

DOI:10.19431/j.cnki.1673-0062.2023.06.006

基于 CNN-BiLSTM 的脑电情绪分类模型及验证

杨晓同¹, 易灿南², 蒋复量^{1,2*}, 胡 鸿², 张 勉¹, 李昊宇¹, 吴 文²

(1. 南华大学 资源环境与安全工程学院, 湖南 衡阳 421001; 2. 湖南工学院 安全与管理工程学院, 湖南 衡阳 421002)

摘 要:情绪在人们的思考、行为和交流方式中起着重要作用。为提高脑电信号的情绪识别准确率,充分利用脑电信号的频率、空间和时间维度上的信息,提出一种基于 CNN-BiLSTM(convolutional neural networks-bidirectional long short term memory)的脑电情绪分类神经网络模型。该模型由卷积神经网络和多层特征融合的双向长短时神经网络构成,卷积神经网络用于学习脑电信号的频率和空间特征,双向长短时神经网络则从卷积神经网络的输出中挖掘脑电切片之间的时序信息。借助离散情绪模型的 SEED(sjtu emotion eeg dataset)数据集和连续情绪模型的 DEAP(database for emotion analysis using physiological signals)数据集来进行情绪分类实验。实验结果表明,在 SEED 和 DEAP 两个数据集上,CNN-BiLSTM 模型均取得了目前最好的情绪分类性能。此外,该模型的时序信息挖掘模块性能优于单层长短时神经网络,能够学习更多的时序信息。

关键词:情绪分类;脑电;神经网络;特征融合

中图分类号:TP183 **文献标志码:**A

文章编号:1673-0062(2023)06-0041-09

EEG Emotion Classification Model and Validation Based on CNN-BiLSTM

YANG Xiaotong¹, YI Cannan², JIANG Fuliang^{1,2*}, HU Hong²,
ZHANG Mian¹, LI Haoyu¹, WU Wen²

(1. School of Resources Environment and Safety Engineering, University of South China, Hengyang, Hunan 421001, China; 2. School of Safety and Management Engineering, Hunan Institute of Technology, Hengyang, Hunan 421002, China)

Abstract: Emotions play an important role in people's thinking, behavior and communication styles. In order to improve the accuracy of emotion recognition from EEG (electroencephalogram) signals and make full use of the information in the frequency, spatial and

收稿日期:2023-09-01

基金项目:湖南工学院省级应用特色学科开放基金(KFB22028;KFB22029);湖南省教育厅项目(21C0811);国家级大学生创新创业训练计划项目(2022211528009)

作者简介:杨晓同(1999—),男,硕士研究生,主要从事脑电信号识别与分类方面的研究。E-mail:2312383000@qq.com。*通信作者:蒋复量(1978—),男,教授,主要从事安全工程方面的研究。E-mail:jfljfd@163.com

temporal dimensions of EEG signals, a CNN-BiLSTM (convolutional neural networks-bidirectional long short term memory)-based neural network model for EEG emotion classification is proposed. The model consists of a convolutional neural network and a bi-directional long-short-time neural network with multi-layer feature fusion, in which the convolutional neural network is used to learn the frequency and spatial features of the EEG signals, and the bi-directional long-short-time neural network mines the temporal information between EEG slices from the output of the convolutional neural network. Emotion classification experiments are conducted with the help of the SEED dataset for the discrete emotion model and the DEAP dataset for the continuous emotion model. The experimental results show that the CNN-BiLSTM model achieves the best emotion classification performance so far on both SEED and DEAP datasets. In addition, the temporal information mining module of the model outperforms single-layer LSTM (long short term memory, LSTM) and is able to learn more temporal information.

key words: emotion classification; electroencephalogram (EEG); neural network; feature fusion

0 引言

情绪是人对客观事物的态度体验及相应的行为反应,在人们的日常生活中起着重要的作用,影响着人们的认知和决策^[1]。近年来,由于情绪识别在自然人机交互、人工智能、医疗和军事等领域的潜在应用而受到越来越多的关注,情绪识别已成为当下炙手可热的研究方向^[2]。

目前普遍认为情绪量化模型分为两类,离散模型和连续维度的情绪量化模型^[3]。面部表情^[4]、语音语调^[5]、姿势动作^[6]和生理信号^[7]都可以作为情绪识别的数据来源,前三种包含的信息较少而且很容易被伪装,但是以眼电、脑电(electroencephalogram, EEG)和功能性近红外光谱为代表的生理信号是人体自发生成的,客观不易被伪装,包含的信息也更丰富^[8]。考虑到中枢神经系统(大脑)通过调节和控制自律神经系统参与情绪变化过程^[2],因此脑电信息是研究情绪认知机制和识别情绪状态的有效手段。

