

DOI: 10.13382/j.jemi.B2104154

基于 CNN 时-空卷积优化的 EM-EEG 识别方法研究*

黄永庆^{1,2} 周强^{1,2}

(1. 陕西科技大学电气与控制工程学院 西安 710021; 2. 陕西科技大学陕西省人工智能联合实验室 西安 710021)

摘要:针对当前情绪脑电信号(emotion electroencephalogram, EM-EEG)识别研究中时间域信息的时间尺度难以把握和空间域信息易被忽视致使辨识度停滞不前,以及采集 EM-EEG 时通道过多导致信息冗余和信息处理成本增加等问题,提出了基于 CNN 的时-空卷积优化融合网络进行 EM-EEG 识别研究。该融合网络由提取 EM-EEG 时域信息的长卷积(long convolution, L-Conv) CNN 和提取 EM-EEG 空域信息的 CNN 并联组成,在 CNN 模型时-空优化中使用粒子群算法(particle swarm optimization, PSO)对时域 CNN 中的 L-Conv 尺度进行了优化,并使用短时功率谱(short time power spectrum, STPS)的相关分析方法进行空域 CNN 模型通道数目优化,深层且有效地提取了 EEG 中的时间域和空间域特征。结果表明,提出的时-空卷积优化融合 CNN 在 SEED IV 数据集上对平和、悲伤、恐惧、高兴 4 种情绪最终准确率可以达到 90.13%,相比传统单一 CNN 的识别准确率提高了 4.76%,并且通道数目由 62 路降低至 33 路,缩减了 46.77%,证实了本方法的可行性。

关键词: EM-EEG; 时-空卷积优化; 粒子群算法; STPS 相关分析; SEED IV 数据集

中图分类号: TP391; TN911 **文献标识码:** A **国家标准学科分类代码:** 520.20

Research on EM-EEG recognition method based on CNN time-space convolution optimization

Huang Yongqing^{1,2} Zhou Qiang^{1,2}(1. School of Electrical and Control Engineering, Shaanxi University of Science and Technology, Xi'an 710021, China;
2. Shaanxi Joint Laboratory of Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, China)

Abstract: In view of the current emotional electroencephalogram (EM-EEG) identification research on time scales is difficult to grasp the time domain information and the spatial domain information is easy to ignore the recognition rate is stagnant, and collect the EM-EEG with too many channels in the excessive information redundancy and increasing cost of information processing problems, it puts forward the space-time convolution based on CNN optimization study on EM-EEG identification fusion network. The fusion network is composed of a parallel long convolution (L-Conv) CNN that extracts EM-EEG time domain information and a CNN that extracts EM-EEG spatial information. Particle swarm optimization (PSO) is used in the time-space optimization of the CNN model. The L-Conv scale in CNN has been optimized, and use the short time power spectrum (STPS) correlation analysis method of the spatial CNN channel number optimization model, temporal and spatial domain features in EEG are extracted deeply and effectively. The results show that the proposed optimization of space-time convolution integration CNN on SEED IV data set for peace, sadness, fear, happy four final accuracy can reach 90.13%, compared with the traditional single CNN recognition accuracy rate increased by 4.76%, and channel number from 62 to 33 road, shrank by 46.77%, confirmed the feasibility of this method.

Keywords: emotional electroencephalogram (EM-EEG); time-space convolution optimization; particle swarm optimization (PSO); STPS correlation analysis; SEED IV data set

0 引言

情感^[1]是人与人之间交流过程中不可或缺的一部分,是综合了感觉、感知、思维和行为的心理活动过程,能反映一个人当下的心理状态,对人们的认知、沟通和决策等产生重要影响^[2]。情感的表达可以分为脑电、肌电、皮肤电、心电等生理信号和面部表情、语言表达等非生理信号。相比于非生理信号易于伪装的特点,生理信号更能够反映个体的真实情绪状态。神经心理学研究发现脑电信号(electroencephalogram, EEG)不仅可以反映大脑组织的各种脑电活动及大脑的功能状态,而且可以反应人类情感状态的有效信息^[3],并且情感的产生或活动与大脑皮层的活动相关度很大^[4],对情绪的变化有着更准确的表征^[5],所以情绪脑电信号(emotion electroencephalogram, EM-EEG)作为反映情绪状态的主要载体在情绪辨识研究中受到广泛关注。随着 5 G 技术、物联网、人机交互等技术的不断发展,在医疗保健、媒体娱乐、信息检索、教育以及智能可穿戴设备等领域都有着广阔的应用前景^[6]。

在情绪识别研究中,影响准确率的因素主要包括采集信号质量、采集环境等外部因素和数据处理、特征提取等内部因素。为了得到有效的判别准确率,学者们从不同角度对 EM-EEG 辨识展开了一系列实验探究。文献[7]从时域角度出发,提取 EEG 的功率谱密度(power spectral density, PSD)作为特征,分类器选用 SVM,对开心、生气、悲伤、欣喜 4 种情绪状态进行分类,准确率达到 82.29%;文献[8]从频域角度提取频谱特征,使用 KNN 分类器,对积极、消极、中性 3 种情绪分类的准确率达到 85%;文献[9]将原始 EEG 分解为若干本征模函数,并计算其样本熵,经过极限学习机训练和测试,在 DEAP 数据集上平均识别率为 88.93%。以上研究主要通过传统机器学习方法提取特征进而分类,始终受到手动提取 EM-EEG 特征这一盲目、繁琐、低效工作的困扰。因此,随着近年来深度学习的兴起,越来越多的研究开始向深度学习聚焦,由于深度学习具有自动提取深层次特征的能力,从而避免了传统机器学习繁杂的人工特征提取环节。深度信念网络(deep belief network, DBN)、卷积神经网络(convolutional neural network, CNN)、长短时记忆网络(long short-term memory, LSTM)、循环神经网络(recurrent neural networks, RNN)等多种深度网络模型也逐渐被用于大规模 EEG 的分类和识别领域。由于 CNN 具有局部感知和参数共享两个特点,能够很好地自动挖掘 EEG 中深层情感特征且模型复杂度较低,所以以 CNN 为代表的研究最为普遍。如 Rui 等^[10]使用 CNN 的方法对 DEAP 数据集中的 EEG 进行了情感识别,达到了 87.27%

的平均识别精度;Machot 等^[11]使用 CNN 的方法分别对 DEAP 数据集和 MAHNOB 数据集中的 EDA 信号进行情感识别,达到了 82%和 78%的识别精度;Tripathi 等^[12]使用 CNN 在 DEAP 数据集^[13]效价以及唤醒度上进行情感识别,平均准确率分别为 81.41%和 73.36%;Mei 等^[14]研究了一种用 CNN 直接利用函数连通矩阵和特征提取的方法,同时完成了二分类任务、三分类任务和四分类任务,其平均准确率分别达到了 85%、78%、75%。

