

用于稳态视觉诱发电位目标识别的多尺度特征融合卷积神经网络方法

胡勤伟¹, 陶庆¹, 王妮妮¹, 陈清正¹, 吴腾辉¹, 张小栋²

(1. 新疆大学机械工程学院, 830047, 乌鲁木齐; 2. 西安交通大学机械工程学院, 710049, 西安)

摘要: 针对传统稳态视觉诱发电位(SSVEP)脑电信号目标识别方法分类精度低、提取特征不充分、方法复杂且耗时等问题,提出一种基于多尺度特征融合卷积神经网络的SSVEP信号分类识别方法(SSVEP-MF)。利用小波变换将多通道SSVEP信号整合转化为二维图像作为输入样本集;建立多尺度特征融合卷积神经网络模型(MFCNN),该模型利用三层二维卷积核实现图像样本不同尺度特征的充分提取,构建多尺度特征融合单元对不同层级特征进行融合,并通过全连接等操作完成模型的训练;将样本集输入到MFCNN模型中实现脑电信号特征自适应提取及端到端分类。所提SSVEP-MF方法能够充分提取信号各层级特征,实现短时间视觉刺激下SSVEP信号的有效识别,并具有较高的目标识别效率。实验结果表明,在1 s刺激时长时,相比传统功率谱密度分析方法、典型相关分析方法以及普通卷积结构方法,所提方法的识别准确率分别提升了18.57%、20.08%及7.03%,有效提高了基于稳态视觉诱发电位范式下脑机接口的信号识别性能。

关键词: 稳态视觉诱发电位;目标识别;多尺度特征融合;卷积神经网络;小波变换;

脑机接口

中图分类号: TP23 文献标志码: A

DOI: 10.7652/xjtub202204020 文章编号: 0253-987X(2022)04-0185-10



OSID 码

A Multi-Scale Feature Fusion Convolutional Neural Network Approach for Steady-State Visual Evoked Potential Target Recognition

HU Qinwei¹, TAO Qing¹, WANG Nini¹, CHEN Qingzheng¹, WU Tenghui¹, ZHANG Xiaodong²

(1. School of Mechanical Engineering, Xinjiang University, Urumqi 830047, China;

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

Abstract: A multi-scale feature fusion convolutional neural network-based SSVEP signal classification and recognition method is proposed to solve the problems of low classification accuracy, inadequate feature extraction, complex and time-consuming methods of traditional steady-state visual evoked potential (SSVEP-MF) signal target recognition methods. Firstly, the wavelet transform is used to integrate the multi-channel SSVEP signals into two-dimensional images as the input sample set; secondly, a multi-scale feature fusion convolutional neural network model (MFCNN) is established, which uses a three-layer two-dimensional convolutional kernel to achieve sufficient extraction of features at different scales of image samples, constructs multi-scale feature fusion units to fuse features at different levels, and completes the training of the model through operations such as full connectivity; finally, the sample set is input to the

MFCNN model to achieve adaptive extraction of EEG signal features and end-to-end classification. The proposed SSVEP-MF method can fully extract the features at each level of the signal, achieve effective recognition of SSVEP signals under short-time visual stimulation, and have high target recognition efficiency. The experimental results show that the recognition accuracy of the proposed method is improved by 18.57%, 20.08% and 7.03%, respectively, compared with the traditional power spectral density analysis method, typical correlation analysis method and common convolutional structure method at 1 s stimulus duration, which effectively improves the signal recognition performance of brain-machine interface based on the steady-state visual evoked potential paradigm.

Keywords: steady-state visual evoked potentials; target recognition; multi-scale feature fusion; convolutional neural network; wavelet transform; brain-computer interface

脑-机接口(BCI)是构建大脑与外部信息交流的一种技术方式,通过采集神经系统的电活动信号,将其分类到特定指令的不同类别中,来实现解码用户意图并与计算机等外部设备进行通信的人机接口技术^[1]。近年来,BCI技术在人机交互、智能感知、医疗康复等领域得到了广泛的应用^[2-4]。在脑电波采集过程中,人的大脑活动会受到不同类型刺激的影响而产生预期的信号反应,稳态视觉诱发电位(SSVEP)是脑机接口中被广泛应用的外源性刺激诱发信号类型之一,它是外界周期性视觉刺激诱发大脑视觉中枢系统产生的周期性响应^[5],在大脑中产生由视网膜接收到的 3.5 Hz~75 Hz 范围相同频率或倍数频率的电生理活动^[6],具有信噪比高、无需训练、稳定性高等优势。

由于脑电信号微弱,易受到外界环境因素的影响,以及其他生物电信号等背景噪声的干扰,从而导致 SSVEP 信号识别困难^[7]。如何进一步提高目标识别效率是 SSVEP-BCI 系统中的一个重要问题,在算法研究方面,特征提取算法优化以及分类器改进一直是 SSVEP 信号分类研究的重点。在传统 SSVEP 信号目标识别方法中,通常是以人工提取信号的特征信息,再对特征向量进行有监督的分类方式来进行识别^[8]。文献[9]结合经验模式分解(EMD)和功率谱峰值分析(PSPA)方法,以基频功率谱峰的突出度和宽度的最大比值与从窄频带分量中计算的谐波来对 SSVEP 信号进行分类,但是分解后的各基本模式间并非严格正交,存在混叠现象。文献[10]利用典型相关分析(CCA)方法,运用结合叠加平均和快速傅里叶变换,提取与多通道 SSVEP 信号相关的频率信息进行相关分析,仅提高了低频率刺激下的识别精度,存在局限性。文献[11]利用小波包分解(WPD)对脑电信号进行特征提取,结合

