

DOI:10. 19651/j. cnki. emt. 2314602

# 基于多域信息融合的脑电情感识别研究

# 王泽田1 张学军1,2

(1.南京邮电大学电子与光学工程学院、柔性电子(未来技术)学院南京 210023;2.南京邮电大学射频集成与微组装技术国家地方联合工程实验室南京 210023)

**摘** 要:脑电信号识别方法较少将空间、时间和频率信息相融合,为了充分挖掘脑电信号包含的丰富信息,本文提出 一种多域信息融合的脑电情感识别方法。该方法利用二维卷积神经网络和一维卷积神经网络相结合的并行卷积神经 网络(PCNN)模型学习脑电信号的空间、时间和频率特征,来对人类情感状态进行分类。其中,2D-CNN用于挖掘相 邻 EEG 通道间的空间和频率信息,1D-CNN用于挖掘 EEG 的时间和频率信息。最后,将两个并行卷积模块提取的信 息融合进行情感识别。在数据集 SEED 上的情感三分类实验结果表明,融合空间、时间、频率特征的 PCNN 整体分类 准确率达到了 98.04%,与只提取空频信息的 2D-CNN 和提取时频信息的 1D-CNN 相比,准确率分别提高了 1.97%和 0.60%。并于最近的类似工作相比,本文提出的方法对于脑电情感分类具有一定的优越性。

关键词:脑电信号;多域信息融合;情感识别;并行卷积神经网络

中图分类号: TP391 文献标识码: A 国家标准学科分类代码: 510.40

# Research on EEG emotion recognition based on multi-domain information fusion

Wang Zetian<sup>1</sup> Zhang Xuejun<sup>1,2</sup>

 School of Electronic and Optical Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210023, China;
 Nation-Local Joint Project Engineering Lab of RF Integration & Micropackage, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

**Abstract**: EEG signal recognition methods rarely integrate spatial, temporal and frequency information, in order to fully explore the rich information contained in EEG signals, this paper proposes a multi-information fusion EEG emotion recognition method. The method utilizes a parallel convolutional neural network model (Parallel Convolutional Neural Network, PCNN) that combines a two-dimensional convolutional neural network (2D-CNN) and a one-dimensional convolutional neural network(1D-CNN) to learn the spatial, temporal, and frequency features of the EEG signals to categorize the human emotional states. Among them, 2D-CNN is used to mine spatial and frequency information of EEG. Finally, the information extracted from the two parallel CNN modules is fused for emotion recognition. The experimental results of emotion triple classification on the dataset SEED show that the overall classification accuracy of the PCNN fusing spatial, temporal, and frequency features reaches 98.04%, which is an improvement of 1.97% and 0.60%, respectively, compared to the 2D-CNN extracting only null-frequency information and the 1D-CNN extracting temporal-frequency information. And compared with recent similar work, the method proposed in this paper is superior for EEG emotion classification.

Keywords: EEG; multi-domain information fusion; emotion recognition; parallel convolutional neural network

# 0 引 言

情感是大脑高级认知功能之一,在人们日常生活中起 着重要的作用。积极的情绪能够提高工作学习的效率,有 益于身心健康,而消极的情绪可能会影响人的精神,甚至导 致健康问题<sup>[1]</sup>。在人与人之间的交互中,感知他人的情绪 很容易,但是要让机器理解人类复杂的情绪依然很难。随 着人工智能等技术的发展,情感识别已经成为了人机交互、

收稿日期:2023-09-17

模式识别和认知科学等领域的研究热点<sup>[2]</sup>,受到了学术界 和工业界的广泛关注。

情感信息的研究,主要分为非生理信号研究和生理信 号研究,其中非生理信号研究包括面部表情、语音以及身体 姿势等<sup>[3-4]</sup>,虽然这些特征很容易获取,但是容易被人故意 的掩饰,所以这些特征是不可靠的。生理信号研究主要包 括脑电图(EEG)、心电图(ECG)和肌电图(EMG)等,这些 信号都是人体对外部刺激所自发产生的神经生理信号,能 够客观地反映人的情感状态。近年来神经科学研究表明, 大脑中一些特定的区域与情感存在着密切的联系,因此越 来越多的研究者采用基于 EEG 的方法来进行情感识别 研究。