截至当前, CNN 在情绪辨识研究中虽然可以达到一定精度,但是常用的 CNN 矩形卷积核结构在特征提取中难以精确提取 EM-EEG 的强时间特征^[15],导致情绪辨识准确率徘徊不前。另外,单一结构的 CNN 网络提取特征的能力有限,难以获得 EEG 完整的特征信息,尤其是空间特征信息最易被忽略。与此同时,随着 EEG 检测技术的升级,EEG 采集通道数不断增加,通道数的增加一方面使采集到的 EEG 信息量增大,另一方面也使得通道之间的距离缩小,造成通道之间的信号耦合干扰增强,并引入更多的噪声干扰^[16];其次,采集设备的成本也显著增大,因此在确保 EM-EEG 辨识精度的同时缩减通道数量成为目前一个重要研究方向。

针对以上问题,本文提出了基于 CNN 时-空卷积优化的 EM-EEG 识别方法,能够有效提取 EEG 中时间域和空间域特征进行分类辨识。在利用 CNN 对 EM-EEG 的时域和空域信息进行卷积的特征提取过程中,引入了长卷积(long convolution, L-Conv)提取 EM-EEG 的时间信息,并基于粒子群算法(particle swarm optimization, PSO)对 L-Conv 的尺度进行优化;为了有效提取 EM-EEG 的空间信息,利用了短时功率谱(short time power spectrum, STPS)的互相关分析技术对输入 CNN 的 EM-EEG 通道数进行优化,最终将时间域和空间域的两个网络的信息进行融合。对于完成训练的网络模型,使用 SEED IV 数据库^[17](含平和、悲伤、恐惧、高兴 4 种情绪数据)进行实验。结果表明,本文提出的 CNN 时-空卷积优化方法有效提升了情绪的分类准确率,在将 EM-EEG 通道数由 62 路降至 33 路的情况下,辨识准确率达到 90.13%。

1 本文方法思路

1.1 EM-EEG 研究中的问题及难点

1) EM-EEG 长卷积核的时间尺度难以把握,空间信息易被忽略

由于 EM-EEG 在时间维度上含有很强的特征信息,多数研究人员在时域上利用 CNN 对 EM-EEG 进行研究,虽然传统 CNN 中的矩形卷积核对 EM-EEG 的时间信息具有一定的表征能力,但是由于对时间尺度无法精确把握,难以准确捕捉 EM-EEG 中本质的特性,因此辨识效果

欠佳。与此同时,在对 EM-EEG 的辨识研究中,研究人员往往只从单一的时域或频域角度出发进行特征提取及情绪分类,而忽略了 EM-EEG 中的空间信息。

2) EEG 采集通道过多导致信息冗余,EM-EEG 数据运算量过大

随着脑电技术的发展,采集通道数目逐渐增加,当前的 EM-EEG 通道通常有 64 路、128 路甚至达到 256 路。虽然数量不断增多的采集通道提供了更丰富且全面的脑电信息,但也造成了 EM-EEG 数据维度过高、信息冗余,

使得数据运算量过大,使得当前的 EM-EEG 辨识方法对计算机性能要求苛刻。

1.2 时-空卷积优化融合 CNN 的 EM-EEG 识别方法

针对以上难点,本文方法研究思路如图 1 所示,由信号获取、数据输入、特征提取、特征融合和辨识分类等部分组成。其中,在时域特征提取中利用粒子群算法进行时间尺度优化,在空间信息矩阵输入中利用 STPS 的相关分析实现空间优化。

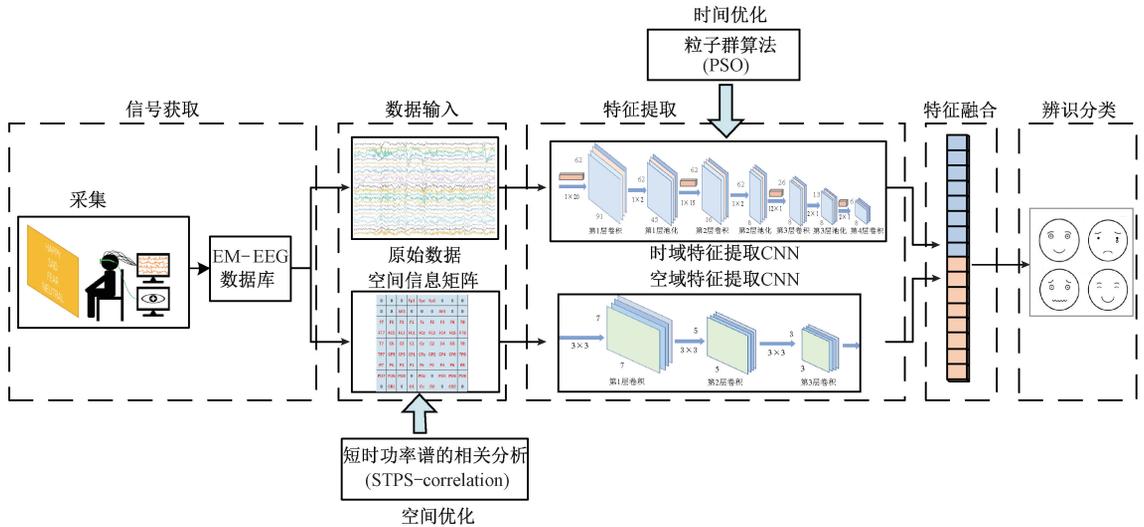


图 1 总体研究思路框架图

Fig. 1 Overall research thinking frame diagram

1) 构建提取 EM-EEG 时-空特征的 CNN 融合网络

①时域特征提取 CNN 中引入 L-Conv 模块

在 AlexNet、VGGNet、ResNet 等 CNN 分类框架中,卷积核大小多数为 3×3 或 5×5 。本文使用 CNN 作为情绪分类主要架构,结合 EM-EEG 以时序为主的特点,将传统应用中矩形卷积核的维度设定为 $1 \times n$ 或 $n \times 1$,使得在进行卷积特征提取的过程中,能够成功提取到时间域上的特征信息。L-Conv 的应用有效地避免卷积过程分布不规范的通道相互之间信息的混杂,对每一个通道在时间域上进行分步处理。

②EM-EEG 空间信息矩阵的构造

采集 EEG 时,虽然不同位置采集得到的信号不尽相同甚至差异巨大,然而在 EM-EEG 的研究中,空间位置信息却容易被忽略,因此本文在对 EM-EEG 进行识别的研究中考虑位置空间信息,根据电极位置将 62 个通道数据排布成 9×9 的二维矩阵作为 CNN 的输入,而其中没有电极对应的单元填充 0,如图 2 所示。EM-EEG 的二维矩阵作为 CNN 的输入,对空间位置特征进行提取。

