通常需要事先定义时间窗口, 而这种固定窗口不足以适应 SSVEP 信号的动态特性, 使其在充分利用特征信息方面受到限制。2023 年 Li 等<sup>[6]</sup>提出一种基于 eCC 的改进方法, 该方法在参考信号中加入 SSVEP 刺激范式设计参数, 用于提高 SSVEP 信号的频率识别性能。然而, 该算法通常基于线性关系的假设, 限制了对与非线性频率信息的捕捉, 此外, 它们无法多维度学习除 SSVEP 信号之外其他可供判别的信息成分, 而且, 不能针对被试者间的个体差异进行强化学习。因此, 近年来研究人员逐渐将监督算法应用于 SSVEP 信号的分类任务之中, 深度学习理论作为目前最强大的监督学习方法之一, 在相关领域引起广泛关注。

近些年来, 深度学习理论(deep learning, DL)在计算机视觉(computer vision, CV)、自然语言处理(natural language processing, NLP)和推荐系统(recommendation system, RS)等诸多领域快速发展, 并取得巨大成功<sup>[7-9]</sup>。与传统方法不同, 基于深度学习理论的相关方法可在端到端场景下, 同时进行特征提取和分类器训练, 有效避免手工设计特征所需的经验设计。在过去的十年中, 新兴的深度

学习理论引起人们对大脑信号分析越来越多的兴趣, 它为解码大脑状态提供一种极具潜力的手段<sup>[10-12]</sup>。目前, 基于深度学习理论的 SSVEP 信号分类模型主要以卷积神经网络(convolutional neural network, CNN)为主。2019 韩向可等<sup>[13]</sup>提出一种全新的多模态混合脑机接口系统, 该系统通过 STFT 将多通道的时域信息转化为具有时-频-空域三维特征的二维特征图, 并通过 CNN 完成特征分类。尽管 STFT 广泛应用于时频分析, 但该系统在设计中, 注重频率信息的利用, 相对于相位信息的利用较少。2021 年 Keerthi 等<sup>[14]</sup>提出一种通过生成频谱图像对脑电信号进行预处理的新方法, 以实现基于 CNN 的 MI 识别。该方法基于脑电信号的变分模态分解(variational mode decomposition, VMD)模式, 将所有模式的短时傅里叶变换(short time fourier transform, STFT)排列起来形成脑电频谱图像, 从而对脑电信号进行分类。2022 年 Zhang 等<sup>[15]</sup>提出一种基于决策融合反向传播(back propagation, BP)神经网络, 该方法使用自回归模型(autoregressive model, AR)和功率谱密度(power spectral density, PSD)将 EEG 信号的频域特征进行分类, 并将其应用在处于监控关键位置操作员的警觉性检测, 避免警觉性降低导致的潜在灾难。2023 年谢平等<sup>[16]</sup>提出一种基于多任务学习的快速序列视觉呈现脑电图分类模型。该模型通过提取与任务关联的深度特征, 并结合分类任务和超球约束任务联合学习, 从而优化特征。采用多任务学习策略的设计以提高模型的泛化性能。尽管出现众多以频域信号为输入的深度学习类 EEG 识别方法, 但有些模型采用频谱图、时域频谱图或功率谱图等形式作为输入, 仅利用频域幅度信息, 对相位信息利用不足; 有些模型使用原始数据作为输入, 对频域特征的提取不够充分。上述方法会导致系统识别的稳定性和准确率均出现下降。一般来说, 一个高性能的 BCI 系统很大程度上依赖于校准数据的数量, 然而, 收集校准数据是一件费时、费力的工作, 长时间实验很容易导致被试者出现疲劳, 从而降低收集数据的质量<sup>[17]</sup>。为了让 EEG 信号实现更加高效、准确的分类效果, 并保证跨被试者、多分类情况下的准确性和可靠性, 需要综合考虑 EEG 信号在被试者跨试次以及被试者个体间更深层次相位、频率特征的差异性, 如果忽略这两个因素中的任何一个, 都会导致 EEG 信号分类效果的降低。因此, 必须充分挖掘、利用被试者本身不同的实验试次以及被试者个体间在频域上表示相位、频率特征信息的潜在关联性, 以实现对 EEG 信号更加准确的分类。而迁移学习通过自适应和知识迁移可实现对多源数据信息的学习, 从而对神经网络进行进一步调整<sup>[18]</sup>。例如 Guney 等<sup>[19]</sup>提出一种新颖的深度神经网络架构(deep neural network, DNN), 并使用迁移学习的微调策略技术以提升个体内的分类能力。受此启发, 本文基于迁移学习思想从频域角度出发充分挖掘 SSVEP 信号蕴含的相位、频率信息, 并对其进行利用, 以期改善 SSVEP 信号的识别准确率, 从而提高

网络整体对 EEG 信号的识别能力。

基于以上分析,本文针对 SSVEP 信号的相位、频率信息充分挖掘、利用问题,运用迁移学习思想,提出一种新颖的深度神经网络模型—PLFC-DNN (phase & frequency feature classification)。它由两部分构成:全局相频特征模块和局部相频特征模块。前者用于学习所有被试者相位、频率的共性特征;后者用于减少各被试者个体间相位、频率特征的个体性差异。通过将两大模块相互嵌套,利用全局相频特征模块学习所有被试者共性特征得到的权重参数,逐层赋值给局部相频特征模块。接着,通过局部相频特征模块的再一次强化学习,进一步微调权重参数,使 PLFC-DNN 能从不同角度学习 SSVEP 信号蕴含的相位、频率特

征信息,实现对二者的充分挖掘与利用,以提高其分类精度。实验结果表明:PLFC-DNN 可有效提高 SSVEP 信号中相位、频率特征的分类准确率和信息传输率,相比 CCA、eCCA、eTRCA 等主流算法的分类准确率分别提升 21.74%、6.87%、7.81%,信息传输率各提升 29.05、11.46、13.09 bit/min。与其他方法相比,本文构建的分类模型取得了显著优越的分类效果。通过考虑从全局和局部多个维度对特征进行深入挖掘和充分利用,提高了分类准确率和信息传输率。