基于脑电信号的情感识别主要包括信号采集、数据预 处理、特征提取和情感分类4个步骤,其中最重要的是特征 提取和情感分类,如何从脑电信号中提取具有鉴别性的情 感特征以及如何建立更有效的情感分类计算模型是 EEG 情感识别领域中两个重要的技术挑战<sup>[5]</sup>。通常提取脑电信 号特征主要从时域、频域、时频域和空间域 4 个方面分 析<sup>[6]</sup>。时域方面通常是对脑电信号的时间变化特征来进行 分析,如信号统计量,高阶过零分析等。频域分析是对各个 频带的脑电信号进行分析研究,一般将脑电信号频带分为 5 类,分别为δ频带(1~3 Hz)、θ频带(4~7 Hz)、α频带 (8~13 Hz), β频带(14~30 Hz)以及 γ1 频带(31~50 Hz)。 Duan 等<sup>[7]</sup> 在这5个频带中提取了功率谱密度(power spectral density, PSD)、微分熵(differential entropy, DE) 等特征分别用于情感识别,发现微分熵特征的效果最好,还 证实了高频带上的 EEG 信号与情感状态的关系比其他频 带更密切。时频域分析通常使用短时傅里叶变换 (shorttime Fourier transform, STFT), 小波变换 (wavelet transform, WT) 和小波包变换(wavelet packet transform, WPT)来进行特征提取。刘珑等<sup>[8]</sup>对基于小波变换和小波 包变换的脑电信号的滤波降噪方法,和基于小波包变换的 脑电信号特征提取方法进行了研究,验证了基于小波包变 换的降噪方法的优越性和特征提取方法的有效性。空间域 分析主要包括脑电通道电极的位置信息和脑区的不对称信 息,脑电通道电极的位置信息可以通过二维空间矩阵<sup>[9]</sup>、等 距方位投影<sup>[10]</sup>和 EEG 图片<sup>[11]</sup>等方法来构造,脑区的不对 称特征主要是由左右大脑或前后大脑之间的差异信息表 示,如微分不对称(differential asymmetry, DASM)特征、 理性不对称(rational asymmetry, RASM)特征和微分尾部 (differential caudality, DCAU)特征。目前,大多数方法都 是以脑电信号的频率特征为主,经研究表明,DE 是脑电情 感识别领域中最适合和最稳定的特征之一[12]。

综上,传统的 EEG 特征提取主要是从时域、频域、时频 域和空间域 4 个方面的一种来分析,在多域特征融合上的 研究较少。近年来,随着计算能力的提高,深度学习的技术 被越来越多的研究者应用到了情感识别领域,因此多域特

征融合也逐渐成为 EEG 特征提取方面的研究热点。Yang 等<sup>[13]</sup>提出了一种脑电信号的三维表示方法(CCNN),在保 留通道间空间信息的同时,又结合了不同频带信号的特征, 在 DEAP 数据集的唤醒度和效价二分类平均准确率分别 达到 90.24%和 89.45%。Wang 等<sup>[14]</sup>提出了一种名为 EmotioNet 的 EEG 三维卷积神经网络(convolutional neural network, CNN),利用 3D 卷积核从原始 EEG 信号 中同时提取空间和时间特征,在 DEAP 数据集的唤醒度和 效价两个维度的准确率分别达到 73.1%和 72.1%。Yang 等<sup>[15]</sup>提出一种并行卷积循环神经网络(PCRNN),其利用 CNN 模块将链状脑电序列转换为二维矩阵来挖掘物理相 邻脑电信号之间的通道位置信息,利用长短时记忆网络 (long short-term memory, LSTM)模块挖掘脑电时间信 息,将两模块得到的空时特征融合进行情感分类,在 DEAP 数据集的唤醒度和效价二分类平均准确率分别达到 91.03%和 90.80%。Li 等<sup>[16]</sup>提出了一种新的双半球差异 模型(BiHDM)来学习两个脑区半球之间的不对称差异,在 基于 SEED 数据集的情感三分类平均准确率达到了 93.12%。苗敏敏等[17]提出一种自适应优化空频微分熵 (AOSFDE)特征,通过稀疏回归算法对多重局部空间-频域 内的 DE 特征进行优化选择,在基于 SEED 数据集的积极/ 消极、积极/中性、中性/消极这3个情绪二分类场景平均准 确率分别达到 91.8%,93.3%,85.1%。Gao 等<sup>[18]</sup>提出一 种基于多层卷积神经网络(multilayer convolutional neural network, MCNN)并结合 DE 和脑网络的情感分类方法, 利用连续小波变换得到 EEG 的时频表示,之后在每个频段 构建脑网络,并在 DE 的提取过程保留了通道之间的位置 的信息,最终准确率达到了91.45%。