2) CNN 融合网络的时-空优化

①时间优化:L-Conv 的尺度优化

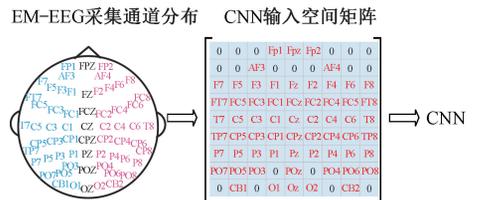


图 2 EM-EEG 空间信息矩阵构造示意图

Fig. 2 Schematic diagram of EM-EEG spatial information matrix construction

本文在 EM-EEG 时间域特征提取中,提出了在时间维度上使用 L-Conv 逐通道进行卷积,由于 EM-EEG 大尺度上的时变性和小尺度上的时间惯性,决定了 EM-EEG 数据在一定时间尺度上的依赖关系,那么在时间域的 L-Conv 尺度优化是准确提取时间特征的关键。本文利用 PSO 优化算法,将 L-Conv 的尺度作为优化参数,CNN 模型训练后的最低偏差率作为优化目标,借助于该群体智能优化算法实现 CNN 网络的时域优化。

②空间优化:通道简化

为了解决 EM-EEG 采集通道数目过多带来的信息冗

余、数据维数过高等问题,本文在研究 EM-EEG 空间特征相似性的基础上进行了通道优化,以此来降低通道数目。首先,利用 STPS 提取了不同通道的 EEG 时间-频率-能量密度分布,然后计算任意两个通道间 STPS 的 Pearson 相关系数,进而得到所有 EM-EEG 通道间的 STPS 互相关系数和 STPS 自相关系数,将各通道自相关系数和互相关系数的绝对值累计,累积量反映出该通道含有的信息量大小,累积量大于某阈值的通道将被保留。

2 本文方法的实现

2.1 构建 CNN 时-空融合网络

1) 搭建 L-Conv 时域 CNN 网络

本文在传统 CNN 基础上,对其卷积核进行了相应的改进,在卷积特征提取环节引入了 L-Conv,利用 L-Conv 在 EM-EEG 的时间维度上进行特征信息提取。图 3 为 EM-EEG 时域特征提取网络结构,该网络由 4 层卷积、3 层池化组成,其中卷积过程的卷积核选用 $1 \times n$ 或 $n \times 1$ 的 L-Conv。前 4 层网络的卷积和池化均是利用 L-Conv 对每一个单一通道进行处理,之后的卷积和池化主要是对数据通道进行降维,以减少运算量。

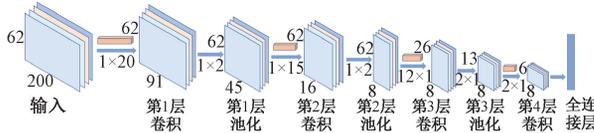


图 3 EM-EEG 时域特征提取 CNN 结构

Fig. 3 EM-EEG time domain feature extraction of CNN structure

2) 搭建空域 CNN 网络

大量的研究表明,不同情绪诱发会在头部不同空间位置显示不一样的刺激反应^[18],因此,电极的空间位置也包含着很重要的信息,本文中根据电极的空间位置信息,将 EEG 按照空间位置排布成二维数组输入 CNN,通过 CNN 的层层特征提取和组合获得通道间的空间关联。其网络结构如图 4 所示,由于输入数据维数本身较小,所以在空域提取特征中仅使用 3 层卷积,将 9×9 的输入最终降维成 3×3 的特征矩阵,之后通过全连接层将特征参数变成一维数组。

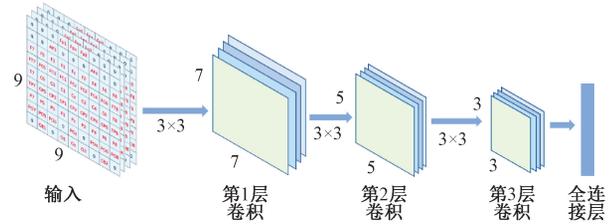


图 4 EM-EEG 空域特征提取 CNN 结构

Fig. 4 EM-EEG spatial feature extraction of CNN structure

如图 5 所示,将分别提取时间信息和空间信息的两个 CNN 网络的全连接层进行信息融合,随后经过降维处理,最终进行情绪状态的辨识。

2.2 基于 PSO 的 L-Conv 尺度优化

1) L-Conv 尺度优化

为了获取 EM-EEG 采集过程中的时域特征,本文搭建了 L-Conv 时域 CNN 模型,模型中的 L-Conv 尺度对于有效提取隐含在 EM-EEG 中(不同时间跨度数据关联

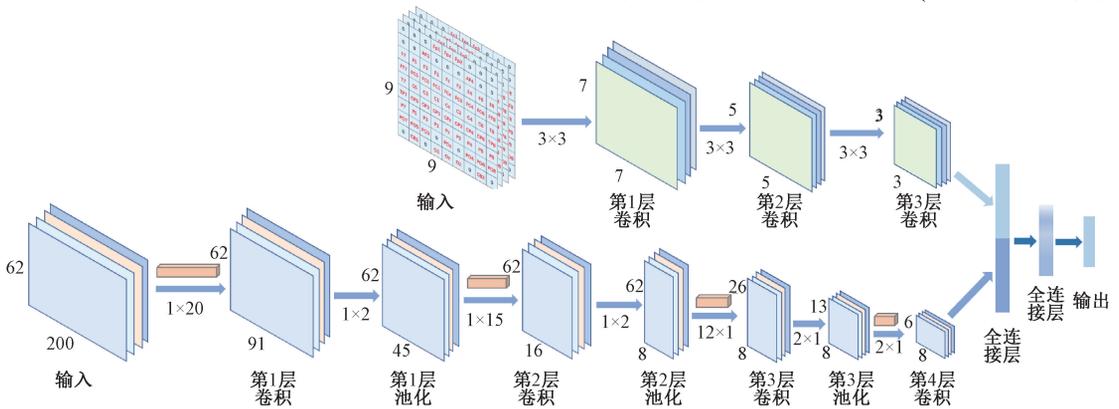


图 5 EM-EEG 时-空域特征提取和融合的 CNN 结构

Fig. 5 EM-EEG time-spatial feature extraction and fusion of CNN structure

所造成)的时域特征至关重要,因此有必要对各层 CNN 的 L-Conv 尺度进行优化。本文以 CNN 训练结果的偏差率 P (与准确率 ACC 等价)最低为优化目标,以 CNN 中的 4 个卷积层的 L-Conv 尺度 $L1, L2, L3, L4$ 为优化参数,