时域和频域中的特征信息,进一步提高了信号识别准确率,但是小波基函数的选择多依赖于经验,结果随函数影响较大。

目前,深度学习方法在近期的许多研究中被成功地应用于学习特征和分类不同类型的数据,然而在 BCI 的应用中采用深度学习方法的研究数量较为有限^[12]。卷积神经网络(CNN)是一种通过多个滤波器及其特殊卷积结构进行信号特征提取、细化的深度学习方法,所得特征在一定程度上具有平移、翻转及缩放不变性,相较于传统算法具有分类精度高且能够自动学习特征和解码等优势,有效避免了人工特征提取过程中存在的信息丢失等不确定性问题,有望实现更高的目标识别性能。文献[13]提出一种 CNN 与多层感知器(MLP)并行的网络结构,分别处理静态能量和动态能量,并且取得了较高识别准确率。文献[14]构建了一种时频 CNN 模型运用到 SSVEP 信号分类中,由两个空域和时域滤波器卷积层以及快速傅里叶变换组成,将卷积层抽象得到的特征转换至频域特征进行分类。基于该研究,文献[15]提出以 CNN 与支持向量机(SVM)组合分类器的方法,进一步提高了网络对脑电信号的识别正确率。文献[16]构建了一种多尺度卷积核的 CNN 脑电分类模型,通过不同大小的卷积核来增加特征提取时的视野维度,提高了分类精度,但是由于数据转换成的是二维数字矩阵,并不像图像数据那样具有丰富的信息,因此存在一定的局限性。

上述基于 CNN 的分类识别方法虽然均能达到较好的识别效果,但是若只考虑单一尺度或单一级下的特征,将会导致特征提取过程中存在特征缺失或冗余等问题,从而限制了识别精度。因此,本文提出一种基于多尺度特征融合卷积神经网络(MF-CNN)的信号分类识别方法,利用卷积神经网络在

图像处理中的强大优势,将 SSVEP 信号数据转化为二维时频图像作为输入,基于 MFCNN 模型框架有效提取目标信号多尺度级联特征,实现 SSVEP 信号目标自适应特征提取及端到端识别。

1 基于多尺度特征融合卷积神经网络的目标识别方法

1.1 卷积神经网络

CNN 的网络结构一般由输入层、交替叠加的卷积层与池化层、全连接层和分类器组成^[17]。卷积层将对输入样本的局部区域和滤波器内核之间进行内积运算,每个滤波器都通过相同的内核来获得输入信号的局部特征。设前层图像大小为 x_l ,则卷积过程描述如下

$$y_{l+1,m} = w_{l,m} \otimes x_l + b_{l,m} \tag{1}$$

式中: $w_{l,m}$ 和 $b_{l,m}$ 分别为第 l 层中第 m 个滤波器内核的权值矩阵和偏置项; $y_{l+1,m}$ 为第 $l+1$ 层中第 m 个滤波器卷积后的输出; $\otimes$ 表示卷积计算。

卷积之后,利用激活函数来获得非线性特征,本文采用 LeakyReLU 激活函数加速 CNN 的收敛过程,该函数在 ReLU 基础上保留部分小于零的输入,激活后的输出值 $a_{l+1,m}$ 表达式如下

$$a_{l+1,m} = f(y_{l+1,m}) = \max(0, y_{l+1,m}) + Z_{\text{leak}} \min(0, y_{l+1,m}) \tag{2}$$

式中: $a_{l+1,m}$ 为 $y_{l+1,m}$ 激活后的输出值; Z_{leak} 为负值区域的非零斜率。

池化层将提取到的特征降维简化,减少卷积后的特征空间和网络参数,本文采用最大池化来降低数据维度和网络复杂度,输出采样区域的最大值,如下式

$$p_{l+1,m} = \max_{(n-1)w+1 \leq t \leq nw} \{q_{l,m}(t)\} \tag{3}$$

式中: $q_{l,m}(t)$ 为第 l 层中第 m 个帧中第 t 个神经元的值, $t \in [(n-1)w+1, nw]$; $p_{l+1,m}$ 为池化操作后第 $l+1$ 层中神经元对应的值; w 为池化区域的宽度。

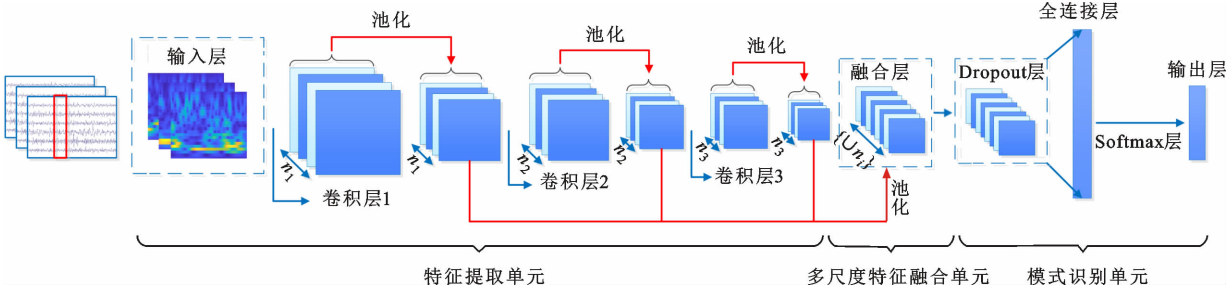


图 1 多尺度特征融合卷积神经网络目标识别模型

Fig. 1 MFCNN target recognition model

1.2 时频化处理

由于 SSVEP 信号的非平稳、非线性、时变特性^[18],传统的时域或者频域分析方法具有信号处理的不确定性,谱分析无法反映各个频段能量随时间的变化情况,也不能完整地表达信号的特征参数。小波变换在传统傅里叶变换基础上,引入尺度和时间变量来分析信号的不同频率成分,同时提供时域和频域的信息特征^[19],适合处理短时间内的突变信号。因此,本文利用小波变换对 SSVEP 信号进行时频化图像处理,将时频域特征生成时频图,以不同颜色分布表征不同类别的信号差异,如下式

$$WT_f(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(\tau) \psi\left(\frac{t-\tau}{a}\right) d\tau \tag{4}$$