## 1 数据集及预处理

本文研究内容的总体架构,如图 1 所示。

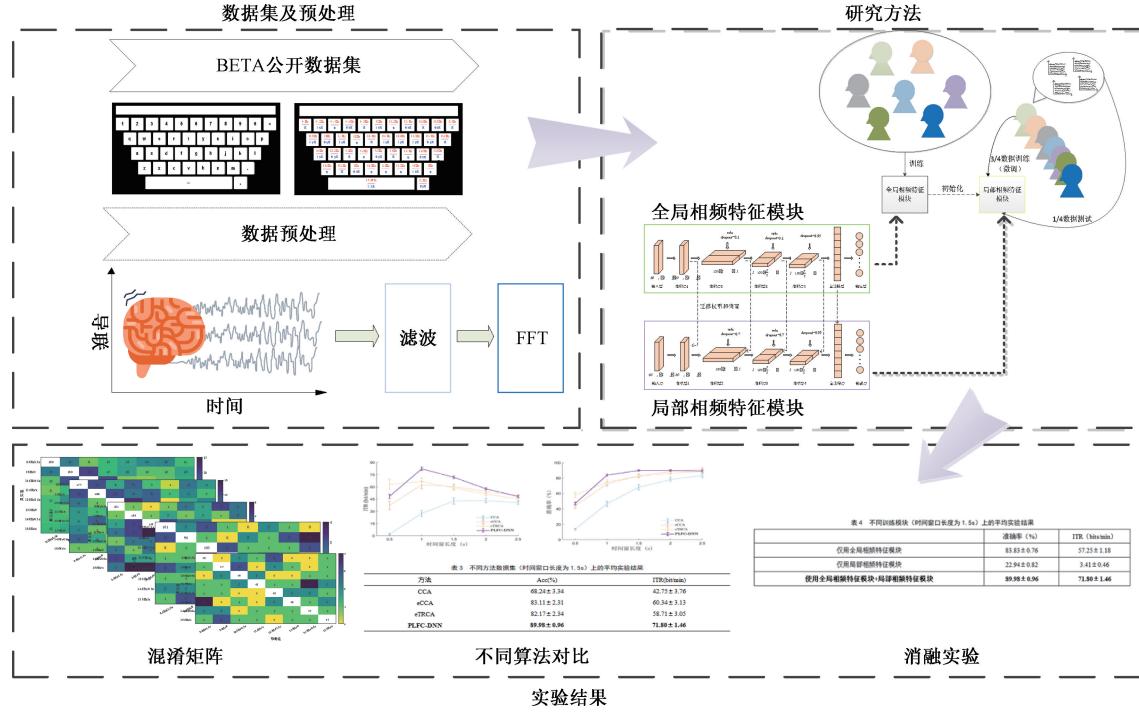


图 1 研究内容总体架构

### 1.1 BETA 公开数据集

本文采用 Liu 等<sup>[20]</sup>公开的数据集进行测试。该数据集是面向 SSVEP-BCI 应用的大型基准数据库,数据使用 Neuroscan 公司生产的 64 导联 SynAmps2 记录仪进行采集,导联分布采用国际系统电极位置 10-10。BETA 数据集是为实际应用而开发,该数据集使用的范式是带有 40 种刺激目标的虚拟键盘拼写器,如图 2 所示。40 个目标包括:10 个数字、26 个字母和 4 个符号(点、逗号、退格、空格),最上面的空白矩形用于结果反馈,其中,图 2(a)为虚拟键盘图,即所呈现的刺激界面,图 2(b)为各虚拟按键所对应的刺激频率和相位信息,其中,频率以 0.2 Hz 为步长,从 0.8 Hz 逐渐递增到 15.8 Hz,并且使用 0、0.5π、π 和 1.5π 这 4 种相位对刺激进行联合频率—相位调制(joint frequency-phase modulation,JFPM)<sup>[20]</sup>。该刺激范式呈现

在 27 英寸的 LED 显示器上,分辨率为 1 920 pixels × 1 080 pixels,刷新率为 60 Hz。

该数据集共招募 70 名被试者,其平均年龄为 25 岁。实验前所有被试者均已签署书面协议,保证实验的合规性,符合赫尔辛基宣言,并获得清华大学伦理委员会的批准。在数据集发布之前,清华大学对原始的脑电数据进行过简单处理和整合,这包括将采样率从 1 000 Hz 降采样为 250 Hz,采用 0.15~200 Hz 的硬件滤波,并内置陷波滤波器以消除 50 Hz 的工频干扰。在数据采集过程中,所有电极的阻抗都保持在 10 kΩ 以下,并使用顶点 Cz 电极作为参考电极。经过简单处理后的脑电数据,以 4 维张量形式存储在 MATLAB 的.mat 文件之中。其中,张量的 4 个维度分别表示通道(channel)、时间点(time point)、块(block)、条件(condition)。每次实验都包括刺激前 0.5 s

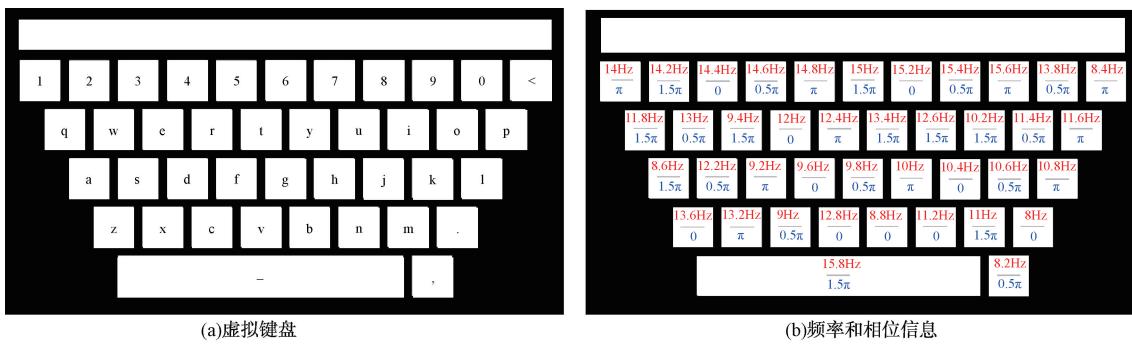


图 2 BETA 公开 SSVEP 数据集刺激范式界面

的数据, 刺激期间 2 s 或 3 s, 以及刺激后 0.5 s 的数据。对于前 15 个被试者, 刺激期间的时间窗口长度为 2 s; 对于后 55 个被试者, 时间窗口长度为 3 s, BETA 数据集的具体描述如表 1 所示。

表 1 BETA 数据集描述

特征	BETA 数据集描述
被试者数量	70
EEG 通道数	64
原始采样率/Hz	1 000
降采样后/Hz	250
trials 数量	40
block 数量	4
参考电极	Cz
单个 trial 的时间/s	3(S1~S15) 4(S16~S70) 9(10~20 岁) 53(20~30 岁)
被试者不同年龄段人数	4(30~40 岁) 3(40~50 岁) 1(60~70 岁)

## 1.2 数据预处理

数据集在公开之前清华大学虽已做简单处理, 但为进一步提高脑电数据的质量并减少干扰和伪相关, 使脑电数据更加可靠, 同时, 也为后续 SSVEP 信号的相位、频率特征提取建立基础, 因此, 本文将对该数据集做相应的预处理。使用滤波、去噪技术, 以减少眼电和肌电信号的干扰, 从而捕捉到 EEG 信号中包含的真实脑电活动。具体而言: 首先, 对数据集进行筛选, 滤除幅值超过  $\pm 100 \mu\text{V}$  的数据, 从而过滤异常值或异常幅度的数据点, 降低由肌肉活动引起的伪相关。考虑到网络模型结构、训练的数据量以及采样率等因素, 本实验选用 S16~S70(55 名被试者)的数据, 并选用每个被试者全部 4 个 block 数据 40 类刺激中的 8 类刺激信号数据, 所选刺激的频率、相位分别为  $8.6 \text{ Hz} + 1.5\pi, 9.6 \text{ Hz} + 0, 10.6 \text{ Hz} + 0.5\pi, 11.6 \text{ Hz} + \pi$ ,