优化目标为优化目标函数可表示为

$$P_{\min}(L1, L2, L3, L4) = \min_{L1, L2, L3, L4} \|O^i - O_{CNN}^i(L1, L2, L3, L4)\|_2^2 \quad (1)$$

其中, P_{\min} 是 CNN 训练结果的最小偏差率, Q^i 是

CNN 的期望输出, Q_{CNN}^i 是 CNN 的实际输出。

2) PSO 优化算法

由于式(1)的优化目标函数含有结构复杂的 CNN 模型,传统优化方法难以胜任,因此本文使用 PSO 算法。PSO 是 Eberhart 和 Kennedy 提出的一种迭代的群优化算法^[19],该方法以其强大的全局搜索能力和非线性处理能力,以及简单易行、收敛速度快、设置参数少等优点,在复杂函数优化、模式分类和神经网络等领域有着广泛的应用,所以本文使用 PSO 群智能算法实现 CNN 网络 L-Conv 尺度优化,优化过程如图 6 所示。

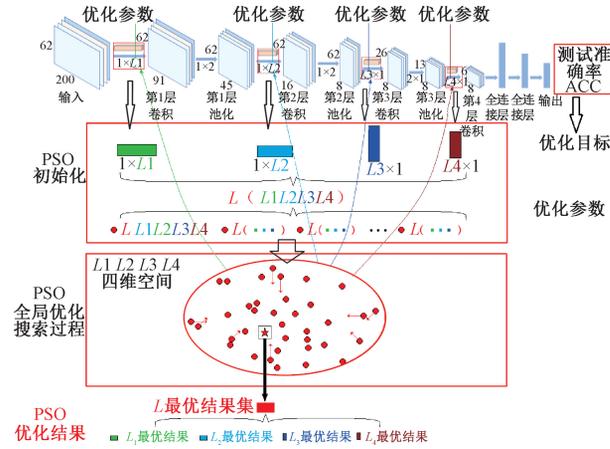


图 6 基于 PSO 的 L-Conv 尺度优化示意图

Fig. 6 Schematic diagram of L-CONV scale optimization based on PSO

粒子群在优化参数 L_1, L_2, L_3, L_4 的四维空间中以群体形式进行迭代搜索,每个粒子在迭代过程中根据式(2)和(3)更新其在四维空间中的搜索速度和当前位置,在解空间中寻找最优解。

粒子 i 的第 d 维空间上的速度更新公式为:

$$v_i^{k+1} d = w v_{id}^k + c_1 r_1 (L_{opt id}^k - L_{id}^k) + c_2 r_2 (L_{opt d}^k - L_d^k)$$

$$d = \{1, 2, 3, 4\} \quad (2)$$

粒子 i 的第 d 维位置更新公式为:

$$L_{id}^{k+1} = L_{id}^k + v_{id}^{k+1} \quad (3)$$

其中, v_{id}^k 为第 k 次迭代粒子 i 速度的第 d 维分量, L_{id}^k 为第 k 次迭代粒子 i 位置的第 d 维分量, $L_{opt id}^k$ 是粒子 i 局部最优值的第 d 维分量, $L_{opt d}^k$ 是全局最优值的第 d 维分量, w 是惯性权重, c_1 和 c_2 是学习因子, r_1 和 r_2 是 $(0, 1)$ 的随机数。

设定优化目标为寻找最小值,则 $L_{opt i}^k$ 和 L_{opt}^k 的更新方式为:

$$L_{opt i}^{k+1} = \begin{cases} L_i^{k+1} & P_{\min}(L_i^{k+1}) < P_{\min}(L_{opt i}^k) \\ L_{opt i}^k & P_{\min}(L_i^{k+1}) \geq P_{\min}(L_{opt i}^k) \end{cases} \quad (4)$$

$$L_{opt}^k = \arg \min_{1 \leq i \leq N} (P_{\min}(L_{opt i}^k)) \quad (5)$$

其中, $P_{\min}(L_d, d=1, 2, 3, 4)$ 是目标函数, N 为粒子群中粒子的总数。

3) 优化过程

本文将 CNN 通道中的 4 个 L-Conv 作为 PSO 粒子的四维分量,如表 1 前 3 列所示,以 CNN 训练并测试得到的测试误差作为适应度函数值,由 PSO 算法得到的最终结果即为自动选择的 L-Conv 尺度最优值。

表 1 长卷积核参数范围及 PSO 基本参数设置

Table 1 Set parameters range of long roll core and basic parameters of PSO

粒子坐标	参数	取值的范围	基本参数	数值
L_1	$L1$	[3,50]区间的整数	种群规模	4 000
L_2	$L2$	[3,50]区间的整数	最大迭代次数	15 000
L_3	$L3$	[3,30]区间的整数	w	0.85
L_4	$L4$	[3,20]区间的整数	c_1, c_2	0.1, 0.05

具体实现步骤如下:

步骤 1: PSO 基本参数设置如表 1 所示,包括 PSO 的种群规模,最大迭代次数及惯性权重 w ,学习因子 c_1, c_2 。

步骤 2: 将表 1 中的第 2 列参数作为 PSO 算法粒子分量,初始化粒子。

步骤 3: 以每个粒子所在的四维空间位置作为 CNN 模型的 L-Conv 尺度,随机初始化 CNN 权重,通过训练确定 CNN,将 CNN 测试误差作为该粒子的适应度函数值,确定 $L_{opt i}$ 和 L_{opt}^k 。

步骤 4: 若 L_{opt} 的适应度函数值小于给定阈值或达到最大迭代次数,则停止迭代;否则,利用式(2)、(3)更新粒子,返回步骤 3。

最终得到的 L_{opt} 即为最佳 CNN 结构,最佳尺度参数如表 2 所示。

表 2 长卷积核最优参数

Table 2 Optimal parameters of long roll kernel

参数	最优参数
$L1$	20
$L2$	15
$L3$	12
$L4$	2

2.3 基于 STPS 互相关分析的 EM-EEG 通道优化

1) 基于 STPS 估计的 EM-EEG 通道相似性分析

STPS 是处理非平稳信号的一种有效方法,它结合了短时傅里叶变换和功率谱估计的特点,所以本文以 STPS 互相关系数来量化通道间 EM-EEG 信号的相关性,并以任一通道的自相关系数和互相关系数的绝对值之合来反映该通道的信息量大小和重要性,从而作为 EM-EEG 通道选择判别依据。