式中: a 为尺度参数, $a > 0$, $1/\sqrt{a}$ 为保持能量守恒的规范因子; t 为时间平移量。

1.3 多尺度特征融合卷积神经网络

小波变换将 SSVEP 信号转化成频率-时间-小波系数图(时频图),不同颜色分布的图像作为输入,利用 CNN 进一步提取图像特征进行学习并分类。

CNN 利用浅层单元能够从图像中提取边缘、颜色等细节特征,该类特征分辨率较高,但是伴随较多噪声,随着卷积层的加深,网络能在浅层特征的基础上进一步提取图像的深层全局语义特征,但是同步增加的卷积等操作也导致模型过饱和、信息丢失等问题。因此,本文考虑利用从不同层次的特征来构建 MFCNN 网络模型,以获取图像各层级不同尺度特征加强网络分类效果。模型主要由特征提取、多尺度特征融合以及模式识别 3 个单元组成,整体结构如图 1 所示。

特征提取单元将 SSVEP 信号分段利用小波变换构造时频化图像输入,图像大小为 32 像素 $\times$ 32 像素,样本图像依次经过 3 个卷积层。第 1 个卷积层的卷积核大小为 5 $\times$ 5、步长为 2,得到输出特征图大小为 16 $\times$ 16 $\times$ n , n 为卷积核数,卷积时进行边界补零填充,避免提取过程中的未知信息丢失^[20];第 2

个卷积层的卷积核大小为 5×5 ,步长为 1;第 3 个卷积层的卷积核大小为 3×3 ,步长为 1。每个卷积层后都跟有一个激活函数层和池化层,同时加入批量归一化层,防止训练过程出现过拟合现象,最大池化层降低数据维度并保留主要的特征信息,池化核大小为 2×2 ,步长为 2。

多尺度特征融合单元保证网络同步提取细节及全局特征,该部分以特征提取单元的输出作为输入,在保证特征多样化表达的同时进行融合,其中设计池化层来统一参与融合特征向量的维度,第 1 层级池化核大小为 4×4 ,步长为 4;第 2 层级池化核大小为 2×2 ,步长为 2;第 3 层级池化核大小为 2×2 ,步长为 1,最后输入到融合层中进行特征衔接,得到融合特征,表达式为

$$\boldsymbol{F} = \{\cup \boldsymbol{F}_i\}$$

(5)

式中: $\boldsymbol{F}$ 为融合特征; $\boldsymbol{F}_i$ 为第 i 级特征。设需融合的两个向量分别为 $\boldsymbol{F}_1(x,y,z)$ 和 $\boldsymbol{F}_2(m,p,q)$, $x=m$, $y=p$,则融合后得到特征向量 $\boldsymbol{F}=(x,y,z+q)$ 。

模式识别单元将融合特征作为输入进行识别,全连接层与 Softmax 分类器输出 N 分类的识别准确率,表达式为

$$A_j(z) = \frac{e^{z_j}}{\sum_N e^{z_N}}$$

(6)

式中: z_j 为第 j 个类别的线性概率; z_N 为 N 个类别的线性概率之和; $A_j(z)$ 为类别归一化结果。

1.4 网络模型的参数选择

MFCNN 模型的前向传播与 CNN 相同,采用误差的反向传播算法加以训练,利用前向传播特征传递,对训练数据在网络中前向输出值与期望输出值之间进行误差计算,并反向传播来修正各层的网络权值和偏置^[16]。本文网络训练采取了小批量训练,权值的优化调整采用随机梯度下降(SGD)优化器,通过最小化损失来调整训练参数,以优化网络模型各环节并判断模型是否训练终止,损失越小,网络拟合效果越好。在训练中加入 L2 正则化提高泛化能力^[21],防止网络过拟合。针对数据集统计模型训练参数,并从输入图像大小和融合特征层网络结构进行分析,重复实验记录平均识别准确率及训练时长,对比得到 MFCNN 针对该数据集时的训练参数如表 1 所示。

1.5 目标识别模型

基于 MFCNN 的 SSVEP 目标识别方法对应的识别流程如图 2 所示。首先,采集受试者在不同閃

表 1 MFCNN 针对该数据集时的训练参数

Table 1 MFCNN training parameters table

训练参数	数值设定
学习率	0.02
迭代轮次	20
批尺寸	32
衰减周期	6
衰减因子	0.999
L2 正则化	0.000 1

烁目标刺激下的多通道 SSVEP 信号,经过预处理整合后输入至网络模型中,利用小波变换构造 SSVEP 信号为二维时频图像样本集,随机选取数据集的 70% 为训练集对 MFCNN 模型进行训练,30% 为测试集,之后通过 MFCNN 中特征提取单元、多尺度特征融合单元进行信号特征的提取与融合,并输入至分类识别单元计算模型最终解码性能,实现 SSVEP 信号自适应特征学习及端到端识别。