以随机森林^[9]、k 近邻算法^[10]和支持向量机^[11]为代表的传统机器学习模型在脑电情绪分类上取得了非常好的效果。但是机器学习的局限性在于其复杂的手动特征提取以及浅层模型,前者提高了处理脑电信号的知识储备门槛而后者限制了机器学习在情绪分类模型上的精确度。得益于深度学习在计算机视觉领域出色的表现以及脑机接口的快速发展,一些研究者开始将深度学习应用于脑电的情绪分类。N. Thammasan 等^[12]利

用深度信念网络成功对脑电情绪分类,其在效价和唤醒度的准确率均优于支持向量机和决策树。Y. X. Yang 等^[13]提出了一种新的通道频率卷积神经网络,该网络对情绪的识别准确率达到 92.24%。Y. Wang 等^[14]提出并实现了一个名为 EmotioNet 的三维卷积神经网络,在没有任何手动提取特征的情况下,在 DEAP 数据集的唤醒和价态分类准确率分别为 73.3% 和 72.1%。S. Alhagry 等^[15]提出了一种基于长短时记忆神经网络(long short term memory, LSTM)的两层循环神经网络,在 DEAP 公开数据集上的唤醒、价态和喜欢三个维度的准确率分别为 85.65%, 85.45% 和 87.99%。F. Y. Shen 等^[16]借鉴计算机视觉的思维,将不同通道、不同频率的微分熵(differential entropy, DE)特征转化为 4 维图片,利用卷积神经网络(convolutional neural networks, CNN)和 LSTM 识别不同情绪,该模型在 DEAP 数据集的价态和唤醒度两个维度的准确率分别为 $(94.22 \pm 2.61)\%$ 和 $(94.58 \pm 3.69)\%$, 在 SEED 数据集上四分类准确率为 $(94.74 \pm 2.32)\%$ 。这些研究的着重点大多在于脑电信号的空间和频率,而对时序的研究较少,大多神经网络忽略了时序信息或者只使用了一层 LSTM 层,这对于极高时间分辨率的 EEG 信号来说,并不能充分提取情绪特征。

基于此,笔者拟提出一种综合考虑 EEG 信号的自身特点,充分利用脑电信号的频率和空间信息,重点挖掘时序信息的脑电情绪分类模型,以期提高情绪分类的准确度。

1 模型构建

本文提出了一种基于卷积神经网络和双向长短时神经网络混合的用于脑电情绪识别的模型(4 dimensional-inception convolutional neural networks, 4D-INN)。提出此模型的出发点在于充分挖掘脑电信号所携带的时序信息并且将脑电信号的频率、电极空间所携带的有效信息融合起来。为了提取脑电信号在时域上的更深层次的特征,使用双向长短时神经网络(bidirectional long short term memory, BiLSTM)代替 LSTM,在此基础上借鉴 Inception^[17] 结构将 BiLSTM 提取到的特征融合,充分挖掘脑电信号在时域上携带的信息,并将其命名为 Inception-BiLSTM。

图 1 说明了 4D-INN 的总体架构,它由创建四维数据、卷积神经网络层、双向特征融合的 Inception-BiLSTM 模块和最后的分类器组成。

1.1 创建四维数据

创建四维数据的过程如图 2 所示,首先将每个被试的 EEG 数据等长的划分为 T_s 长的 N 个片

段,片段不重叠,片段的标签与原数据相同。然后利用带通滤波器将原始数据分解为 θ (4 ~ 8 Hz)、 α (8 ~ 12 Hz)、 β (14 ~ 31 Hz) 和 γ (31 ~ 45 Hz) 四个频段,随后以 0.5 s 的不重叠滑动窗口提取每个通道的 DE 特征,DE 已被证明是情绪识别中最显著最稳定的特征^[2]。