STPS 通过对非平稳信号分别进行加窗平稳化和功

率谱估计,并通过时间窗的移动,对每个时刻都进行频率能量估计,其计算公式为:

$$P_{S_i}(t, f) = \left| \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} S_i(u) h^*(u-t) e^{-j2\pi fu} du \right|^2 \quad (6)$$

式(6)中, $h^*(\cdot)$ 是观测窗 $h(u-t)$ 的共轭函数,本文采用汉明(Hamming)窗,窗函数为:

$$h(u-t) = 0.54 - 0.46\cos\left[2\pi\left(\frac{u-t}{H}\right)\right] \quad (7)$$

式中: H 为汉明窗宽度,时刻 t 和 u 分别代表观测窗的移动和对观测窗内部的扫描,从而获得被分析信号的能量在时-频域内的分布。本文方法利用式(6)、(7)对受试者的 FP1、FPZ 等 62 个通道的 EM-EEG 进行 STPS 时-频分析,生成如图 7 所示的部分时-频谱图。

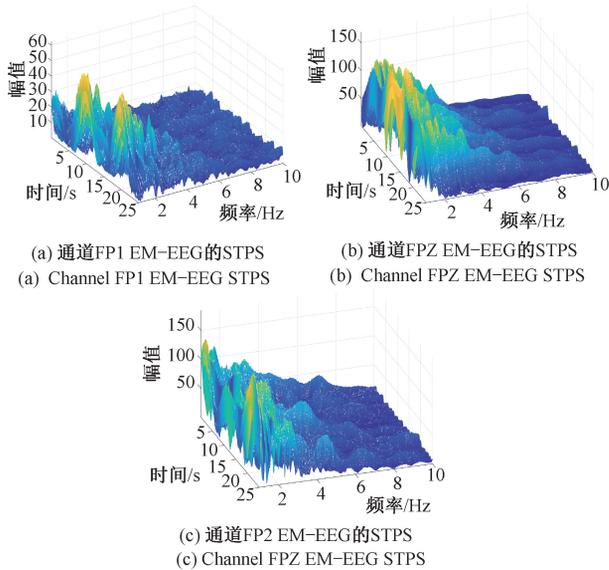


图 7 部分通道的 STPS 时-频谱图

Fig. 7 STPS time-spectrum diagram of some channels

由图 7 的 STPS 谱图,可以看到各通道 EM-EEG 在不同时刻和不同频率上的能量分布,为了描述 EM-EEG 通道的相关性,本文对各通道 EM-EEG 的 STPS 进行互相关运算,其互相关函数可以全面地描述两两通道之间的相似性,计算公式为:

$$\rho(P_{S_i}, P_{S_j}) = \frac{E[(P_{S_i}(t, f) - \overline{P_{S_i}})(P_{S_j}(t, f) - \overline{P_{S_j}})]}{\sqrt{\sum_{i=1}^{62} (P_{S_i} - \overline{P_{S_i}})^2} \sqrt{\sum_{j=1}^{62} (P_{S_j} - \overline{P_{S_j}})^2}} \quad (8)$$

$i, j = \{1, 2, 3, \dots, 62\}$

式中: $\rho(P_{S_i}, P_{S_j})$ 为任意两个通道 i, j 间 STPS 值的互相关系数,取值范围为 $[-1, 1]$, E 为数学期望, $\overline{P_{S_i}}$ 为所有 P_{S_i} 的平均值, $\overline{P_{S_j}}$ 为所有 P_{S_j} 的平均值。

利用式(8),计算出本文所使用的 62 路 EM-EEG 通道间 STPS 的 Pearson 相关系数构成的相关系数矩阵为:

$$\rho_{[]} = \begin{bmatrix} \rho_{P_{S_1}, P_{S_1}} & \rho_{P_{S_1}, P_{S_2}} & \cdots & \rho_{P_{S_1}, P_{S_{61}}} & \rho_{P_{S_1}, P_{S_{62}}} \\ \rho_{P_{S_2}, P_{S_1}} & \rho_{P_{S_2}, P_{S_2}} & \cdots & \rho_{P_{S_2}, P_{S_{61}}} & \rho_{P_{S_2}, P_{S_{62}}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{P_{S_{61}}, P_{S_1}} & \rho_{P_{S_{61}}, P_{S_2}} & \cdots & \rho_{P_{S_{61}}, P_{S_{61}}} & \rho_{P_{S_{61}}, P_{S_{62}}} \\ \rho_{P_{S_{62}}, P_{S_1}} & \rho_{P_{S_{62}}, P_{S_2}} & \cdots & \rho_{P_{S_{62}}, P_{S_{61}}} & \rho_{P_{S_{62}}, P_{S_{62}}} \end{bmatrix} = \begin{bmatrix} 1 & 0.229 & \cdots & 0.411 & 0.397 \\ 0.229 & 1 & \cdots & 0.356 & 0.251 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0.411 & 0.356 & \cdots & 1 & 0.241 \\ 0.397 & 0.251 & \cdots & 0.241 & 1 \end{bmatrix} \quad (9)$$

图 8 将式(9)中的矩阵以三维柱状图形式表示,可以量化各通道之间的相关程度。

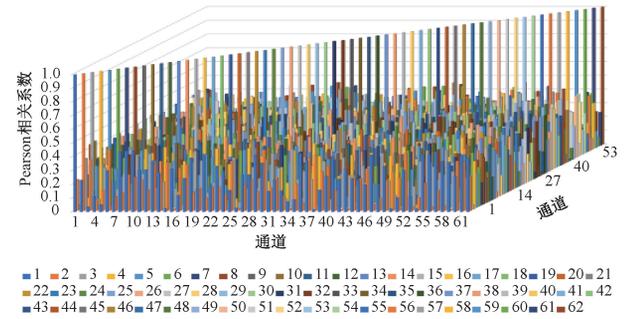


图 8 通道间互相关系数三维柱状图
Fig. 8 Three-dimensional histogram of the number of cross-relations between channels

2) 通道优化

从任一通道 i 与其他通道 $j(i \neq j)$ 的 STPS 互相关系数以及与自身的 STPS 自相关系数,反映出该通道含有的信息量和重要性,因此将通道 i 的所有 STPS 互相关系数和自相关系数的绝对值累计量,计算公式为:

$$P_{\sum S_i} = \sum_{j=1}^{62} |\rho(P_{S_i}(t, f) P_{S_j}(t, f))| \quad (10)$$

设置阈值 P_{\sum} ,对 $P_{\sum S_i} \geq P_{\sum}$ 的通道予以保留,同时忽略 $P_{\sum S_i} < P_{\sum}$ 的通道。本文的阈值 P_{\sum} 取值为 16.5,通道选择结果如图 9 所示。