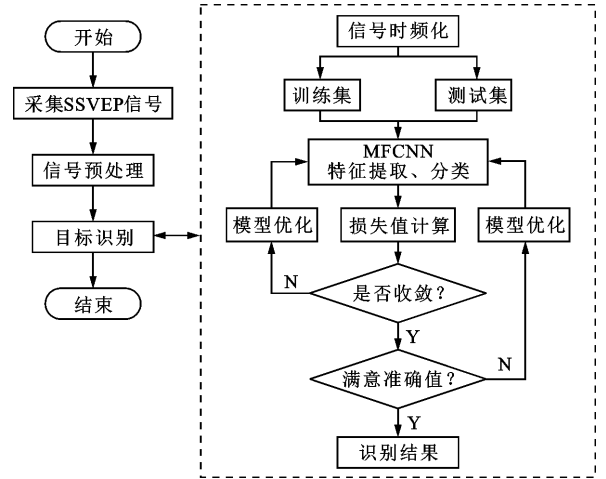


图 2 基于 MFCNN 的 SSVEP 识别流程

Fig. 2 SSVEP identification flow chart based on MFCNN

2 稳态视觉诱发电位实验分析

2.1 实验范式设计

为了验证本文方法对 SSVEP-BCI 目标识别的实际效果,实验范式设计引用基于场景动画的稳态视觉刺激诱发范式^[22],该范式将生活情景划分为相应的独立刺激场景图,经过灰度标准化处理后再通过方波调制模式对黑白反转色图片进行视觉刺激。刺激范式界面的设计和呈现由 PsychoPy3 软件实现,场景动画诱发范式如图 3 所示,实验所使用的显示器分辨率为 $1\,920 \text{ 像素} \times 1\,080 \text{ 像素}$,刷新频率为 60 Hz,屏幕背景色为深灰色,共有 4 张刺激图片分

别以上、右、下、左的顺序对应于 6 Hz、7.5 Hz、8.57 Hz、10 Hz 的频率进行闪烁以诱发 SSVEP 信号。

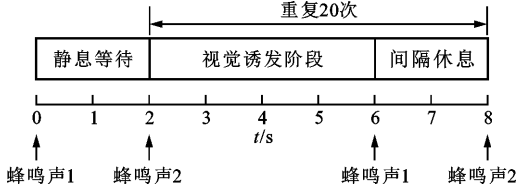


图 3 场景动画诱发范式示意图

Fig. 3 Schematic diagram of the scene animation evocation paradigm

2.2 实验流程

实验招募 7 名视力正常或矫正视力正常的健康受试者(S1~S7)作为受试对象,包括 4 名男性,3 名女性,年龄均在 22~26 岁之间。实验处于安静的房间内进行,并要求受试对象保持舒适坐姿,与显示器的距离约为 70 cm。由于视觉皮层中央放大效应,受试者需将视觉注意力注视点集中在特定的目标上,尽量避免肢体的移动和眼部的运动。每名受试者按照流程顺序依次注视闪烁频率为 6 Hz、7.5 Hz、8.57 Hz、10 Hz 的刺激目标,对应标签 1、2、3、4。图 4(a)给出了每轮实验的时序流程安排,图 4(b)给出了实验场景。实验开始时,受试者有 2 s 的



(a)实验时序图



(b)实验场景图

图 4 实验时序图与场景图

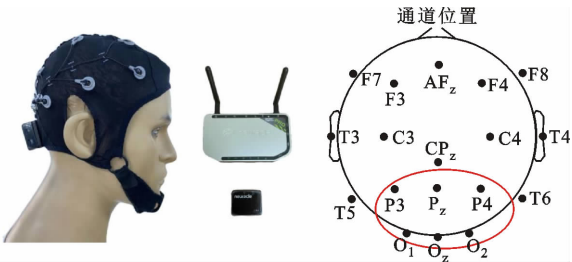
Fig. 4 Experimental timing diagram and scenario diagram

静息准备时间,当 PC 机提示音响起,正式开始闪烁,重复 20 次为一轮,单次刺激目标的呈现时间为 4 s,随后提示音响起闪烁停止,受试者间隔休息 2 s 后再次进行下一次闪烁,每个刺激目标呈现 3 轮,轮与轮之间也给予受试者充足的休息时间。

2.3 信号采集与预处理

实验采用图 5(a)所示 Neuracle 公司生产的 NeuSen W16 无线脑电信号采集系统完成受试者信号的采集,采样频率为1 000 Hz,电极分布按照国际标准 10/20 电极安放法放置,视觉刺激通过视觉通路传递到大脑皮层,处在枕叶区的视觉皮层负责处理视觉信号的感知等一系列任务,故实验采集枕叶区 P3、Pz、P4、O1、Oz、O2 共 6 个通道的信号,接地电极放置在 AFz 处,参考电极放置在 CPz 处,脑电图记录的通道位置如图 5b 红圈所围部分,实验过程中通道阻抗保持在 10 kΩ 以下。

为了获得脑电信号中的有效信息,对采集到的原始数据进行预处理,以平均电位作为参考,对信号进行工频干扰消噪以及 5~45 Hz 的带通滤波来提高 SSVEP 信号的信噪比,并按照标签进行分段处理,以时序、通道数、次数的 3 维数据矩阵结构进行存放。



(a)Neuracle 脑电采集设备 (b)采集通道位置

图 5 Neuracle 脑电采集设备和通道位置

Fig. 5 Neuracle EEG acquisition equipment and channel locations

2.4 传统方法的对比分析

本文针对 SSVEP 信号提出识别方案,为验证该方法的有效性,利用同组数据集设置以下 3 种方法作为对比:

(1)典型相关分析方法^[23],信号预处理后,直接使用 CCA 方法计算相关系数并进行无监督分类,设置谐波数为 2;

(2)功率谱密度分析(PSD)结合支持向量机(SVM)方法^[24],采用 welch 法提取功率谱密度特征,特定频段 δ 、 θ 、 α 、 β 节律的功率信息,并得到特征向量输入至 SVM 中训练,接着对测试集进行有监

督测试；

(3)CNN 方法,普通 CNN 网络仅从前层进一步进行特征学习提取深层特征,在网络后半部容易产生信息遗漏,以 MFCNN 相同设定参数搭建普通 CNN 网络结构进行识别。

定义识别准确率 A 以及信息传输率 v_{ITR} 作为评价指标。识别准确率表达式见式(6)。 v_{ITR} 代表单位时间周期内所传递的总信息量,表达式为

$$v_{\text{ITR}} = \frac{60}{T} \left[\text{lb}K + A\text{lb}A + (1 - A)\text{lb}\left(\frac{1 - A}{K - 1}\right) \right]$$