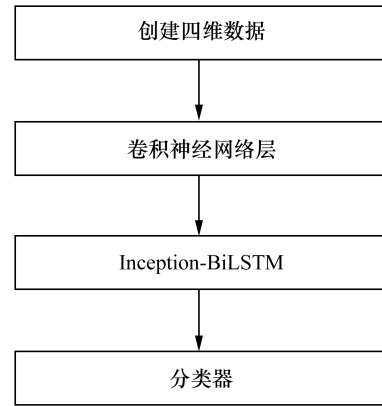


图 1 4D-INN 总体架构

Fig. 1 4D-INN overall architecture

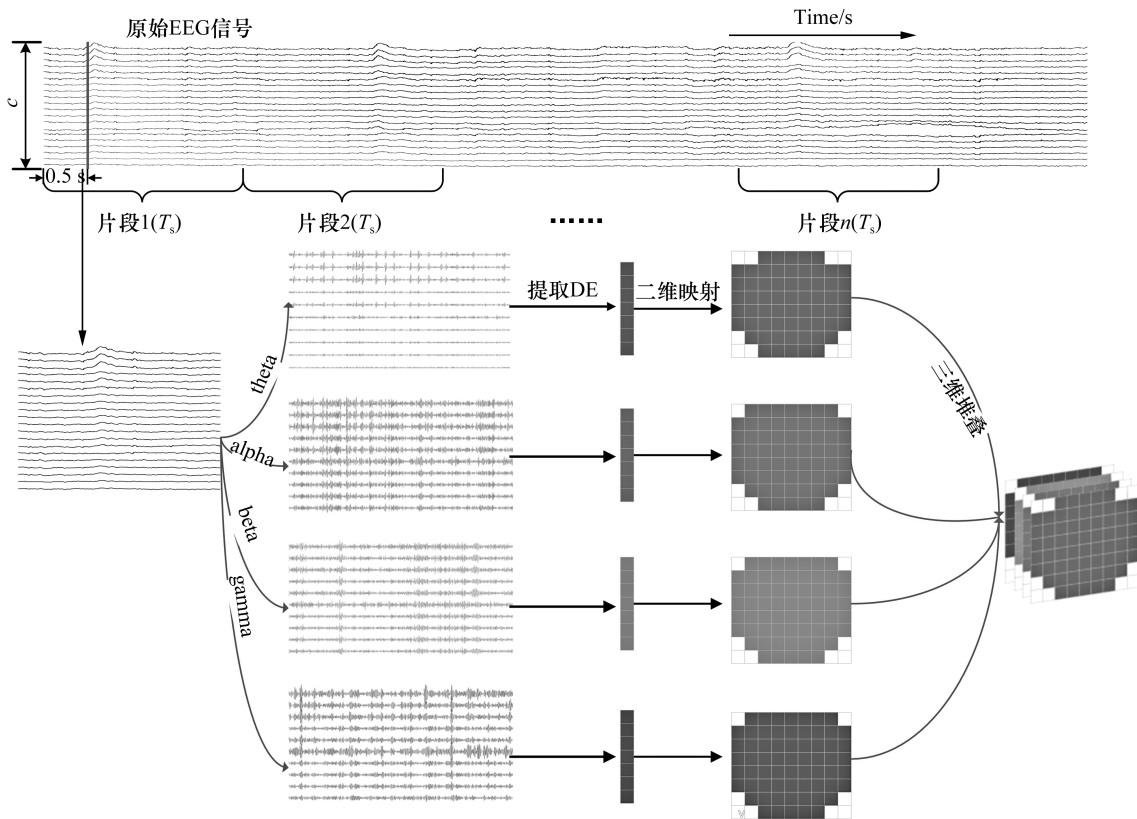


图 2 创建四维数据

Fig. 2 Create four-dimensional data

D_e 被定义为

$$D_e = - \int_a^b p(x) \log(p(x)) dx \quad (1)$$

其中 $p(x)$ 表示连续信息的概率密度函数, $[a, b]$ 表示信息取值的区间。对于一段特定近似服从高斯分布 $N(\mu, \sigma_i^2)$ 的 EEG 信号, 其 DE 可以表示为

$$D_e = - \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x-\mu)^2}{2\sigma_i^2}} \times \log\left(\frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x-\mu)^2}{2\sigma_i^2}}\right) dx = \frac{1}{2} \log(2\pi e \sigma_i^2) \quad (2)$$

考虑到脑电信号是一段连续信号, 因此标签是一段连续时间的标签, 那么一段时间内的某一个时刻的脑电信号也可以带有待识别标签的特征^[18]。有研究指出^[19], EEG 电极的位置信息可