最终选定分布在额叶、颞叶、枕叶等区域的 FP1、FP2、T7、T8、O1、O2 等 33 路通道,而省略顶叶等区域的 C1、C2、FC3、FC2 等 29 路通道,保留通道为原有通道的 53.23%。

3 实验过程及结果分析

3.1 实验条件及平台

1) 硬件平台:使用 Intel (R) Core (TM) i9-9900k

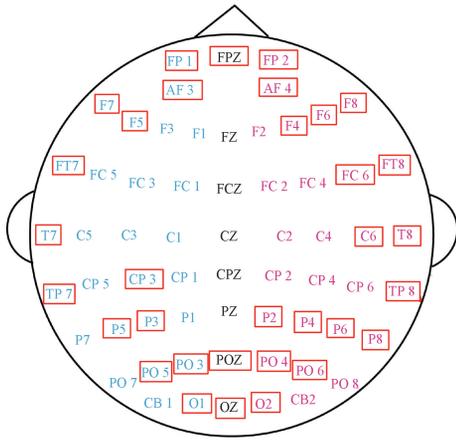


图 9 通道选择结果
Fig. 9 Channel selection result

3.36 GHz 处理器、32 GB 运行内存、GPU 采用 NVIDIA RTX 2080Ti。

2) 软件平台:使用 Python3.7 编程语言、采用 Keras 深度学习框架。

3.2 实验数据集

SEED IV 是由上海交通大学吕宝粮教授于 2015 年 10 月公开发布的基于 62 导采样频率为 200 Hz 的 EEG 情感计算数据集,包含平和、悲伤、恐惧和高兴 4 种情绪。数据集选择了 168 个电影视频片段作为 4 种情绪素材库,15 名健康右利手参与者(7 名男性和 8 名女性)参加实验。每次实验前,参与者都会被告知实验目的和步骤以及整个实验的无害性。每名参与者在不同的时间做了 3 次实验,每次实验使用的刺激完全不同,因此本数据集共有 45 个实验。每次实验每名参与者需要接受 24 个试验(每种情绪 6 个试验),试验过程如图 10 所示。每个电影剪辑都有一个 5 s 的开始提示和 45 s 的自我评估,在每个剪辑后使用积极消极情感量表(positive affect and negative affect scale, PANAS),根据评分来确定受试者在观看 2 min 电影片段的真实感受。目前使用 SEED IV 数据集的应用和研究机构数量已分别达 1 800 和 700 多个。关于 SEED IV 四分类情绪数据集详细介绍可参考文献[17]。

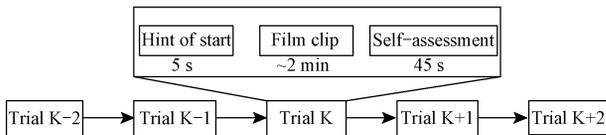


图 10 情绪采集实验过程
Fig. 10 Emotion collection experiment process

3.3 实验过程

1) 网络结构优化

首先,对所搭建的时-空融合 CNN 模型中的时间特征提取 CNN 和空间特征提取 CNN 分别进行了优化。在时间特征提取 CNN 中的优化对象为 L-Conv 的尺度,利用 PSO 算法对设置的尺度参数范围进行自动寻优,得到最终的 L-Conv 尺度,如表 2 所示。在空间特征提取 CNN 中主要优化对象为空间通道的数目,利用 STPS 的互相关技术对通道数目进行降维,最终结果如图 9 所示。

2) 网络超参数确定

本文采用网格搜索法对融合 CNN 网络模型的超参数进行自动选择,根据提前设置的超参数范围进行搜索查找来确定最优参数,如表 3 所示。第 1 列为需要优化的超参类别,包含学习率、正则化的权值系数、丢弃率、迭代次数和卷积层中卷积核的个数;第 2 列为超参数的预设空间,通过网格搜索自动调参算法,迭代选取最佳超参之后对网络进行动态训练;第 3 列为经验证集测试之后得到的最优参数。

表 3 网格搜索的最优超参

Table 3 Optimal super parameter for grid search

超参类别	预设空间	最优参数(时域 CNN/空域 CNN)
学习率	[0.001 0.005 0.01 0.015 0.02]	0.005/0.01
正则化系数 λ	[0.002 0.006 0.01 0.05 0.1]	0.01/0.01
第 1 层丢弃率	[0.1 0.2 0.3 0.4 0.5]	0.4/0.3
迭代次数	[50 100 150 200]	200/200
第 1 层卷积核个数	[32 64 128 256]	64/64
第 2 层卷积核个数	[32 64 128 256]	128/128
第 3 层卷积核个数	[32 64 128 256]	128/256
第 4 层卷积核个数	[32 64 128 256]	256/无
第 2 层丢弃率	[0.1 0.2 0.3 0.4 0.5]	0.3/0.2

3) 网络训练过程

本文在 SEED IV 数据集上使用本文提出的时-空融合 CNN 网络,并使用 PSO 算法将 4 个卷积层的 L-Conv 尺度 $L1, L2, L3, L4$ 作为优化参数在时域中进行优化,经过 200 次迭代训练最终训练集辨识准确率达到 93.56%,验证集达到 92.35%,训练过程的 loss 曲线最终趋于平稳,训练集 loss 低于 0.2,如图 11 所示。在时间优化的基础上又进行空间优化的时-空优化融合 CNN 网络训练过程如图 12 所示,训练集和验证集分别可以达到 92.4% 和 89.79%,loss 曲线降低并最终趋于稳定。

3.4 实验结果分析

使用本文所提方法对时域和空域分别进行优化之后的时-空优化融合 CNN 在经过 200 次训练之后得到的测试集准确率可以到达 90.13%,其输出混淆矩阵如表 4 所示。从表 4 中可以看出平和情绪的识别准确率最高,可以达到 95.56%,而恐惧情绪识别准确率最低,其中有

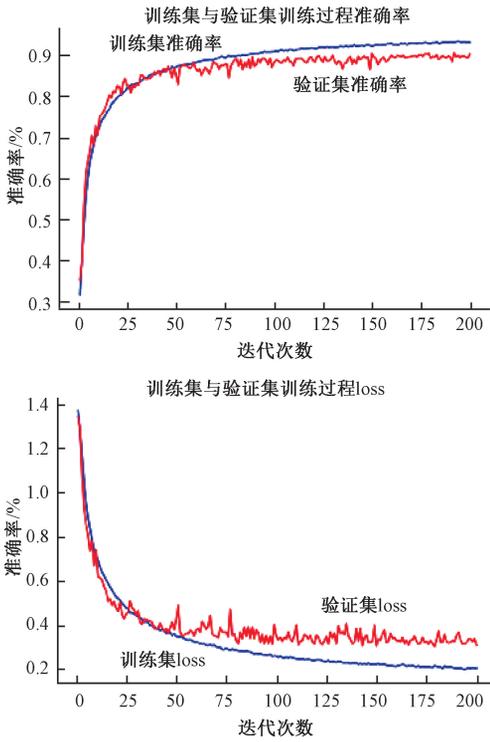


图 11 时域优化-空融合 CNN 训练 acc 和 loss 曲线

Fig. 11 Time domain optimization and space fusion CNN training acc and loss curves