$12.6 \text{ Hz} + 1.5\pi, 13.6 \text{ Hz} + 0, 14.6 \text{ Hz} + 0.5\pi, 15.6 \text{ Hz} + \pi$ 。使用 MATLAB 中 EEGLAB 工具箱对脑电数据进行带通滤波处理, 滤波范围为 6~50 Hz, 以减少噪声影响, 保留感兴趣频段。本实验选用 64 导联中 30 个导联的 SSVEP 脑电数据, 30 个导联电极分布在脑顶叶区域和枕叶区域, 如图 3 所示, 导联的序号为 34~42 及 44~64。对于该数据集, 在已有研究中视觉延迟时间被认为是 0.13 s<sup>[19]</sup>, 因此, 在剔除刺激前 0.5 s 提示和刺激后 0.5 s 休息间隔后, 提取 0.13~2.63 s 之间的数据。

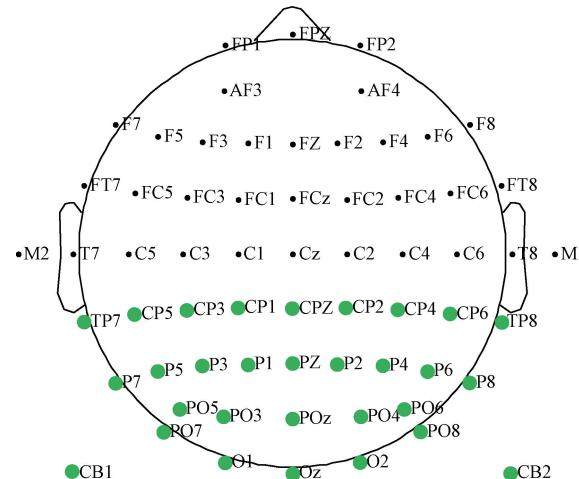


图 3 选择导联数

在数据划分方面, 由于 EEG 信号的非线性和不平稳特性, 不同被试者对相同视觉刺激频率反应的起始时间、时长存在差异。这种差异性会导致模型对不同被试者的识别效果差别较大, 从而降低系统的稳定性。因此, 本文对训练数据集采用 0.5、1.0、1.5、2.0、2.5 s 这 5 个时间窗口, 并按照 0.5 s 的步长对数据进行划分。接着, 将划分好的数据集沿着各个导联方向进行 FFT 处理, 其形式如图 4 所示。模型中使用的信号长度为 0.5、1.0、1.5、2.0、2.5 s, 采样率为 250 Hz, 因此, 每个数据段的采样点数分别为 125、250、375、500、625。根据傅里叶变换的共轭对称性, 只需采用单边频率特征, 所以, 最终输入模型中的频率点  $f_i$ ,  $i \in \{1, 2, \dots, 5\}$  个数分别为 63、126、188、251、313。为充分

挖掘、利用各个采样点的相位、频率信息,采用 FFT 变换后得到的复数实、虚部向量,并将它们沿通道方向拼接成一个 2 维频域特征矩阵,其中,该特征矩阵中相位信息由实、虚部共同决定,  $\varphi_n = \arctan(b_n/a_n)$ ,  $n \in \{f_1, f_2, f_3, f_4, f_5\}$ , 频率信息为该复数的索引值。最后,为满足网络模型对训练数据维度的要求,使其能够像处理图像一样进行卷积操作,使用 reshape 函数为训练数据增加 1 个维度,最终得到的训练数据样本维度为  $60 \times f_i \times 1$ 。由于提出的模型是针对多个用户,因此,在预处理的最后阶段,需将所有被试者的数据进行拼接,形成跨被试者的 SSVEP 信号数据集。

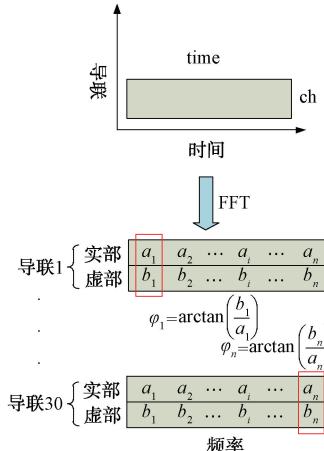


图 4 FFT 处理流程图

通过上述操作,成功对脑电数据进行预处理,并且成

功提取 SSVEP 信号的相位、频率特征。这样本文得到每个训练样本的三维特征矩阵,为后续的模型设计、训练和测试奠定基础。

## 2 研究方法

### 2.1 网络整体结构

为改善 SSVEP 信号的相位、频率特征分类问题,本文将提出一种新颖的神经网络模型—PLFC-DNN。该模型设计受到迁移学习思想启发,共分成两个模块,利用它们各自的训练优势,以多角度充分挖掘、利用相位、频率特征信息,从而提高 SSVEP 信号在 BCI 系统应用中的分类性能。PLFC-DNN 模型整体结构,如图 5 所示,它主要包括 2 个模块:全局相频特征模块和局部相频特征模块,具体参数细节如表 2 所示。该网络模型采用紧凑的网络结构,首先,构建全局相频特征模块,以学习所有被试者个体间的相位、频率信息的共性特征;然后,构建局部相频特征模块,利用训练好的全局相频特征模块各个网络层中的权重和偏置参数,对局部特征模块进初始化操作;接着,通过对单个被试者数据进行训练,微调局部特征模块中的初始化参数,以实现对全局相频特征模块深层次的强化学习。需要说明的是两个模块设计的结构和层数相同,不同之处在于 dropout\_1 和 dropout\_2 两个层设置,全局相频特征模块的丢弃概率为 0.1,而局部相频特征设置为 0.7。通过 dropout 层在训练过程中随机丢弃一部分神经元,以降低模型的复杂性,并减少模型对特定神经元的依赖。本网络模型所处理的数据维度为  $60 \times 188 \times 1$ ,共包含 8 个类别。