上述的方法中,大多都只融合了脑电信号中的一种或 两种信息,为了更加充分的提取脑电信号蕴含的多种信息, 本文提出了一种基于并行卷积神经网络(parallel convolutional neural network, PCNN)的多域信息融合 EEG 情感识别方法,将脑电信号的空间、时间和频率信息 相融合作为情感识别网络模型的输入。该方法首先将脑电 信号进行分段,并分别构造空频特征结构和时频特征结构。 对于空频特征结构:分别提取5个频带的DE特征,再根据 电极空间位置转化为二维特征矩阵,最后将5个频带的二 维特征矩阵堆叠构造成三维矩阵。对于时频特征结构:对 分段的脑电信号进一步按照特定时间窗口计算出不同时刻 的5个频带的 DE 特征,再按照时间顺序组织成一维特征 向量,然后进行滑动平均滤波(moving average filtering, MAF)消除特征不稳定性。之后,针对两种特征结构搭建 了 2D-CNN 和 1D-CNN 网络,分别学习脑电信号的空频信 息和时频信息,最后将两种 CNN 模型提取的信息融合进 行脑电情感分类。在 SEDD 公共数据集上对所提方法进行 评估和讨论,并通过对比实验,证明了该方法对情感分类的 有效性。

# 1 数据与方法

#### 1.1 数据集和数据预处理

本文选用上海交通大学的 SEED 脑电情感数据集<sup>[12]</sup>, 数据集的构建流程如图 1 所示。数据集包含平均年龄在 23 岁左右的 15 名受试者(7 男 8 女)的脑电数据,采用国际 10~20 体系,以1000 Hz 的采样率从 62 个脑电电极记录 EEG。每个受试者在 3 个不同的时间段分别采集脑电数 据,每个时间段间隔1周左右。在每个时间段的数据采集 中,受试者观看15个电影片段,每个电影片段约4min左 右,其中积极、中性和消极的电影片段各5个。此外,还在 受试者观看电影片段后对每个受试者进行额外的45s主观 自评,以确保所收集的脑电数据与呈现给受试者的电影片 段具有相同的情感状态,在自我评估后休息 15 s,然后观看 下一个视频。对采集后的数据进行去噪处理后下采样至 200 Hz,并应用 0~75 Hz 的带通滤波器进行滤波。研究表 明,高频的 EEG 信号对情感识别的影响大于低频信号<sup>[7]</sup>, 本文中,不同于以往文章只提取最大到 50 Hz 的频带,而也 将 50~75 Hz 的频带应用到本文方法中。为了保证数据格 式的统一,截取每个脑电数据的其中120s作为实验数据, 预处理后的 SEED 数据集格式如表 1 所示。



图 1 SEED 数据集的构建流程

	表 1	表	1 划分	·后的	SEED	<b>致</b>	集格	) えんし しんし しんし しんし しんし しんし しんし しんし しんし しんし
--	-----	---	------	-----	------	----------	----	---

名称	形状	组成
数据	$45\! imes\!62\! imes\!24$ 000	视频×通道×采样点
标签	$45 \times 1$	视频×标签

#### 1.2 特征构造

1) 微分熵特征

微分熵是一种脑电情感识别领域中常用且有效的特征,它是香农信息熵在连续变量上的推广形式<sup>[6]</sup>,微分熵的 计算公式为:

$$h(X) = -\int_{X} f(x) \log(f(x)) dx$$
(1)

其中,X 表示时间序列,f(x)表示连续信息的概率密 度函数。对于一段固定长度的近似服从  $N(\mu,\sigma^2)$  高斯分 布的脑电信号,其微分熵可以表示为<sup>[7,12]</sup>:

$$h(X) = -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(x-\mu)^2}{2\sigma^2} \log \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(x-\mu)^2}{2\sigma^2} \log \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(x-\mu)^2}{2\sigma^2} dx = \frac{1}{2} \log 2\pi e \sigma^2$$
(2)