以为情感识别提供增益信息。基于此, 为保留电极空间位置信息, 借鉴计算机视觉中图片的红绿蓝三通思想, 每个通道的图片都会带有图片中待识别物体的特征, 可以将预处理好的每个标签对应的 T_s 的脑电数据以通道数为单位进行堆叠, 其中电极空间位置作为像素点位置, DE 值作为像素值, 频率作为颜色通道数构建三维数据结构。

利用电极的空间位置, 以二维映射的方式把每个频段的 DE 数据结构从一维转为二维, 具体映射方式见图 3, 最后将每个频段的二维 DE 特征图堆叠成一个三维数据。至此, 每一个片段的脑电数据可以表示成一个四维数据 $X_n \in R^{h \times w \times f \times 2T}$, $n=1, 2, 3, \dots, N$ 。其中 N 表示片段的个数, h 和 w 为二维映射时的长和宽, f 表示频带数, $2T$ 表示片段长的两倍。

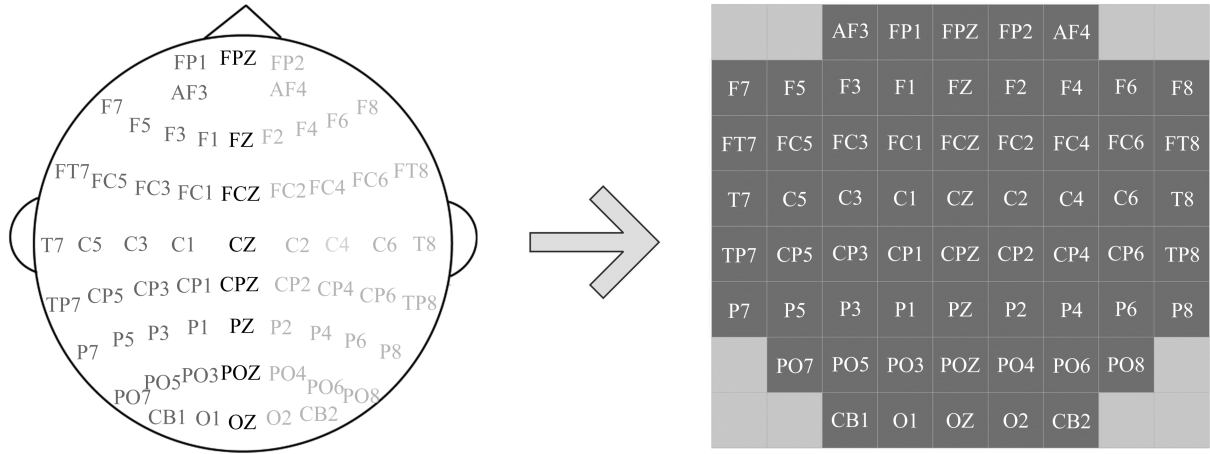


图3 2D映射关系图

Fig.3 2D mapping relationship diagram

1.2 卷积神经网络

对于每一个样本 X_n , $n=1, 2, 3, \dots, N$, 通过 CNN 提取每一个 0.5 s 时间切片中的频率和空间信息。4D-INN 摒弃了传统卷积神经网络中的池化层, 池化层存在的意义是仿照人的视觉系统进行降采样, 以牺牲信息量为代价获取更高层的抽象特征, 但是对于小尺寸脑电切片来说, 最好的方式是保存更多的信息而不是降维, 因此在提取频率和空间信息时并没有加入池化层。卷积神经网络模块如图 4 所示, 该模块包括四个卷积层、一个 Flatten 层和一个全连接层。第一层是用 64 个 2×2 的卷积核将脑电切片升维到 $8 \times 9 \times 64$, 目的是为了获取更多的脑电特征。第二和第三卷积层均使用

两个 3×3 的卷积核, 卷积核个数分别为 128 和 256, 这里借鉴了 VGG (vision geometrical group) 模型的思想, 利用两个 3×3 的卷积核代替 5×5 的卷积核, 在感受野相同的情况下获得更多的非线性特征。最后一层为了融合通道特征信息, 使用了 64 个 1×1 的卷积核。将输出平坦化处理后送入具有 512 神经元的全连接层, 最后每个片段的输出可以表示为 $C_n = R^{512 \times 2T}$, $n=1, 2, 3, \dots, N$ 。

该模块中每一层的卷积操做之后, 都会有一个批标准化层用来加快模型训练速度, 为避免卷积操作中边缘信息丢失, 选择零填充。每一次卷积操做的激活函数为线性整流单元激活函数 (rectified linear unit, ReLU), 步长为 1。