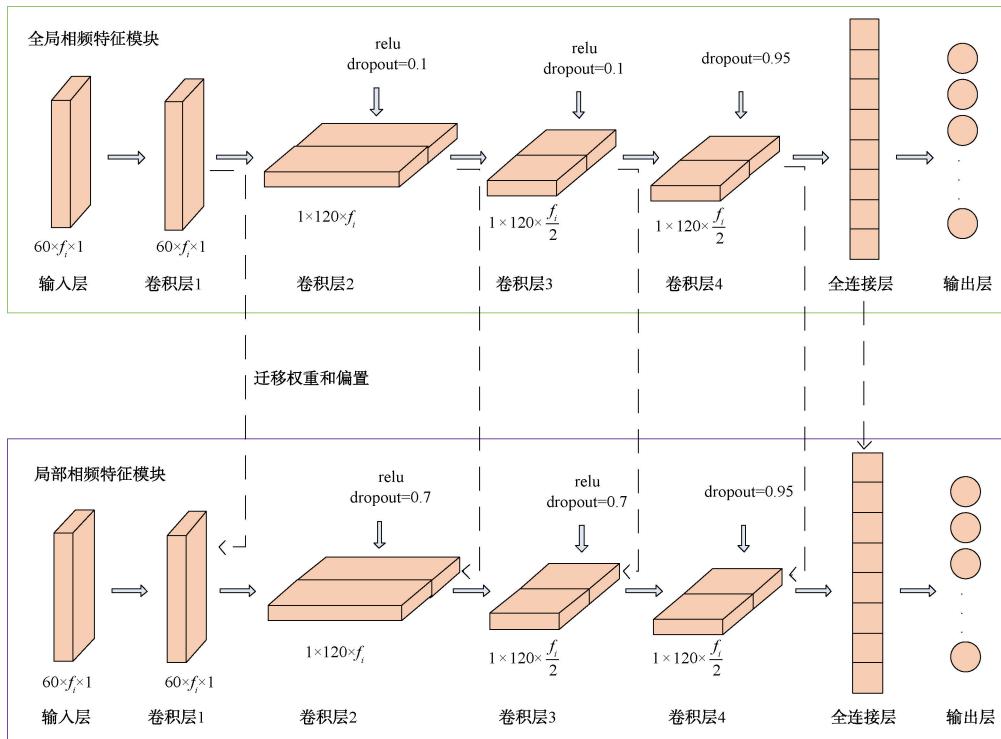
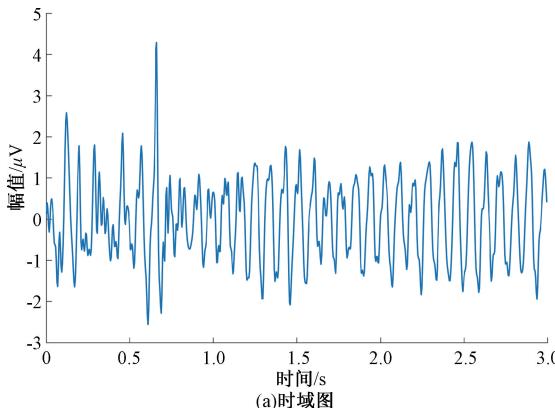


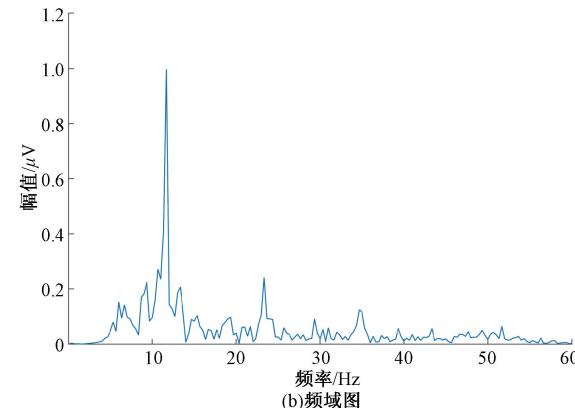
图 5 PLFC-DNN 模型整体结构图

表2 PLFC-DNN模型详细结构及参数设置

网络层类型	卷积核数量、尺寸、步长	Dropout率	激活函数
输入层	$60 \times 188 \times 1$	—	—
卷积层1	Conv_1	$1,1 \times 1,1 \times 1$	—
卷积层2	Conv_2	$120,60 \times 1,1 \times 1$	ReLU
Dropout	dropout_1	—	0.1
卷积层3	Conv_3	$120,1 \times 2,1 \times 2$	ReLU
Dropout	dropout_2	—	0.1
卷积层4	Conv_4	$120,1 \times 10,1 \times 1$	—
Dropout	dropout_3	—	0.95
FC	全连接层	8	softmax
输出层	8	—	—



(a)时域图



(b)频域图

图6 POz通道上11.6 Hz的刺激

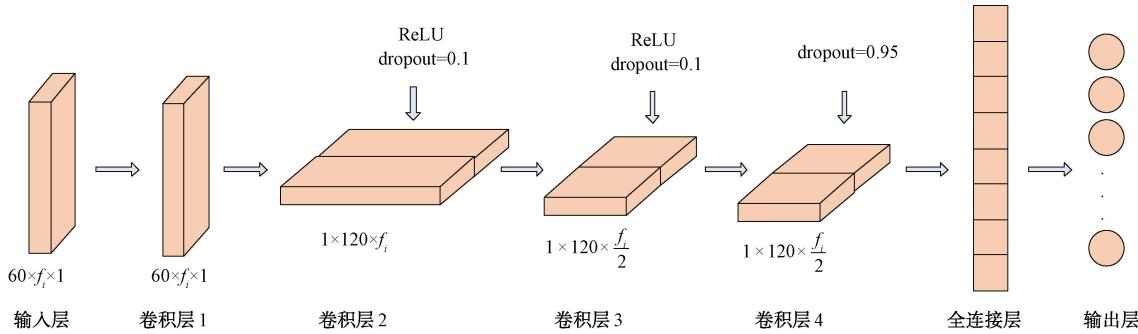


图7 全局相频特征模块

其中,全局相频特征模块每一层网络具体描述如下:

1)输入层:该层作为整个网络的输入,接收经过特征提取得到的特征矩阵,其大小为 $60 \times f_i \times 1$ 。该矩阵表示该层输入30个导联的数据,每个导联中包含 $f_i$ 个相位、频率特征。

2)卷积层1:该层使用 $1 \times 1$ 大小的卷积核,步长为 $1 \times 1$ ,对输入层的数据进行卷积操作,采用2D卷积滤波器对

## 2.2 全局相频特征模块

考虑到所使用的BETA数据集采用相位调制,这使得信号同时具备相位、频率特征。相较于时域,频域上的SSVEP信号能够展示明显的特征,如图6(b)所示。由于信号是由固定频率的视觉闪烁刺激引发,因此,在频率域中呈现出与刺激频率和倍频近似的正弦信号。相位和频率信息在解码SSVEP信号时至关重要。相位提供事件的时序信息,反映信号在时间上的变化规律,而频率信息反映信号的周期性和振荡特性。通过分析和提取相位和频率信息,可以识别出特定频率的SSVEP信号。鉴于相位、频率信息的重要性,本文首先设计全局相频特征模块,该模块包括1个输入层、4个卷积层、1个全连接层和1个输出层,如图7所示。利用1.2节中提取的相位、频率信息的3维特征矩阵作为输入,并通过一系列卷积操作学习每个导联中相位、频率的权重信息,从而学习所有被试者SSVEP信号中的相位、频率的共性特征。

不同导联的特征信息进行线性组合变换,从而得到更丰富的特征表示,并输出1个大小为 $60 \times f_i \times 1$ 的特征图。

3)卷积层2:使用120个大小为 $60 \times 1$ 、步长为 $1 \times 1$ 的2D卷积滤波器,在每个输入通道上以无填充的方式进行滑动。激活函数采用ReLU,并使用随机初始化的权重。这些卷积滤波器旨在学习上一层输入中的相位、频率特征。利用不同大小的卷积核能够从输入信号中提取到不