(7)

式中: K 为可选目标数; T 为识别一次刺激所需时间。

2.5 实验结果分析

实验计算在 Matlab2019 环境中进行,采用配置为 Intel(R) Core(TM) i5-4200H CPU @ 2.80 GHz、8 GB 运行内存的计算机,Matlab Deep Network 工具箱^[25]用于设计和测试所提出的网络,以图 1 所示模型结构搭建深度学习网络。为了获得受试者在不同时间窗长度刺激下的目标识别准确率,本文取单次刺激的前 3 s 数据划分成 6 个不同时间段进行分析,分别按照前 0.5 s、前 1 s、前 1.5 s、前 2 s、前 2.5 s 以及前 3 s 的时间窗口截取数据长度。

MFCNN 针对受试者 S1 在 1 s 时间窗长度的训练集最终训练结果如图 6a 所示,模型训练准确率最终收敛至 99.38%,损失值收敛至 0.031 9。利用训练完成如图 6b 所示 MFCNN 验证测试集的目标识别结果混淆矩阵,绿色块代表模型对该类标签样本的识别准确数以及该准确数的占比(准确数与标签样本总数的比值),红色块代表模型对该类标签样本的识别错误数以及该错误数的占比,浅灰色块表示对应标签样本的识别准确率和错误率。可以分析出识别后所得的预测标签结果与真实标签基本一

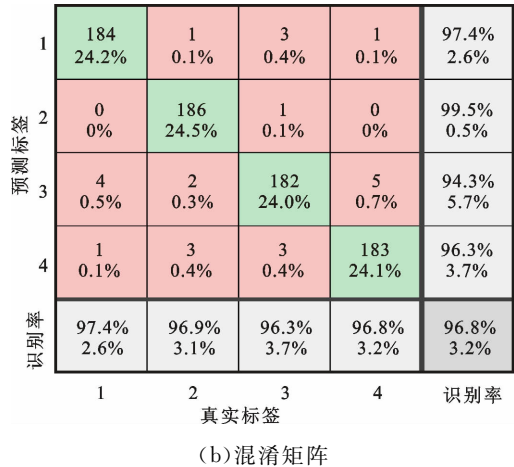
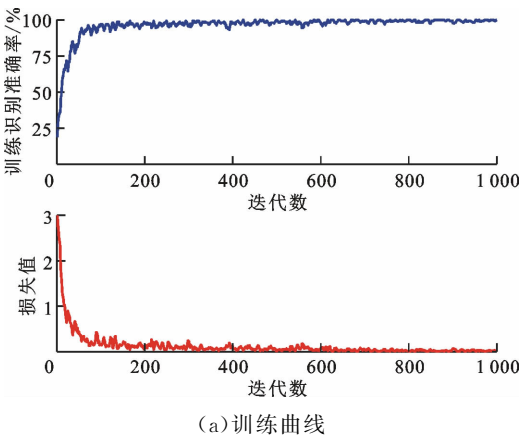


图 6 MFCNN 训练曲线与混淆矩阵

Fig. 6 MFCNN training curve and confusion matrix

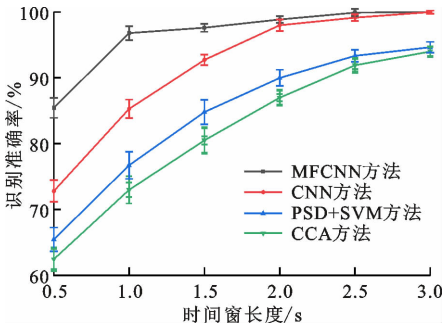
致,识别正确率为 96.8%,最多是在 8.57 Hz 频率中出现 7 个错误分类,占该类样本的 3.7%、整个测试样本的 0.9%。

针对受试者 S1~S7 的数据集利用 CCA 方法、PSD 结合 SVM 方法、CNN 方法 3 种对照方法来验证本文模型识别效果和特征提取对识别结果的影响,对不同时间窗长度下的数据重复进行 10 次实验以避免偶然性,统计不同受试者在不同时间窗长度的识别准确率、标准差以及多受试者平均识别准确率,实验结果如图 7 所示。

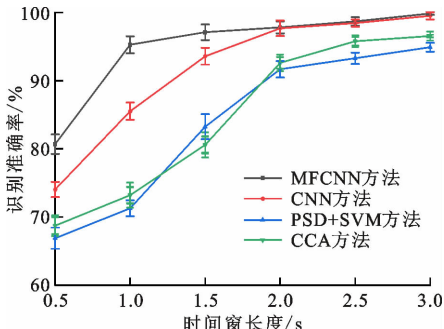
从图 7 所示实验结果可以看出,MFCNN 方法在对不同受试者的不同时间窗长度下均具有最高的识别准确率,与 PSD 结合 SVM 方法、CCA 方法相比提升效果显著。尤其是对于受试者 S1、S3、S5,本文方法相比 CCA 方法最少提升了 5.98%,最多提升了 24.81%,相比 PSD 结合 SVM 方法最少提升了 5.36%,最多提升了 21.73%。对于其他受试者 S2、S4、S6、S7,在不同时间窗长度下,识别准确率均有所提升,在 0.5 s、1 s 的短时间窗长度时最为显著,验证本文所提方法相较于这两种传统方法能够提高解码性能。从所有受试者的平均识别准确率实验结果可知,在 1 s 时间窗长度下,相较于这两种传统方法,本文方法对平均识别准确率的提升分别为 18.57%和 20.08%,且标准差为 0.988 1,低于两种传统方法的1.781 2、1.596 4,可见本文方法在提高了识别准确率的同时也具有更好的稳定性,优于 CCA 方法、PSD 结合 SVM 方法。