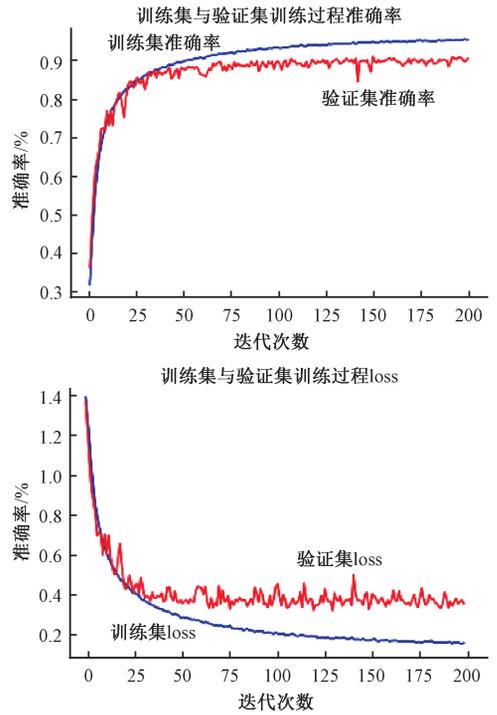


图 12 时-空优化融合 CNN 训练 acc 和 loss 曲线

Fig. 12 Time-space optimization fusion of CNN training acc and loss curve

8.05%的恐惧情绪被识别成平和情绪。

表 4 时-空优化融合 CNN 的测试集混淆矩阵

Table 4 Test set obfuscation matrix for time-space optimization fusion of CNN

真实	预测			
	平和	悲伤	恐惧	高兴
平和	95.56%	1.76%	1.34%	1.34%
悲伤	3.40%	89.36%	3.40%	3.84%
恐惧	8.05%	5.90%	83.58%	2.47%
高兴	4.15%	3.83%	2.31%	89.70%

本研究中搭建时-空优化融合 CNN 过程中的几种 CNN 网络进行了测试,最终得到的测试准确率对比如图 13 所示。从图中可以看出未经任何优化的时-空融合 CNN 相比单一 CNN 准确率提高了 5.5%;而在时-空融合 CNN 的基础上对 L-Conv 尺度进行了时域优化之后准确率最高,可以达到 92.35%;在时域优化的基础上又进行了空域的优化得到了本文所提出的时-空卷积优化融合 CNN,将原来的 62 路采集通道数目降低至 33 路,虽然其准确率略微下降,达到 90.13%,但是相对降低 46.77%的通道数目带来的优势而言可以忽略。此外对几种 CNN 的综合性能对比如表 5 所示,虽然经过时域优化后的时-空融合 CNN 识别准确率最高,但从综合性能角度考虑经

过时域和空域优化后的时-空融合 CNN 综合性能最优。

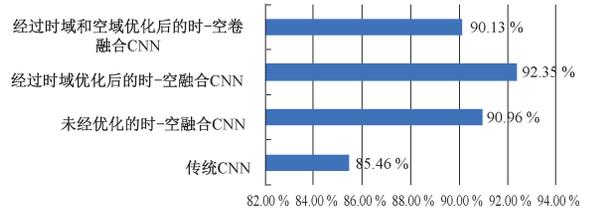


图 13 时-空优化前后各种 CNN 准确率对比

Fig. 13 Comparison of various CNN accuracy rates before and after time-space optimization

CNN网络结构	传统CNN	未经优化的时-空融合CNN	经过时域优化后的时-空融合CNN	经过时域和空域优化后的时-空卷积融合CNN
识别准确率	85.46%	90.96%	92.35%	90.13%

表 5 时-空优化前后各种 CNN 综合性能对比

Table 5 Comparison of comprehensive performance of CNN before and after time-space optimization

CNN 网络	精度	通道数	信息冗余度	运算时间/h
时域和空域优化后的时-空融合 CNN	较高	33	低	2.1
时域优化后的时-空融合 CNN	高	62	高	4.3
未经优化的时-空融合 CNN	较高	62	高	5.7
传统 CNN	低	62	高	5.5

此外,还将本文所提出的时-空优化融合 CNN 与其他文章方法进行了对比,如表 6 所示,由表可知,本文所提方法在降低了采集通道数目的基础上也有效保证了 EM-EEG 的识别准确率。

表 6 本文方法与其他实验方法对比

Table 6 The method in this paper is compared with other experimental methods

论文	分类方法	情绪状态	准确率/%
[20]	递归特征消去;SVM	控制、喜好、 唤醒、愉悦	78.49
[21]	CNN	积极、消极、中性	85.79
[22]	二次鉴别+K- 最邻近;SVM	开心、伤心、愤怒、平和	83.87
[23]	多尺度注意力 残差网络	压力、平静、轻松、 沮丧、快乐	85.20
[10]	CNN	愉悦度、唤醒度	87.27
本文	时-空优化融合 CNN	高兴、悲伤、恐惧、平和	90.13

4 结 论

本文针对在 EM-EEG 辨识研究中时域信息表征不明显和空间信息被忽略的问题,提出了时-空融合 CNN 网络,有效提取了 EM-EEG 的时间域和空间域特征并进行两类特征的融合。在公开数据集 SEED IV 上进行实验表明:相比传统单一 CNN 识别准确率提升了 5.5%,证实了时-空融合网络的可行性。此外,在时-空融合网络的基础上分别基于 PSO 算法和基于 STPS 的 Pearson 相关系数对时域上 L-Conv 尺度和空域 EM-EEG 采集通道分布进行了优化,空域优化后缩减了 46.77%的通道数目,且识别准确率依旧可以保持较高,达到 90.13%,证实了本文通道选择方法的有效性。

由于不同个体之间情绪的诱发会因为生理、心理等方面存在差异,而本文在 SEED IV 整体数据集上进行实验,忽略了个体间情绪差异,在未来的工作中可考虑构建更加优良的情感识别模型,解决个体差异问题。

参考文献

[1] 姜晓光, 陈兰岚, 宋振振. 基于多源域迁移学习的脑电情感识别[J]. 计算机工程与设计, 2020, 41(7): 2011-2018.
LOU X G, CHEN L L, SONG ZH ZH. EEG emotion recognition based on multi-source domain transfer learning[J]. Computer Engineering and Design, 2020, 41(7): 2011-2018.