其中, e 是欧拉常数,  $\mu$ 和  $\sigma$ 是 x 的均值和标准差。

2) 空间-频率特征构造

为了增加训练数据量,将原始的120s脑电信号以Ts 划分片段作为一个样本,且无重叠,划分后的每一片段与原 始试验分配相同的标签。

空间-频率特征构造的步骤为:(1)将划分片段后的样本 根据式(2)分别计算出 $\theta(4 \sim 7 \text{ Hz}), \alpha(8 \sim 13 \text{ Hz}), \beta(14 \sim 13 \text{ Hz})$ 30 Hz)、y1(31~50 Hz)、y2(51~75 Hz)5 个频带的 DE 特 征。(2)为了保持脑电电极位置的空间结构信息,采用构造 二维矩阵的方法,根据脑电电极的空间位置和不同电极之 间的相对位置,将 62 电极脑电信号的一维 DE 特征向量转 换为二维 DE 特征矩阵,对于没有值的位置采用 0 进行填 充,因为0值对于后续神经网络权值的调整不起作用,所以 添加 0 值对神经网络的识别分类也不会有影响<sup>[19]</sup>,其转换 过程如图 2 所示。(3)为了结合脑电信号不同频率的信息, 对于来自每个脑电片段不同频带的5个二维特征矩阵,将 它们堆叠起来构成三维特征矩阵<sup>[13]</sup> $X_n \in \mathbb{R}^{h \times w \times b}, n = 1,$  $2, \dots, N, N$  表示总样本的数量, h 和 w 的值分别表示电极 在垂直、水平方向上的最大值,b 表示不同频带的数量,此 三维特征矩阵包含了脑电信号的空间和频率信息,构造过 程如图 3 所示。在本文中,h,w 和b 的值分别为 9,9,5。



图 2 二维空间特征矩阵的构造

#### 3) 时间-频率特征构造

时间-频率特征构造的步骤为:(1)将划分后的每一个时 间长度为 *T*s 的脑电样本,进一步使用长度为 *W*s 的时间窗 口将其分割为 *T*/W 个小片段,并确保连续相邻片段之间没 有重叠。(2)对于每一段样本中的 *T*/W 个 *W*s 的脑电片段, 计算出其 $\theta$ , $\alpha$ , $\beta$ , $\gamma$ 1, $\gamma$ 2 这 5 个频带的 DE 特征。(3)将所有 通道的每一个频带的 DE 特征按照时间顺序构成一维特征 向量 *Y*<sub>n</sub>  $\in \mathbb{R}^{62\times k\times \frac{T}{W}}$ , *n*=1, 2,…, *N*,*N* 表示总样本的数量, 此一维特征向量包含了脑电信号的时间和频率信息。在本 文中,样本片段长度 *T* 和时间窗口长度 *W* 分别为 6 和 1。



图 3 三维空频特征矩阵的构造

由于脑电信号具有非平稳性的特点,构造的一维时频 特征向量也会存在一些波动较大的现象,为了减小一维时 频特征向量的波动,从而使特征向量更加平滑稳定,本文采 用了滑动平均滤波的方法。滑动平均滤波也称移动平均滤 波,是一种有限冲激响应型(FIR)低通滤波器,其原理是对 一段序列数据在固定长度的滑动窗口中做局部平均,以消 除或减小序列数据的波动。滑动平均滤波的一般表达 式为:

$$x(k) = \frac{1}{N} \sum_{i=0}^{N-1} y(k+i)$$
(3)

经过多次实验比较,最终选取滑动窗口长度 N 为 20。 图 4(a)、(b)分别为某一段时频特征向量滑动平均滤波前 后的结果,由此可见,滑动平均滤波有效的抑制了原始特征 向量的波动,并且滤波后的特征向量很好的保留了原始特 征向量的趋势,与常用的累加平均算法比较,滑动平均滤波 具有更快的计算速度。

#### 1.3 并行卷积神经网络模型

本节提出了一种基于 1D-CNN 和 2D-CNN 的时空频 特征融合的 PCNN 模型。该模型包括时频特征提取模块、 空频特征提取模块以及特征融合分类模块。其中,利用 1D-CNN 对构造的时频特征序列进行建模,利用 2D-CNN 挖掘重构的多通道脑电信号的空频信息。在此基础上,将 两个模块得到的深度特征进行融合,最终输入到分类器中 进行情感识别。图 5 描述了并行卷积神经网络模型的 结构。

1) 1D-CNN 网络模型

1D-CNN 网络一共有 3 个卷积层、3 个最大池化层、 1 个展平层、1 个 Dropout 层和 1 个全连接层。卷积层通过 参数共享和局部感知来挖掘上一层的特征,卷积层的核数 量依次为 16、32、64,所有卷积层的卷积核大小都为 18,卷 积步长都为 1,采用 ReLU 函数作为激活函数。在每一个 卷积层后都有一个最大池化层,用来控制参数量,减小计算 成本,池的大小为 2,步长也为 2。此外,在每个池化层后都 添加了批归一化(batch normalization, BN)操作,BN 可以 加快训练速度,并且还有轻微正则化效果,可以抑制过拟