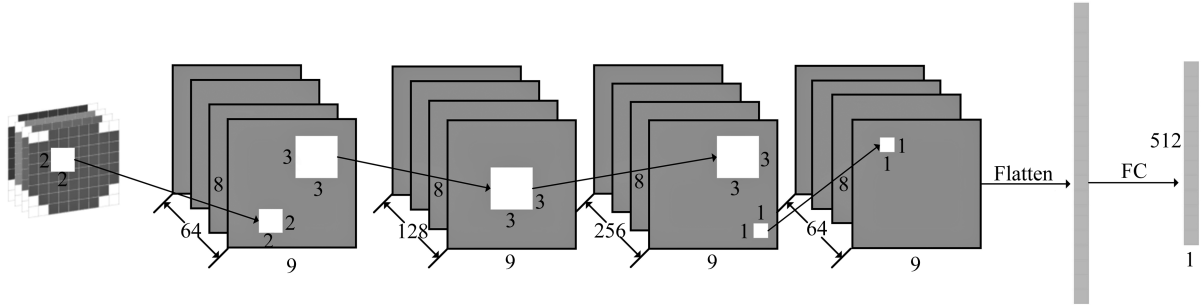


图 4 卷积神经网络

Fig. 4 Convolutional neural network

1.3 Inception-BiLSTM

脑电信号是一种高分辨率的时序信号,时切片之间隐藏着额外的时间特征。可以利用 Inception-BiLSTM 模块从 CNN 的输出中挖掘时间信息。Inception-BiLSTM 模块结构如图 5 所示,它包含四个 BiLSTM 层和两次特征融合,BiLSTM 的输入为 T_s 内 CNN 输出的序列集合 $C_n = (C_1, C_2 \dots C_{2T})$ 。

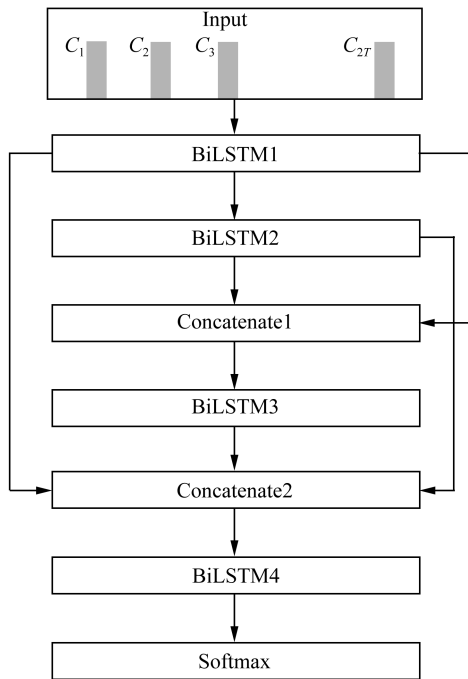


图 5 Inception-BiLSTM 模块

Fig. 5 Inception-BiLSTM module

LSTM 本质上是一种特殊的循环神经网络,它能够学习长序列信息前后的依赖和关联关系,记录间隔或延迟较长的历史信息 and 重要事件,目前已成功应用于语音识别和情感分析等领域。BiLSTM 即为双向 LSTM,将两个方向相反的单向

LSTM 连接到同一输出端。与单向 LSTM 相比,双向 LSTM 能保存过去和未来的信息,这使其能够更加准确的挖掘情绪发生时脑电信号前后之间的时序特征。输入时间序列首先会被依次送入 BiLSTM1 和 BiLSTM2,然后将 BiLSTM1 和 BiLSTM2 输出特征融合,以此来弥补两次长短时神经网络遗漏的时序信息。将融合特征送入 BiLSTM3,随后将三次 BiLSTM 的输出进行第二次特征融合,原因同上。第二次特征融合的信息将会被送入 BiLSTM4 中,此模块的 BiLSTM 层节点数均为 128,因此最终输出为 $y_n = R^{128}$,其包含了 T_s 时间段内频率、空间和时间信息,将会作为输入送入分类模块。