同类型的判别特征,并对输入数据的低、高频特征做出响应。通过该层卷积操作,将会输出 120 个大小为  $1 \times f_i \times 120$  的特征图。

4) 卷积层 3: 为增加模型灵活性,以适应不同类型、规模数据,本层采用 120 个大小为  $1 \times 2$ 、步长为  $1 \times 2$  的 2D 卷积滤波器。激活函数选择 ReLU, 并使用窄正态分布初始化权重, 以更接近于零的初始值促进模型快速收敛和性能优化。最终输出 120 个大小为  $1 \times f_i/2 \times 120$  的特征图, 用于提取输入数据的抽象特征。

5) 卷积层 4: 该层增加对输入数据的边缘填充处理, 采用边缘填充为 0 的方式, 以保持输出特征图的大小与输入特征相同。同时, 它也使用窄正态分布来初始化权重, 与上一层相同。通过这些参数的组合和设置, 该卷积层可以从输入图像中提取不同类型特征, 并输出 120 个大小为  $1 \times f_i/2 \times 120$  的特征图。

6) 全连接层: 在卷积层 4 后引入 1 个全连接层, 将卷积层中所有神经元与全连接层中所有神经元连接起来。通过线性变换, 使用权重矩阵将卷积层的输出转换为一维向量, 以建立特征与类别之间的映射关系。通过调整权重和偏置, 最小化损失函数, 从而提高特征与类别之间的映射准确性。

7) 输出层: 输出层由 8 个节点组成, 每个节点代表原始 SSVEP 数据中的一个刺激频率类别。全连接层输出与输出层进行连接, 并采用 softmax 函数作为输出层的激活函数。通过 softmax 函数, 输出层的原始数值被转化为表示概率分布的值, 从而确定输入数据属于哪个刺激频率类别的概率最高。

另外, 在本文中还引入批归一化层。在网络训练过程中, 参数变化会导致特征图分布也发生变化。这种变化可能需要降低学习速率和初始参数的大小。特别是在饱和非线性函数的情况下, 这种变化会减慢学习速度并降低系统的学习能力, 即, 内部协变量移位现象。批归一化方法被广泛用来解决这一问题。本文在 Conv\_3 和 Conv\_4 层

之后引入批归一化层, 并对每个批次的输入数据进行归一化, 以保持网络中的激活值在较小范围之内。上述操作非常有益梯度传播和算法收敛性优化。此外, 批归一化还可减少过拟合现象的发生, 提高模型的训练速度、泛化性能和鲁棒性, 有助于模型的训练和优化。

### 2.3 局部相频特征模块

为进一步提高模型的训练效果, 引入迁移学习的思想。迁移学习是将源任务中学习到的知识和经验应用于新任务, 提升新任务的学习性能。本文基于全局相频特征模块, 设计了局部相频特征模块, 两者的共同目标任务是对 SSVEP 信号中的相位和频率进行分类。使用全局相频特征模块预训练的权重和偏置作为局部相频特征模块的初始参数, 通过微调目标任务中与局部相频特征模块相关的权重和偏置进一步优化模型。在微调过程中, 保持预训练模型一部分或全部权重参数不变, 仅更新与目标任务相关的权重参数, 以更好地适应目标任务的特征。这种设计方案有助于提升模型性能, 并使其更适用于目标任务的要求。

局部相频特征模块与全局相频特征模块具有相同的网络结构, 包括 1 个输入层、4 个卷积层、1 个全连接层和 1 个输出层。但局部相频特征模块在卷积层 2 和卷积层 3 上增加了 dropout 层的丢弃概率, 从最初的 0.1 提升至 0.7。dropout 层通过随机丢弃部分神经元的输出, 减少模型复杂度, 降低模型对特定神经元的依赖性, 从而有效降低过拟合风险。

如图 8 所示, 图中展示局部相频特征模块训练的示意图。首先, 该模块使用已训练好的全局相频特征模块参数进行初始化。然后, 利用单个被试者数据微调局部相频特征模块的权重参数, 以获得每个被试者的独立模型。最后, 通过优化后的局部相频特征模块进一步提升分类性能, 从而增强整体模型的分类能力。因此, 局部相频特征模块的训练过程具有重要意义, 可为 PLFC-DNN 模型的个性化学习和分类任务提供有效的支持。

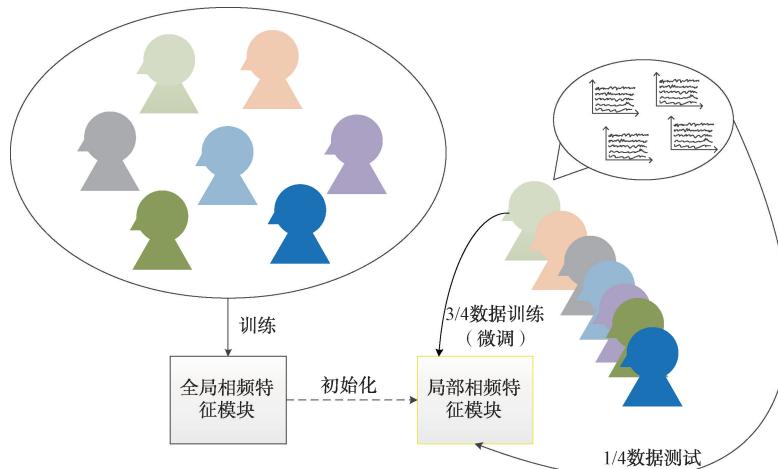


图 8 局部相频特征模块训练示意图

## 2.4 网络训练参数设置

本实验基于 MATLAB 软件进行编程, 硬件设备使用的 CPU: Intel(R) Core(TM) i5-9600KF, 内存: 16 GB, GPU: NVIDIA GeForce GTX 1650, 显存: 4 GB。其中, CCA、eCCA 和 eTRCA 模型代码开源<sup>[21]</sup>, 均在 Matlab 环境下运行。模型使用 Adam 优化器进行训练, 学习率: 0.0001。迭代批次数: 150, 批大小: 100。CUDA 版本: 11.7, cuDNN 版本: 8.4。实验的软、硬件及其相关配置如表 3 所示。

表 3 软硬件配置

硬件或软件	配置
CPU	Intel i5-9600KF
RAM	16 GB
操作系统	Windows 10
GPU	NVIDIA GeForce GTX 1650
开发环境	MATLAB2023a CUDA11.7+cuDNN8.4

## 3 实验结果与分析

### 3.1 评估指标

混淆矩阵是评估分类模型性能的重要工具, 用于计算多种评价指标。它包括: 准确率、精确率、召回率和  $F_1$  分数。准确率和精确率是常用的分类模型评估指标, 用于衡量正样本预测的准确程度和整体预测的准确性。准确率表示预测正确样本占总样本数比例, 用于评估模型对整体样本的准确预测能力。精确率衡量模型在所有被预测为正类样本中, 实际为正类样本的概率。上述两个指标的计算公式如下所示:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$P = \frac{TP}{TP + FP} \quad (2)$$