图 4 滑动滤波前后的时频特征向量

合。最后将所得特征进行展平后输入到全连接层,为了防止过拟合并且提高模型泛化能力,在展平层后加入了 Dropou层,并在每一个卷积层中应用了L2正则化。

1D-CNN 网络模型的输入是构造的一维时频特征向量,其第 *n* 个输入特征向量表示为 *Y<sub>n</sub>* ∈ ℝ<sup>1860</sup>,将其输入到 1D-CNN 网络中:

$$\boldsymbol{P}_{n} = Con v 1 D \left( \boldsymbol{Y}_{n} \right) \tag{4}$$

其中,*Conv*1D 表示一维卷积神经网络运算。然后,将 输出的特征输入到全连接层,得到时频特征向量:

$$\mathbf{TF}_{n} = Dense\left(\mathbf{P}_{n}\right), \mathbf{TF}_{n} \in \mathbb{R}^{1\,024}$$
(5)

其中, Dense 表示全连接层运算。

2) 2D-CNN 网络模型

2D-CNN 网络一共有 4 个连续的卷积层、1 个展平层、 1 个 Dropout 层和 1 个全连接层。卷积层的核数量依次为 64、128、256、64,卷积步长都为 1×1,采用 ReLU 函数作为 激活函数,使用零填充模式防止边缘信息的丢失,对于前 3 个卷积层,卷积核的大小都设置为 4×4,以挖掘更多通道 间信息,第 4 个卷积层采用了 1×1 的卷积核,用来融合前 一卷积层的特征图并降低计算成本。在每一次卷积完成后 都添加了 BN 操作。另外,由于池化的主要功能是以一定 的信息损失为代价降低数据维数,而三维矩阵的尺寸较小,





故该模型没有采用池化的操作。与 1D-CNN 网络模型相同,2D-CNN 网络模型也在展平层后加入了 Dropou 层,并 在每一个卷积层中应用了 L2 正则化,以防止过拟合。

2D-CNN 网络模型的输入是构造的三维空频特征矩阵,其第 n 个输入特征矩阵表示为 $X_n \in \mathbb{R}^{9\times9\times5}$ ,将其输入到 2D-CNN 网络中:

$$\boldsymbol{Q}_{n} = Conv2D\left(\boldsymbol{X}_{n}\right) \tag{6}$$

其中,*Conv2D* 表示二维卷积神经网络运算。然后,将输出的特征输入到全连接层,得到空频特征向量 *SF*<sub>n</sub>:

 $SF_{n} = Dense(Q_{n}), SF_{n} \in \mathbb{R}^{1024}$ (7)
3) 特征融合及分类

将上述两种模型同时提取的时频特征 *TF*<sub>\*</sub> 和空频特 征 *SF*<sub>\*</sub> 进行融合最终得到具有时间、空间和频域信息的特 征向量 *TSF*<sub>\*</sub>:

**TSF**<sub>n</sub> = concatenate (**TF**<sub>n</sub>, **SF**<sub>n</sub>), **TSF**<sub>n</sub> ∈ ℝ<sup>2048</sup> (8) 其中, concatenate 表示拼接操作。为了防止过拟合并 提高模型的泛化能力, 在拼接操作后加入了 Dropout 层。 并将拼接后的特征向量 **TSF**<sub>n</sub> 进一步输入到全连接层:

$$\boldsymbol{V}_n = Dense\left(\boldsymbol{TSF}_n\right), \boldsymbol{V}_n \in \mathbb{R}^{256}$$
(9)

最后,将特征向量 $V_n$ 输入到 softmax 分类器进行情感 三分类:

 $\boldsymbol{R}_{n} = softmax\left(\boldsymbol{V}_{n}\right), \boldsymbol{R}_{n} \in \mathbb{R}^{3}$ (10)

#### 2 实 验

#### 2.1 实验设置

本文实验采用的硬件环境为 NVIDIA Geforce GTX 1650 GPU,软件环境为 Windows 10 操作系统,Python 3.8 编程语言环境,使用 Keras 深度学习框架库搭建网络模型。 采用 Adam 优化器最小化交叉熵损失函数,学习率初始值 设置为 0.001,随后每当连续 10 个 epoch 损失没有下降 时,学习率降为原来的 1/2,epoch 为 150,每次迭代的批大 小为 128。采用所有受试者参与的整体分类实验,此方式 使用多个受试者样本参与模型的训练和测试,更适合用于 情感脑机接口的分析和应用<sup>[20]</sup>,随机打乱所有受试者的脑 电信号段样本,划分训练集和测试集的比例为8:2。