相较于普通 CNN 结构单纯从前层进行特征学习,在后期容易发生信息丢失特点,尤其是在时间窗长度较短、输入信息较少的情况下,频率特征尚不明

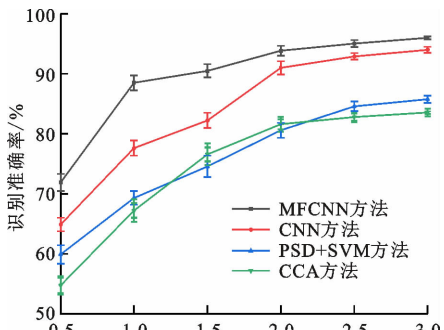
显,对比平均识别准确率实验结果可以看出,当时间窗长度仅为 0.5 s 时,MFCNN 的平均识别准确率为 85.03%,高出 CNN 方法 8.12%;当时间窗长度为 1 s 时,MFCNN 的平均识别准确率为 94.25%,



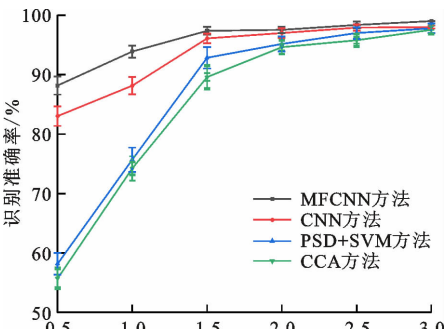
(a) S1



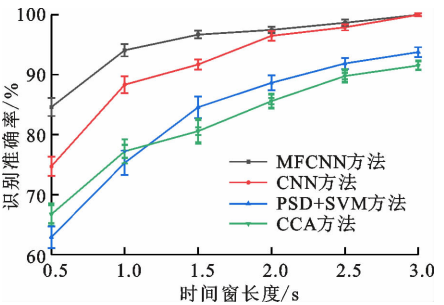
(b) S2



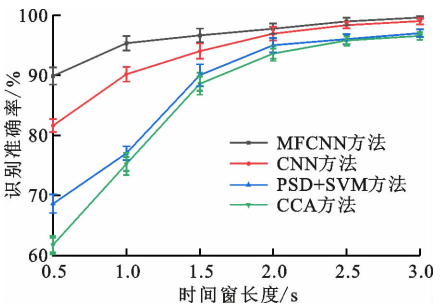
(c) S3



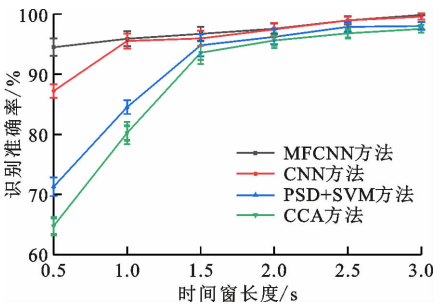
(d) S4



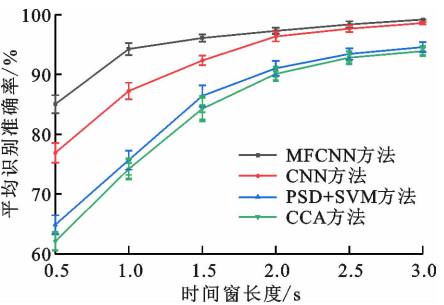
(e) S5



(f) S6



(g) S7



(h) 平均识别准确率

图 7 采用 4 种方法的受试者不同时间窗长度识别准确率的对比

Fig. 7 Comparison of the recognition accuracy of subjects with different time window lengths using the four methods

高出 CNN 方法 7.03%。验证 MFCNN 结合网络浅层细节特征以及深层语义特征,即携带更多的信号特征,能够充分利用信号不同层次特征加强网络学习能力,以提高识别准确率。随着时间窗长度的不

断增加,频率特征逐渐明显,当时间窗长度大于 2 s 后,MFCNN 与 CNN 的识别准确率并无较大差异。分析表明,融合特征能够在短时间窗长度下携带更多的脑电信息,有利于 SSVEP 信号的目标识别,MFCNN 模型要优于普通 CNN 结构特征分类模型。

各受试者在 1 s 时间窗长度时 4 种识别方法的信息传输率如表 2 所示。由于受试者 S7 本身对 SSVEP 响应特征较为明显,本文方法相比于 CNN 方法对其识别准确率的提升有限,存在个体差异性,

信息传输率又同时受识别时间和识别准确率的影响,所得计算结果略低于 CNN 方法,但是综合实验所有受试者平均数据结果可得,MFCNN 的平均信息传输率为 96.86 bit/min,对比 CCA 方法和 PSD 结合 SVM 方法,在信息传输率上有显著提升,分别提升了 53.15 bit/min、49.58 bit/min,对比 CNN 方法提升了 12.15 bit/min,说明 MFCNN 在提升了识别准确率的同时,也能提升信息传输速率,具有较高的目标识别效率,表明本文方法能有效提高 SSVEP-BCI 系统在较短时间窗长度刺激时的识别性能,

表 2 受试者 1 s 时间窗长度时 4 种方法的信息传输率

Table 2 Information transmission rate of 4 methods for each subject with 1 s time window length

受试者	信息传输率/(bit·min ⁻¹)			
	MFCNN	CCA	PSD 结合 SVM	CNN
S1	103.93±15.54	36.56±18.43	50.02±23.13	79.36±11.21
S2	96.58±13.32	51.88±24.57	45.89±15.45	88.21±17.66
S3	81.30±27.43	22.58±15.98	28.31±26.93	56.07±19.27
S4	99.94±9.78	46.02±22.83	50.84±19.34	93.50±15.81
S5	96.01±12.37	43.09±16.28	46.98±14.73	87.68±13.46
S6	98.78±16.33	40.16±26.96	34.06±23.07	79.03±15.83
S7	101.49±14.27	65.67±19.27	74.88±26.75	109.13±21.34
平均值	96.86±15.58	43.71±20.62	47.28±21.34	84.71±16.37