其中,  $Acc$  为准确率,  $P$  为精确率,  $TP$  代表将正样本判断为正,  $TN$  代表将负样本判断为负,  $FP$  代表将负样本判断为正,  $FN$  代表将正样本判断为负。

如果存在不平衡样本情况, 仅仅依赖准确率评估模型是不可行的。因为, 当正类样本数量明显少于负类样本时, 即使将所有样本都预测为负类, 也能获得较高准确率。然而, 这样的模型无法有效地检测到正样本, 因此, 考虑召回率指标是必不可少的。召回率是衡量实际为正样本中被正确预测为正的概率, 体现对于少数类样本的捕捉能力。在脑电特征分类问题中, 召回率反映对脑电特征的有效检测能力。因此, 在面对不平衡样本问题时, 召回率是一个重要的评估指标, 其计算公式如下所示:

$$R = \frac{TP}{TP + FN} \quad (3)$$

其中,  $R$  为召回率,  $TP$  代表将正样本判断为正,  $FN$  代表将正样本判断为负。

$F_1$  分数代表精确率和召回率的调和平均值, 用于综合评估模型性能。它的取值范围在 0~1 之间, 接近 1 的分数表示模型性能较好。因此,  $F_1$  分数是一项重要的综合评估指标, 能客观、全面地衡量模型的准确率和召回率, 其计算公式如下所示:

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (4)$$

在针对多分类模型时, 传统的精确率、召回率和  $F_1$  分数计算公式仅适用于二分类情况。因此, 在多分类模型中, 准确率、召回率、精确率和  $F_1$  分数的定义和计算方式略有不同。本文使用 Macro average 规则对多分类模型结果进行评估, 它将每个类别的评估指标进行算术平均, 以得到整体的评估指标。这样能更准确地评估模型在多分类情况下的相关性能, 不需要考虑类别之间的样本数量差异, 而是平等对待每个类别。其多分类模型的准确率、召回率、精确率和  $F_1$  分数计算公式如下:

$$P_{Macro} = \frac{1}{n} \sum_{i=1}^n P_i \quad (5)$$

$$R_{Macro} = \frac{1}{n} \sum_{i=1}^n R_i \quad (6)$$

$$F_{Macro} = \frac{1}{n} \sum_{i=1}^n F_i \quad (7)$$

$$F_{Macro} = \frac{2 \times P_{Macro} \times R_{Macro}}{P_{Macro} + R_{Macro}} \quad (8)$$

其中,  $P_{Macro}$  为精确率,  $R_{Macro}$  为召回率,  $F_{Macro}$  为  $F_1$  分数。本文使用 5 种不同时间窗长度的 SSVEP 数据对模型进行验证, 根据不同时间窗长度, 生成的样本数量分别包含: 8 800、7 040、5 280、3 520 和 1 760 个, 其中, 每种时间窗长度样本中都包含 8 个类别, 使用刺激目标的相位和频率作为类别标签。不同时间窗长度的训练中, 均选取 3 个 block 数据作为训练集, 1 个 block 数据作为测试集, 并采用留一法对 PLFC-DNN 模型进行交叉验证, 模型进行 150 次迭代, 在最终预测结果中: 使用 1.5 s 时间窗长度样本的准确率达到 91.06%, 精确率: 91.09%, 召回率: 91.06%,  $F_1$  分数: 0.9105。准确率、精确率和召回率相对接近, 这表明模型在多被试者、跨试次以及多类别数据集上表现良好。 $F_1$  分数高表明模型在平衡精确率和召回率方面表现出色。综合以上评估参数的具体情况, 可得出如下结论: PLFC-DNN 模型的整体性能稳定, 无明显的类别偏好, 适用于多被试者、跨试次及多类别分类等任务。

混淆矩阵结果经可视化, 如图 9 所示, 其中, 列代表 SSVEP 信号预测的类别, 行代表 SSVEP 信号的真实类别, 对角线上的数值表示正确预测的样本数量。混淆矩阵对角线上的数值越大, 即, 模型正确预测样本数量越多, 模型的分类效果也越好。实验结果表明: 本文提出的 PLFC-DNN