## 2.2 结果分析

图 6 和 7 分别为 PCNN 网络模型的准确率(Accurary) 和损失(Loss)随 epoch 迭代的变化曲线图。从图中可以看 出,准确率和损失在第 100 个 epoch 之前都有着比较剧烈 的波动,尤其在第 40 个 epoch 左右出现两次较大的波动, 这可能是由于 1D-CNN 和 2D-CNN 两种网络模型分别学 习不同的特征结构融合后的收敛状况不同导致的。此外, 在第 100 个 epoch 之后网络逐渐趋于平稳,在 120 个 epoch 达到平稳状态。



## 1) 消融研究

首先,为了验证滑动平均滤波的方法对 1D-CNN 模型 的性能影响,比较了一维时频特征滑动平均滤波前后的情 感分类准确率,如图 8 所示。可以看出,时频特征经滑动 平均滤波后,输入 1D-CNN 模型的识别准确率达到了 97.44%,比未滤波的识别准确率提高了 1.55%,说明滑动 平均滤波能够有效的消除特征的不稳定性。综合上述章 节的直观图像和评定数据,本文采用的滑动平均滤波在保 留了与原始时频特征相似度的基础上,提高了 1D-CNN 模



图 7 PCNN 网络模型损失曲线

型的情感识别准确率。下文相关模型输入均采用经滑动 平均滤波后的时频特征。



图 8 1D-CNN 模型滑动平均滤波前后的识别准确率

下面进一步从精确率(Precision)、召回率(Recall)和 F1分数(F1 score)3个指标来更全面评价模型的性能。其 中,准确率表示在所有样本中被预测为正确的样本概率, 精确率表示在被所有预测为正的样本中实际为正样本的 概率,召回率表示在实际为正的样本中被预测为正样本的 概率,F1分数则同时考虑精确率和召回率,被定义为精确 率和召回率的调和平均数,评价指标的具体公式为:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
(11)

$$Precision = \frac{TP}{TP + FP}$$
(12)

$$Recall = \frac{TP}{TP + FN}$$
(13)

F1 score = 
$$\frac{2 \times Precision \times Recall}{Precision + Recall}$$
(14)

其中,*TP* 表示预测为正且实际为正,*TN* 表示预测为 负且实际为负,*FP* 表示预测为正但实际为负,*FN* 表示预 测为负但实际为正。

为了证明 PCNN 网络模型的有效性,比较了 2D-

CNN、1D-CNN和PCNN3种网络模型在上述不同指标下的性能,如表2所示。其中,2D-CNN仅用于挖掘EEG的空间、频率信息进行分类,1D-CNN仅用于挖掘EEG的时间和频率信息进行分类。由实验结果可知,结合了2D-CNN和1D-CNN的PCNN网络模型在4个指标中的性能都是最好的,准确率、精确率、召回率和F1分数分别达到了98.04%、98.06%、98.01%和98.03%,比2D-CNN模型分别提高了1.97%、1.88%、1.96%和1.92%,比1D-CNN模型分别提高了0.60%、0.52%、0.54%和0.53%。因此,相比只融合空间、频率信息的2D-CNN网络和时间、频率信息的1D-CNN网络,融合空间、时间、频率3种脑电信息的PCNN网络可以更进一步提高脑电情感识别的性能,由此证明本文所提PCNN网络方法的有效性。

#### 表 2 不同网络模型的评价指标值

模型	准确率	精确率	召回率	F1 分数
2D-CNN	96.07	96.18	96.05	96.11
1D-CNN	97.44	97.54	97.47	97.50
PCNN	98.04	98.06	98.01	98.03

2) 混淆矩阵

为了观察不同网络模型对不同情感的识别效果,引入 了混淆矩阵,图 9(a)、(b)、(c)分别为 1D-CNN、2D-CNN 和 PCNN 3 种网络模型的混淆矩阵结果,从图中可以得到以 下两点:

(1) 通过对比发现,本文所提出的 PCNN 网络模型比 其它两种模型在 3 类情感中的识别效果都要好,且在 3 类 情感中均表现良好,每类情感的正确识别率均在 97%以 上,说明本文所提 PCNN 网络模型的有效性。

(2)通过比较发现,积极的情感比中性和消极的情感 更容易被识别,且准确率都达到了97%以上,此外,消极和 中性情感更容易被混淆,也许受试者更易被积极的情感刺 激材料激发出更明显的感受。