3 结 论

本文针对传统目标识别方法在对 SSVEP 信号进行识别时存在的局限性,提出基于多尺度特征融合卷积神经网络的信号识别方法,利用小波变换对信号进行时频域图像转化,搭建多尺度网络模型进行信号特征学习,提取图像不同层次特征并进行融合,建立多尺度特征向量作为分类依据,最终实现 SSVEP 信号端到端的目标识别。实验结果表明,本文方法对于场景动画刺激范式在较短时间刺激下具有明显的识别优势,当时间窗长度刺激为 1 s 时,平均识别准确率就已达到 94.25%,信息传输率达到 96.86 bit/min,高于无训练 CCA 方法、有训练 PSD 结合 SVM 方法和普通 CNN 方法,说明本文方法适用于 SSVEP 信号,可进一步提高信号的识别准确率,同时还具有较高的信息传输率,为稳态视觉诱发电位脑机接口在特征提取和识别准确率提高等方面提供了参考。

参考文献:

[1] ZAVALA S P, LÓPEZ J L M, CHICAIZA K O, et

al. Review of steady state visually evoked potential brain-computer interface applications: technological analysis and classification [J]. Journal of Engineering and Applied Sciences, 2020, 15(2): 659-678.

[2] 于淑月, 李想, 于功敬, 等. 脑机接口技术的发展与展望 [J]. 计算机测量与控制, 2019, 27(10): 5-12.

YU Shuyue, LI Xiang, YU Gongjing, et al. Development and prospect of brain computer interface technology [J]. Computer Measurement & Control, 2019, 27(10): 5-12.

[3] CHAMOLA V, VINEET A, NAYYAR A, et al. Brain-computer interface-based humanoid control: a review [J]. Sensors, 2020, 20(13): 3620.

[4] 葛松, 徐晶晶, 赖舜男, 等. 脑机接口: 现状问题与展望 [J]. 生物化学与生物物理进展, 2020, 47(12): 1227-1249.

GE Song, XU Jingjing, LAI Shunnan, et al. Brain computer interface: current situation, problems and prospects [J]. Progress in Biochemistry and Biophysics, 2020, 47(12): 1227-1249.

[5] 徐光华, 张锋, 谢俊, 等. 稳态视觉诱发电位的脑机接口范式及其信号处理方法研究 [J]. 西安交通大学学报, 2015, 49(6): 1-7.

- XU Guanghua, ZHANG Feng, XIE Jun, et al. Brain-computer interface paradigms and signal processing strategy for steady state visual evoked potential [J]. Journal of Xi'an Jiaotong University, 2015, 49(6): 1-7.
- [6] VIALATTE F B, MAURICE M, DAUWELS J, et al. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives [J]. Progress in Neurobiology, 2010, 90(4): 418-438.
- [7] 张黎明, 张小栋, 陆竹风, 等. 用于稳态视觉诱发电位特征频率提取的同步压缩短时傅里叶变换方法 [J]. 西安交通大学学报, 2017, 51(2): 20-26.
- ZHANG Liming, ZHANG Xiaodong, LU Zhufeng, et al. A synchrosqueezing short-time Fourier transform method for characteristic frequency extraction of steady state visual evoked potential [J]. Journal of Xi'an Jiaotong University, 2017, 51(2): 20-26.
- [8] 杜光景, 谢俊, 张玉彬, 等. 用于稳态视觉诱发电位脑机接口目标识别的深度学习算法 [J]. 西安交通大学学报, 2019, 53(11): 42-48.
- DU Guangjing, XIE Jun, ZHANG Yubin, et al. A deep learning method for target recognition in steady-state visual evoked potential-based brain-computer interface [J]. Journal of Xi'an Jiaotong University, 2019, 53(11): 42-48.
- [9] ANTELIS J M, RIVERA C A, GALVIS E, et al. Detection of SSVEP based on empirical mode decomposition and power spectrum peaks analysis [J]. Biocybernetics and Biomedical Engineering, 2020, 40(3): 1010-1021.
- [10] CHEN Guang, SONG Dandan, LIAO Lejian. A multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis in the frequency domain [J]. Advances in Intelligent Systems and Computing, 2014, 215: 603-616.
- [11] WU Ting, YAN Guozheng, YANG Banghua, et al. EEG feature extraction based on wavelet packet decomposition for brain computer interface [J]. Measurement, 2008, 41(6): 618-625.
- [12] TABAR Y R, HALICI U. A novel deep learning approach for classification of EEG motor imagery signals [J]. Journal of Neural Engineering, 2017, 14(1): 016003.
- [13] SAKHAVI S, GUAN Cuntai, YAN Shuicheng. Parallel convolutional-linear neural network for motor imagery classification [C]//Proceedings of the 2015 23rd European Signal Processing Conference. Piscataway, NJ, USA: IEEE, 2015: 2736-2740.
- [14] CECOTTI H. A time-frequency convolutional neural network for the offline classification of steady-state visual evoked potential responses [J]. Pattern Recognition Letters, 2011, 32(8): 1145-1153.
- [15] 于鸿伟, 谢俊, 何柳诗, 等. 基于卷积神经网络和支持向量机的事件相关电位识别方法 [J]. 西安交通大学学报, 2021, 55(12): 47-54.
- YU Hongwei, XIE Jun, HE Liushi, et al. An event-related potential recognition method based on convolutional neural network and support vector machine [J]. Journal of Xi'an Jiaotong University, 2021, 55(12): 47-54.
- [16] 李玉花, 柳倩, 韦新, 等. 基于卷积神经网络的脑电信号分类 [J]. 科学技术与工程, 2020, 20(15): 6135-6140.
- LI Yuhua, LIU Qian, WEI Xin, et al. Classification of electroencephalogram signals based on convolutional neural network [J]. Science Technology and Engineering, 2020, 20(15): 6135-6140.
- [17] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [18] 张小栋, 郭晋, 李睿, 等. 表情驱动下脑电信号的建模仿真及分类识别 [J]. 西安交通大学学报, 2016, 50(6): 1-8.
- ZHANG Xiaodong, GUO Jin, LI Rui, et al. A simulation model and pattern recognition method of electroencephalogram driven by expression [J]. Journal of Xi'an Jiaotong University, 2016, 50(6): 1-8.
- [19] 于淑月, 李想, 于功敬, 等. 脑机接口技术的发展与展望 [J]. 计算机测量与控制, 2019, 27(10): 5-12.
- YU Shuyue, LI Xiang, YU Gongjing, et al. Development and prospect of brain computer interface technology [J]. Computer Measurement & Control, 2019, 27(10): 5-12.
- [20] 陈保家, 陈学力, 沈保明, 等. CNN-LSTM 深度神经网络在滚动轴承故障诊断中的应用 [J]. 西安交通大学学报, 2021, 55(6): 28-36.
- CHEN Baojia, CHEN Xueli, SHEN Baoming, et al. An application of convolution neural network and long short-term memory in rolling bearing fault diagnosis [J]. Journal of Xi'an Jiaotong University, 2021, 55(6): 28-36.
- [21] 张鹏. 基于卷积神经网络的多尺度遥感图像目标检测方法研究 [D]. 西安: 西安电子科技大学, 2019: 37-39.