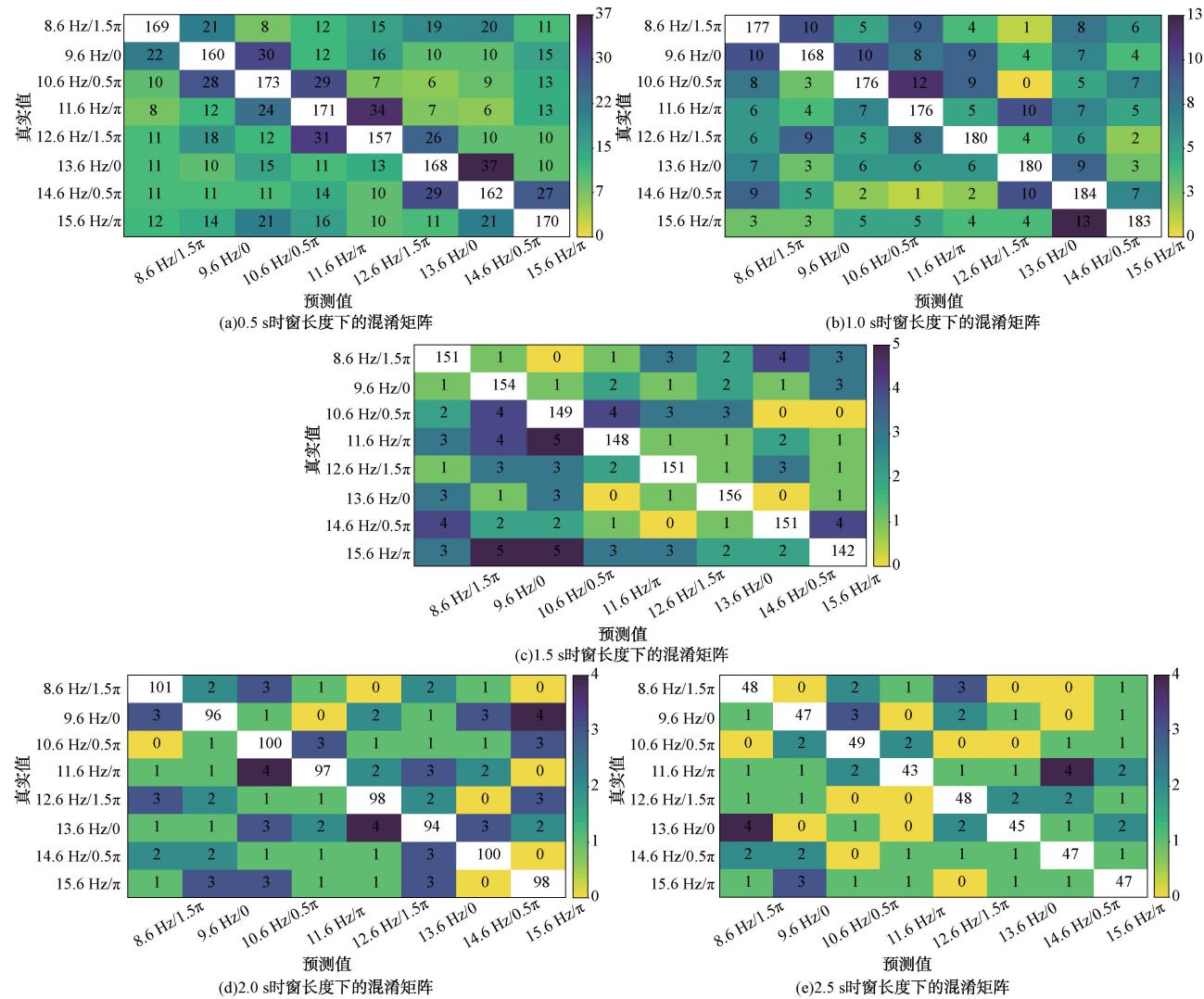


图 9 PLFC-DNN 模型在不同时间窗长度下的混淆矩阵

模型能显著降低 SSVEP 信号中相位和频率特征的误分  
类率。

### 3.2 不同算法对比

实验使用公开数据集 BETA, 通过留一法 (leave-one-subject-out) 交叉验证比较跨试次、跨被试者分类结果。首先, 以所有被试者一个试次的 SSVEP 数据作为测试, 而其他试次的数据用来训练分类器, 反复循环该过程, 直到每一个试次都被测试到; 然后, 使用单个被试者一个试次的 SSVEP 数据进行测试, 微调分类器参数, 再一次重复循环, 直到每名被试者的每一个试次都被测试一次。

下面将本文提出的 PLFC-DNN 与 CCA、eCCA 和 eTRCA 等传统方法进行分类效果比较。为全面评估模型性能, 所有方法均使用相同的数据预处理条件, 并分别对比 5 种不同时间窗长度分段情况下的分类效果。同时, 还使用标准误 (standard error, SE) 对单一时间窗长度划分的数据集进行量化比较。在评估 SSVEP 信号的分类性能方

面, 采用分类准确率和信息传输率 (information transfer rate, ITR) 作为评估指标, 其计算公式如式(9)、(10)所示。分类准确率是衡量模型对不同类别样本的准确识别程度指标, 它通过计算正确分类样本数量与总样本数量的比例得到。ITR 是衡量模型性能的重要指标, 它综合考虑分类准确率、类别数量和单次目标选择时间, 可以比较不同算法的性能, 而不考虑它们的内部性能或优化指标, 这样的评估能够更全面地反映出模型的实际表现。ITR 计算公式如下:

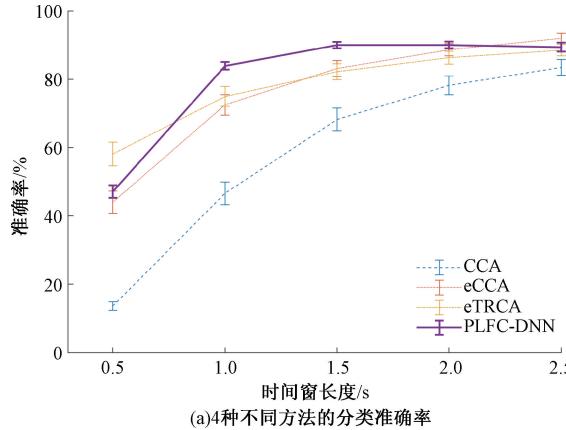
$$B = \log_2 N + p \times \log_2 p + (1 - p) \times \frac{1 - p}{N - 1} \quad (9)$$

$$ITR = B \times \frac{60}{T} \quad (10)$$

式中:  $N$  为分类目标数,  $p$  为分类准确率,  $T$  为单次目标选择时间。

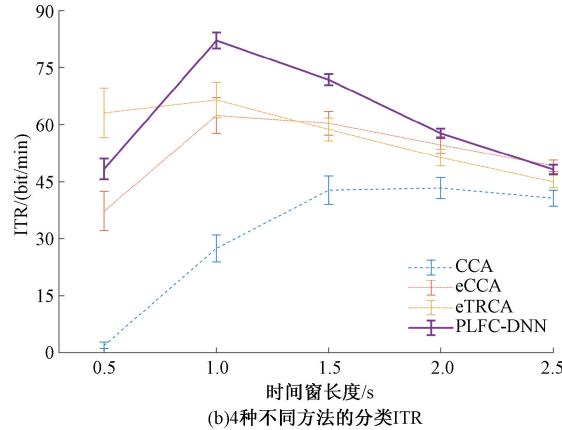
如图 10 所示, 4 种不同方法在不同时间窗口长度数据集上的平均分类准确率和平均 ITR。考虑到 SSVEP 信号

的非线性和不平稳特性,不同被试者对相同视觉刺激频率反应的起始时间和反应时长存在差异。这种差异性会导致模型在不同被试者上的识别效果差异较大,从而降低系统稳定性。因此,为全面评估模型在不同时间尺度上的性能,本文选择 5 个时间窗口长度(0.5、1.0、1.5、2.0 和 2.5 s),以 0.5 s 的步长对数据集进行划分,并使用留一法



(a)4种不同方法的分类准确率

对数据集进行验证。实验结果显示:在不同时间窗口长度数据的实验中,PLFC-DNN 模型在图示的中间部分表现优于其他方法,平均分类准确率最高可达 89.98%,平均 ITR 最高可达 82.10 bit/min。结合结果分析,在 0.5 s 的时间窗口下,较短的窗口长度会导致提取到的特征信息较少,从而导致识别准确率有所下降。



(b)4种不同方法的分类 ITR

图 10 不同方法在不同时间窗长度上的实验结果

如表 4 所示,针对 1.5 s 时间窗长度数据,PLFC-DNN 模型表现优于其他方法,平均分类准确率达到 89.98%,平均 ITR 为 71.80 bit/min。相较于其他算法,本文从全局和局部两个不同维度设计 SSVEP 信号分类模型,深化其内在的学习结构,取得较好效果。PLFC-DNN 能更全面地挖掘、利用数据集中的相位、频率特征信息。此外,PLFC-DNN 模型的卷积层中每个神经元都与一个滤波器相关联。这些滤波器通过在训练数据的局部区域滑动操作,进行卷积运算以提取相位、频率特征。值得注意的是,卷积层中的神经元参数具有共享性质,即同一个滤波器会在输入的不同位置上进行应用。这种参数共享机制使得滤波器更倾向于在不同输入数据中寻找相同的特征信息,从而提高模型的泛化能力,使得模型的分类准确率和 ITR 表现更加优异。

表 4 不同方法数据集(时间窗口长度为 1.5 s)上的平均实验结果

方法	Acc/%	ITR/(bit/min)
CCA	68.24±3.34	42.75±3.76
eCCA	83.11±2.31	60.34±3.13
eTRCA	82.17±2.34	58.71±3.05
<b>PLFC-DNN</b>	<b>89.98±0.96</b>	<b>71.80±1.46</b>