3) 同类相关研究对比

将本文的方法与近几年使用 SEED 情感数据集的已 有研究进行对比,结果如表 3 所示。文献[9]将不同通道的 微分熵特征组织成单频带二维空间图来训练分层卷积神 经网络,最终准确率达到了 88.20%,该方法只考虑了 EEG 的空间信息。文献[21]提出一种基于短时傅里叶变换的 电极频率分布图(electrode-frequency distribution maps, EFDMs),采用基于残差块的深度卷积神经网络用于学习 特征并分类,最终准确率达到了 90.59%,该方法从 EEG 的时频方面进行分析。文献[23]提出了一种卷积+ Transformer 结构的网络,卷积模块学习时空局部特征,后 续连接 Transformer 自注意模块,提取局部时间特征内的 全局相关性,最终准确率达到了 95.30%,该方法从 EEG

%



图 9 不同网络模型的混淆矩阵

的时空方面进行分析。从表 3 中可以看出,本文所提 PCNN方法的情感分类准确率达到了 98.04%,均高于以 上方法,这大概是由于以上方法只考虑了空间、时间、频率 信息中的一种或者两种,而基于以上 3 种 EEG 特征信息的 PCNN 网络更能挖掘出深层次的情感差异。

此外将本文方法与其他融合空间、时间和频率信息的 已有研究成果进行分析对比。文献[22]提出了一种四维 DE 特征矩阵形式,整合了 EEG 的空间、时间和频率信息,

子测量技术

电

表 3 同类相关研究对比

参考文献	方法	信息	准确率/%
文献[9]	HCNN	空	88.20
文献[21]	EFDMs+CNN	时+频	90.59
文献[22]	4D-CRNN	空+时+频	94.74
文献[23]	Conformer	空十时	95.30
文献[24]	AMDET	空+时+频	97.17
本文	PCNN	空+时+频	98.04

并采用 CNN+LSTM 的串行顺序方式进行情感识别,最 终准确率达到了 94.74%,与本文相差 3%以上,导致差异 的原因可能是其四维矩阵较复杂,冗余信息也较多,另一 个重要的原因可能是频带的选取不同导致。文献[24]提 出了一种基于注意力的多维脑电 Transformer 深度模型, 将原始脑电图数据转换为三维时-频-空表征,然后利用频 谱-空间 Transformer 编码层提取脑电图信号中的有效特 征,并对关键时间帧进行时间注意层处理,最终准确率达 到了 97.17%,与本文相差较小,且与本文所选用频带一 致,导致差异的一个可能的原因是其空间通道的表征为一 维方式,没能充分利用脑电的空间信息,但此文的注意力 机制值得后续深入研究。

综上所述,本文方法在 SEED 数据集上取得了较好的 效果,较运用 SEED 数据集的已有研究有一定程度提高, 验证了本文方法的优越性,同时也说明了融合脑电信号多 域信息进行分类的有效性和必要性。

# 3 结 论

为了充分挖掘情感脑电信号中的信息,本文提出了一 种包含空间、时间和频率3个脑电维度信息的 PCNN 情感 识别模型。首先,利用微分熵构造出两种不同的特征结 构,其中空频特征结构是根据电极通道位置将多通道一维 特征向量转换为二维特征矩阵,并将不同频带的二维特征 矩阵堆叠构造成三维空频特征矩阵,充分的考虑了脑电通 道的空间位置和频率信息;时频特征结构是按照时间顺序 将所有通道的不同频带的微分熵特征构成一维向量,充分 的考虑了脑电信号的时间关联性,又考虑到一维时频特征 存在不平稳波动,再对其进行滑动平均滤波处理以消除特 征不稳定性。其次,采用 PCNN 模型分别对构造的空频特 征和时频特征进行深层次的学习。最后,将空频信息和时 频信息进行融合并分类。通过实验对比,结果表明融合两 种网络的 PCNN 模型的效果要好于其中任何一种,说明了 该方法的有效性。并与最近基于同一数据集的方法作了 比较,表明了该模型在情感分类中具有不错的准确性和稳 定性。

# 参考文献

[1] SHU L, XIE J Y, YANG M Y, et al. A review of emotion recognition using physiological signals [J]. Sensors(Basel), 2018, 18(7): 2074.