(下转第 202 页)

- mechanism for bearing fault diagnosis [J]. *Applied Soft Computing*, 2020, 97: 106829.
- [19] 方军, 管业鹏. 基于双编码器的会话型推荐模型 [J]. *西安交通大学学报*, 2021, 55(8): 166-174.
FANG Jun, GUAN Yepeng. A session recommendation model based on dual encoders [J]. *Journal of Xi'an Jiaotong University*, 2021, 55(8): 166-174.
- [20] 曹仰杰, 李昊, 段鹏松, 等. AHNNet: 融合注意力机制的行为识别混合神经网络模型 [J]. *西安交通大学学报*, 2021, 55(5): 123-132.
CAO Yangjie, LI Hao, DUAN Pengsong, et al. AHNNet: human activity recognition based on hybrid neural network combining attention mechanism [J]. *Journal of Xi'an Jiaotong University*, 2021, 55(5): 123-132.
- [21] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//*Proceedings of the European Conference on Computer Vision*. Berlin, Germany: Springer International Publishing, 2018: 3-19.
- [22] 王美华, 吴振鑫, 周祖光. 基于注意力改进 CBAM 的农作物病虫害细粒度识别研究 [J]. *农业机械学报*, 2021, 52(4): 239-247.
WANG Meihua, WU Zhenxin, ZHOU Zuguang. Fine-grained identification research of crop pests and diseases based on improved CBAM via attention [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(4): 239-247.
- [23] NECTOUX P, GOURIVEAU R, MEDJAH K, et al. PRONOSTIA: an experimental platform for bearings accelerated degradation tests [C]//*Proceedings of the IEEE International Conference on Prognostics and Health Management*. Piscataway, NJ, USA: IEEE, 2012: 1-8.
- [24] REN Lei, CUI Jin, SUN Yaqiang, et al. Multi-bearing remaining useful life collaborative prediction: a deep learning approach [J]. *Journal of Manufacturing Systems*, 2017, 43(2): 248-256.
- [25] GUO Liang, LEI Yaguo, LI Naipeng, et al. Deep convolution feature learning for health indicator construction of bearings [C]//*Proceedings of the 2017 Prognostics and System Health Management Conference*. Piscataway, NJ, USA: IEEE, 2017: 1-6.
- [26] 张继冬, 邹益胜, 邓佳林, 等. 基于全卷积层神经网络的轴承剩余寿命预测 [J]. *中国机械工程*, 2019, 30(18): 2231-2235.
ZHANG Jidong, ZOU Yisheng, DENG Jialin, et al. Bearing remaining life prediction based on full convolutional layer neural networks [J]. *China Mechanical Engineering*, 2019, 30(18): 2231-2235.

(编辑 武红江 刘杨)

(上接第 193 页)

- [22] 李睿, 张小栋, 张黎明, 等. 面向假肢的场景动画稳态视觉诱发脑控方法 [J]. *西安交通大学学报*, 2017, 51(1): 115-121.
LI Rui, ZHANG Xiaodong, ZHANG Liming, et al. Brain-controlled prosthesis manipulation based on scene graph-SSVEP [J]. *Journal of Xi'an Jiaotong University*, 2017, 51(1): 115-121.
- [23] 何柳诗, 谢俊, 于鸿伟, 等. 融合眼动追踪和目标动态可调的稳态视觉诱发电位脑机接口系统设计 [J]. *西安交通大学学报*, 2021, 55(10): 87-95.
HE Liushi, XIE Jun, YU Hongwei, et al. A design of brain-computer interface system for steady-state visual evoked potential integrating eye movement tracking and target dynamical adjustment [J]. *Journal of Xi'an Jiaotong University*, 2021, 55(10): 87-95.
- [24] MANSHOURI N, MALEKI M, KAYIKCIOGLU T. An EEG-based stereoscopic research of the PSD differences in pre and post 2D&3D movies watching [J]. *Biomedical Signal Processing and Control*, 2020, 55: 101642.
- [25] PALM R B. Prediction as a candidate for learning deep hierarchical models of data [D]. Lyngby: Technical University of Denmark, 2012: 46-51.

(编辑 武红江 刘杨)