### 3.3 消融实验

为进一步验证 PLFC-DNN 模型设计的合理性,将模型的两个模块分别进行性能对比,并使用 1.5 s 时间窗长度数据进行消融实验。具体来说,分别使用全局相频特性模块、局部相频特性模块及二者相结合的有关模型进行训

练、测试,比较它们的分类准确率和 ITR。同时,消融实验也进一步证明本文所选用设计方式能让两个模块更好地发挥不同角度训练的优势,以获得更好的模型整体性能和训练效果。

如表 5 所示,仅使用全局相频特性模块进行训练时,分类准确率为 83.83%;只使用局部相频特性模块进行训练时,分类准确率仅为 22.94%。两者差距显著,相差高达 60.89%。与之对应,当将两者结合使用时,分类准确率可达 89.98%。相较于仅使用全局相频特性模块和局部相频特性模块时,平均准确率分别提升 6.06% 和 67.04%,平均 ITR 分别提升 14.55 bits/min 和 68.39 bits/min。这些结果表明:局部相频特性模块对全局相频特性模块起到强化的作用,通过将二者结合使用有利于更加充分挖掘、利用 SSVEP 中蕴含的相位、频率特征信息,从而提升 PLFC-DNN 对 SSVEP 信号特征的分类性能,该实验结果进一步表明使用多维度角学习在脑电信号分类中的重要性。

表 5 不同训练模块(时间窗口长度为 1.5 s)上的平均实验结果

训练模块	准确率/%	ITR/(bits/min)
仅用全局相频特征模块	83.83±0.76	57.25±1.18
仅用局部相频特征模块	22.94±0.82	3.41±0.46
使用全局相频特征模块+局部相频特征模块	<b>89.98±0.96</b>	<b>71.80±1.46</b>

## 4 结 论

本文提出一种基于迁移学习思想的深度学习模型—

PLFC-DNN, 用于从多通道、多被试者中识别 SSVEP 信号的相位、频率特征。该模型采用紧凑网络结构, 并结合迁移学习技术, 实现对每一样本相位、频率特征的强化学习。实验采用 BETA 数据集, 结果表明: PLFC-DNN 模型的分类准确率比 CCA、eCCA 和 eTRCA 等主流算法分别提升了 21.74%、6.87%、7.81%, 信息传输率各提升 29.05、11.46、13.09 bit/min。PLFC-DNN 对 SSVEP 信号的相位、频率特征信息实现了更进一步的挖掘、利用, 模型轻量、高效。在未来工作中将继续探索多被试者、跨试次和多分类情况下的 SSVEP 信号分类模型, 以提高 BCI 系统的实际应用价值。

## 参考文献

- [1] WAYTOWICH N R, KRUSIENSKI D J. Multiclass steady-state visual evoked potential frequency evaluation using chirp-modulated stimuli [J]. IEEE Transactions on Human-Machine Systems, 2016, 46(4): 593-600.
- [2] ZHAO D, WANG T, TIAN Y, et al. Filter bank convolutional neural network for SSVEP classification [J]. IEEE Access, 2021, 9: 147129-147141.
- [3] 林艳飞, 瞿博宇, 郭嵘晓, 等. 基于相频特性的稳态视觉诱发电位深度学习分类模型 [J]. 电子与信息学报, 2022, 44(2): 446-454.
- [4] CHEN X, WANG Y, NAKANISHI M, et al. High-speed spelling with a noninvasive brain-computer interface [J]. Proceedings of the National Academy of Sciences, 2015, 112(44): E6058-E6067.
- [5] NAKANISHI M, WANG Y, CHEN X, et al. Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis [J]. IEEE Transactions on Biomedical Engineering, 2017, 65(1): 104-112.
- [6] LI J X, DAI F ZH, YIN D, et al. A method of SSVEP signal identification based on improved eCAA [J]. Instrumentation, 2024(4): 1-11.
- [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] DEVLIN J, CHANG M W, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding [J]. ArXiv Preprint, 2018, ArXiv: 1810.04805.
- [9] BERG R, KIPF T N, WELLING M. Graph convolutional matrix completion [J]. ArXiv Preprint, 2017, ArXiv: 1706.02263.
- [10] FAUST O, HAGIWARA Y, HONG T J, et al. Deep learning for healthcare applications based on physiological signals: A review [J]. Computer Methods and Programs in Biomedicine, 2018, 161: 1-13.
- [11] CRAIK A, HE Y, CONTRERAS-VIDAL J L. Deep learning for electroencephalogram(EEG) classification tasks: a review [J]. Journal of Neural Engineering, 2019, 16(3): 031001.
- [12] ROY Y, BANVILLE H, ALBUQUERQUE I, et al. Deep learning-based electroencephalography analysis: A systematic review [J]. Journal of Neural Engineering, 2019, 16(5): 051001.
- [13] 韩向可, 郭士杰. 基于 SSVEP-SSA 融合的混合脑机接口研究 [J]. 仪器仪表学报, 2019, 40(5): 213-220.
- [14] KEERTHI KRISHNAN K, SOMAN K P. CNN based classification of motor imaginary using variational mode decomposed EEG-spectrum image [J]. Biomedical Engineering Letters, 2021, 11(3): 235-247.
- [15] ZHANG M, LIU D, WANG Q, et al. Detection of alertness-related EEG signals based on decision fused BP neural network [J]. Biomedical Signal Processing and Control, 2022, 74: 103479.
- [16] 谢平, 胡锦城, 江国乾, 等. 基于多任务学习的快速序列视觉呈现脑电图分类 [J]. 仪器仪表学报, 2023, 44(11): 215-223.
- [17] PAN Y, CHEN J, ZHANG Y. A survey of deep learning-based classification methods for steady-state visual evoked potentials [J]. Brain-Apparatus Communication: A Journal of Bacomics, 2023, 2(1): 2181102.
- [18] LI J, WU W, XUE D, et al. Multi-source deep transfer neural network algorithm [J]. Sensors, 2019, 19(18): 3992.
- [19] GUNAY O B, OBLOKULOV M, OZKAN H. A deep neural network for ssvep-based brain-computer interfaces [J]. IEEE Transactions on Biomedical Engineering, 2021, 69(2): 932-944.
- [20] LIU B, HUANG X, WANG Y, et al. BETA: A large benchmark database toward SSVEP-BCI application [J]. Frontiers in Neuroscience, 2020, 14: 627.
- [21] WONG C M, WAN F, WANG B, et al. Learning across multi-stimulus enhances target recognition methods in SSVEP-based BCIs [J]. Journal of Neural Engineering, 2020, 17(1): 016026.

## 作者简介

丛佩超, 博士, 副教授, 硕士生导师, 主要研究方向为移动式机器人的动力学建模与仿真、智能化移动机器人的自主导航与控制问题、工业机械灵巧手的遥操作、空间机器人的相关问题等。

E-mail: cpclzx2022@gxust.edu.cn

陈熙来(通信作者), 硕士研究生, 主要研究方向为下肢外骨骼机器人、脑机接口。

E-mail: xilai\_chen@163.com