- [2] CALVO R A, D' MELLO S. Affect detection: An interdisciplinary review of models, methods, and their applications [J]. IEEE Transactions on Affective Computing, 2010, 1(1): 18-37.
- [3] JAIN D K, SHAMSOLMOALI P, SEHDEV P. Extended deep neural network for facial emotion recognition[J]. Pattern Recognition Letters, 2019, 120: 69-74.
- [4] LIU Z T, WU M, CAO W H, et al. Speech emotion recognition based on feature selection and extreme learning machine decision tree[J]. Neurocomputing, 2018, 273: 271-280.
- LI Y, ZHENG W, WANG L, et al. From regional to global brain: A novel hierarchical spatial-temporal neural network model for eeg emotion recognition[J].
   IEEE Transactions on Affective Computing, 2022, 13(2): 568-578.
- [6] 张冠华,余旻婧,陈果,等. 面向情绪识别的脑电特 征研究综述[J]. 中国科学:信息科学,2019,49(9): 1097-1118.
- [7] DUAN R N, ZHU J Y, LU B L. Differential entropy feature for EEG-based emotion classification [C]. International IEEE/EMBS Conference on Neural Engineering, 2013: 81-84.
- [8] 刘珑,李胜,王轶卿. 基于小波包变换的脑电波信号 降噪及特征提取[J]. 计算机工程与科学,2015,37(4):790-795.
- [9] LI J P, ZHANG Z X, HE H G. Hierarchical convolutional neural networks for EEG-based emotion recognition[J]. Cognitive Computation, 2018, 10(2): 368-380.
- [10] 王璐,梁明晶,石慧宇,等.基于脑电时频空多域特 征融合的情感识别研究[J].计算机工程与应用, 2021:1-7.
- [11] LI Y J, HUANG J J, ZHOU H Y, et al. Human emotion recognition with electroencephalographic multidimensional features by hybrid deep neural networks[J]. Applied Sciences, 2017, 7(10): 1060.
- [12] ZHENG W L, LU B L. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks [J]. IEEE Transactions on Autonomous Mental Development, 2015, 7(3): 162-175.
- [13] YANG Y, WU Q, FU Y, et al. Continuous convolutional neural network with 3D input for EEGbased emotion recognition [C]. The 25th International Conference on Neural Information Processing, 2018.
- [14] WANG Y, HUANG Z Y, MCCANE B. EmotioNet: A 3-D convolutional neural network for EEG-based emotion recognition [C]. 2018 International Joint

Conference on Neural Networks(IJCNN), 2018: 1-7.

- [15] YANG Y L, WU Q F, QIU M, et al. Emotion recognition from multi-channel EEG through parallel convolutional recurrent neural network [C]. 2018 International Joint Conference on Neural Networks (IJCNN), 2018: 1-7.
- [16] LI Y, ZHENG W M, WANG L, et al. A novel bihemispheric discrepancy model for EEG emotion recognition[J]. IEEE Transactions on Cognitive and Developmental Systems, 2021, 13(2): 354-367.
- [17] 苗敏敏,徐宝国,胡文军,等. 基于自适应优化空频 微分熵的情感脑电识别[J]. 仪器仪表学报,2021, 42(3):221-230.
  MIAO M M, XU B G, HU W J, et al. Emotion EEG-recognition based on the adaptive optimized spatial-frequency differential entropy [J]. Chinese Journal of Scientific Instrument, 2021, 42 (3): 221-230.
- [18] GAO Z K, LI R M, MA C, et al. Core-brainnetwork-based multilayer convolutional neural network for emotion recognition [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-9.
- [19] 廉小亲,罗志宏,蔡沫豪,等. 基于卷积神经网络的脑电 情绪识别方法[J]. 计算机仿真, 2022, 39(8): 268-274.
- [20] 董寅冬,任福继,李春彬.基于线性核主成分分析和 XGBoost 的脑电情感识别[J].光电工程,2021, 48(2):15-23.
- [21] WANG F, WU S C, ZHANG W W, et al. Emotion recognition with convolutional neural network and EEG-based EFDMs [J]. Neuropsychologia, 2020, 146: 107506.
- [22] SHEN F Y, DAI G J, LIN G, et al. EEG-based emotion recognition using 4D convolutional recurrent neural network[J]. Cogn Neurodyn, 2020, 14(6): 815-828.
- [23] SONG Y H, ZHENG Q Q, LIU B C, et al. EEG conformer: Convolutional transformer for EEG decoding and visualization [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2023, 31: 710-719.
- [24] XU Y L, DU Y, LI L, et al. AMDET: Attention based multiple dimensions EEG transformer for emotion recognition [J]. IEEE Transactions on Affective Computing, 2023; 1-11.

#### 作者简介

**王泽田**,硕士研究生,主要研究方向为智能信息处理、脑 机接口技术。

E-mail:1021020910@njupt.edu.cn

**张学军**,教授,硕士生导师,主要研究方向为智能信息处理、脑机接口技术、无线射频识别技术等。 E-mail:xjzhang@njupt.edu.cn