[2] 陶建华, 李雅. 情感计算研究进展[J]. 中国计算机学会通讯, 2016, 12(10): 62-70.
TAO J H, LI Y. Advances in affective computing[J]. Communications of China Computer Federation, 2016, 12(10): 62-70.

[3] 孙颖, 马江河, 张雪英. 结合非线性全局特征和谱特征的脑电情感识别[J]. 计算机工程与应用, 2018, 54(17): 116-121.
SUN Y, MA J H, ZHANG X Y. EEG emotion recognition based on nonlinear global features and spectral feature [J]. Computer Engineering and Applications, 2018, 54(17): 116-121.

[4] 李娟, 刘国忠, 高洁. 基于脑电信号的情绪分类[J]. 北京信息科技大学学报(自然科学版), 2017, 32(2): 34-39.
LI J, LIU G ZH, GAO J. Emotion classification based on EEG signal[J]. Journal of Beijing Information Science and Technology University (Natural Science Edition), 2017, 32(2): 34-39.

[5] 王晴. 基于脑电信号的情绪分类[D]. 北京: 北京邮电大学, 2019.
WANG Q. Emotion classification based on EEG [D]. Beijing: Beijing University of Posts and Telecommunications, 2019.

[6] LIU Y, SOURINA O, NGUYEN M K. Real-time EEG-based emotion recognition and its applications [J]. Transactions on Computational Science, 2011, 12: 256-277.

[7] LIN Y P, WANG C H, JUNG T P, et al. EEG-based emotion recognition in music listening [J]. IEEE Transactions on Biomedical Engineering, 2010, 57(7): 1798-1806.

[8] BROWN L, GRUNDLEHNER B, PENDERS J. Towards wireless emotional valence detection from EEG [J]. IEEE Conference Proceeding on Engineering in Medicine and Biology Society, 2011, DOI: 10.1109/IEMBS.2011.6090412.

[9] 田曼, 杨风雷, 张艺. 基于熵的脑电特征选择情绪识别研究[J]. 电子测量技术, 2018, 41(19): 106-111.
TIAN M, YANG F L, ZHANG Y. Research on emotion recognition with EEG signal feature selection based on entropy [J]. Electronic Measurement Technology, 2018, 41(19): 106-111.

[10] RUI Q, QING C, TONG Z, et al. A novel deep-learning based framework for multi-subject emotion recognition [C]. International Conference on Information. IEEE, 2017: 181-185.

[11] MACHOT F A, ELMACHOT A, ALI M, et al. A deep-learning model for subject-independent human emotion recognition using electrodermal activity sensors [J]. Sensors, 2019, 19(7): 1659-1672.

[12] TRIPATHI S, ACHARYA S, SHARMA R D, et al. Using deep and convolutional neural networks for accurate emotion classification on deap dataset [C]. Twenty-Ninth

- IAAI Conference, 2017; 4746-4752.
- [13] KOELSTRA S, MUHL C, SOLEYMANI M, et al. DEAP: A database for emotion analysis; using physiological signals[J]. IEEE Transactions on Affective Computing, 2012, 3(1): 18-31.
- [14] MEI H, XU X. EEG-based emotion classification using convolutional neural network [C]. 2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), 2017; 130-135.
- [15] 李洋. 基于改进转移网络的非线性时间序列分析方法研究[D]. 济南: 济南大学, 2020.
LI Y. Research on nonlinear time series analysis method based on improved transfer network [D]. Jinan: University of Jinan, 2020.
- [16] 张绍荣, 赵紫宁, 莫云, 等. 特征提取对通道选择方法的影响研究[J]. 国外电子测量技术, 2020, 39(9): 1-6.
ZHANG SH R, ZHAO Z N, MO Y, et al. Research on the influence of feature extraction on channel selection method [J]. Foreign Electronic Measurement Technology, 2020, 39(9): 1-6.
- [17] ZHENG W L, LIU W, LU Y F, et al. EmotionMeter: A multimodal framework for recognizing human emotions[J]. IEEE Transactions on Cybernetics, 2018; 1-13.
- [18] KOELSCH S. A coordinate-based meta-analysis of music-evoked emotions [J]. NeuroImage, 2020, 223(2): 117350.
- [19] 白燕燕. 基于粒子群算法优化卷积神经网络结构[D]. 呼和浩特: 内蒙古大学, 2019.
BAI Y Y. Optimization of convolutional neural network structure based on particle swarms optimization [D]. Huhehot: Inner Mongolia University, 2019.
- [20] SHAO H M, WANG J G, WANG Y, et al. EEG-based emotion recognition with deep convolution neural network[J]. 2019 IEEE 8th Data Driven Control and Learning Systems Conference, 2019; 1125-1129.
- [21] ZHENG W L, LIU W, LU Y F, et al. EmotionMeter: A multimodal framework for recognizing human emotions[J]. IEEE Transactions on Cybernetics, 2018; 1-13.
- [22] MA A, ASSEM A A, HA H. EEG-based emotion recognition using a wrapper-based feature selection method [C]. International Conference on Advanced Intelligent Systems & Informatics, 2017; 247-256.
- [23] 柳长源, 孙雨涵, 李文强, 等. 基于 MAREsnet 的脑电情感识别研究[J]. 仪器仪表学报, 2020, 41(7): 235-242.
LIU CH Y, SUN Y H, LI W Q, et al. Research on EEG emotion recognition based on MAREsnet [J]. Chinese Journal of Scientific Instrument, 2020, 41(7): 235-242.

作者简介



黄永庆, 2019 年毕业于陕西科技大学获得学士学位, 现为陕西科技大学电气与控制工程学院在读研究生, 主要研究方向为脑电信号处理。

E-mail: 1906042@sust.edu.cn

Huang Yongqing received his B. Sc. degree in 2019 from Shaanxi University of Science and Technology, now he is a M. Sc. candidate in School of Electrical and Control Engineering, Shaanxi University of Science and Technology. His main research interest includes EEG signal processing.



周强 (通信作者), 1993 年获得学士学位, 2001 年获得硕士学位, 2010 年获得博士学位, 现为陕西科技大学教授, 主要研究方向为智能信息处理技术。

E-mail: 412698495@qq.com

Zhou Qiang (Corresponding author) received his B. Sc. degree in 1993, his M. Sc. degree in 2001 and his Ph. D. degree in 2010. Now he is a professor in School of Electrical and Control Engineering, Shaanxi University of Science and Technology. His main research interest includes intelligent information processing technology.