1.4 分类器

基于 BiLSTM 的最终输出结果 $y_n = R^{128}$,选择使用一层全连接层和 Softmax 激活函数实现最后的预测分类,计算公式如下所示:

$$F_c = \text{Softmax}(W_f y_n + b_f) \quad (3)$$

其中 W_f 为全连接层的权重, b_f 为偏置量。

2 实验数据

在本节中将会简单介绍 SEED 和 DEAP 两个公开数据集,4D-INN 模型实现的具体参数和实验结果的计算方法。

2.1 SEED 数据集

SEED 数据集是由上海交通大学的 BCMI (center for brain-like computing and machine intelligence) 实验室提供的公开脑电情绪数据集。其精心选择了持续时间大约为 4 min 的 15 个电影片段作为情绪诱发刺激源,情绪可以分为三类,分别是正面的、中性的和负面的情绪,每一种情绪对应 5 个电影片段。15 名健康受试者(男性 7 人,女性 8 人,年龄 23.27 ± 2.37) 参加脑电信号实验,每个

被试进行 3 组实验,每个片段的观看过程可以分为四个阶段,5 s 开始提示,4 分钟电影片段,45 s 自我评估以及 15 s 休息时间。实验中产生的脑电信号由 ESI NeuroScan 系统记录。

为了与现有研究方法进行比较,本文采用的实验数据为官方预处理后的数据。SEED 数据集中的脑电数据采样率降至 200 Hz,并去除了噪声和伪影。

2.2 DEAP 数据集

DEAP 数据集是由英国伦敦玛丽皇后大学 Korlstra 等人采集并用于研究人类情绪状态的公开数据集。它使用持续时间为 1 min 的 40 个音乐视频片段作为情绪诱发刺激源,每个片段的情绪类别由效价-唤醒度值(1~9)标记,将 5 作为阈值,将情绪分类转化为两个维度上的二分类问题。32 名健康被试参加脑电实验,每次观看音乐视频片段后被试都会被要求对他们的效价和唤醒度水平自我评估。实验中产生的脑电信号由 32 通道的 Biosemi ActiveTwo 系统记录。

数据集发布前脑电信号已经被去除了肌电和眼电伪影,采样率降至 128 Hz 且应用 4~45 Hz 的带通滤波器消除噪声,采用官方记录四个标签值中的唤醒度(Arousal)、效价(Valence)作为本文的评价标准。

2.3 实验设置

本文中所提出的 4D-INN 深度学习模型是以 4D 片段 $X_n \in R^{h \times w \times f \times 2T}$ 作为输入,其中 h 和 w 为二维映射时的长和宽,由于 F. Y. Shen 等^[16]研究指出稀疏矩阵对于情绪分类的提升并不明显,但是会增加近三倍的运行时间,因此本文中的 h 和 w 分别设置为 8 和 9。W. L. Zheng^[20]和 Y. L. Yang 等^[21]的研究指出,所有频带组合可以相互补充,相较于单个频带在情绪分类方面有更优秀的表现,因此设置 $f=4$ 。

本文对每个被试的输入数据采用十折交叉验证的方式,因此每个被试的平均分类准确率(accuracy, ACC)和标准差(standard deviation, STD)代表了模型在每个被试上的表现。所有被试的平均分类准确率和标准差代表模型在整个数据集上的表现。

4D-INN 深度学习模型由 python3.8, tensorflow-2.9.0 和 keras-2.9.0 实现,使用 Adam 优化器,学习率设置为 0.001,每次训练的样本数量为 64,最大训练批次为 100,使用 NVIDIA GTX 3060 GPU

对模型进行训练。

3 实验结果分析

本节中首先探究了 EEG 片段长度 T 对于识别精确度的影响。然后评估了 4D-INN 的整体性能。随后比较了 4D-INN 与其他传统结构在情绪分类上的表现。最后探究了 Inception-BiLSTM 在 EEG 情绪分类上的效果。

3.1 时长对分类结果的影响

由于 EEG 是时序信号,为探究 EEG 不同片段长度所包含的信息对情绪分类准确率的影响,将 T 分别设置为 1, 2, 3 和 4,表 1 为 4D-INN 在 SEED 和 DEAP 上不同 T 的表现。

从表 1 中数据可以看出 4D-INN 在不同的 T 取值下都获得了较高的精确率。两个数据集最高的分类准确率均出现在 $T=2$ 时,对于 SEED 数据集最高准确率为 $(95.18 \pm 0.28)\%$,对于 DEAP 数据集的 Arousal 和 Valence 两个维度上分类准确率分别为 $(95.45 \pm 1.76)\%$ 和 $(94.86 \pm 1.85)\%$ 。因此 $T=2$ 是 4D-INN 最适合的参数,在本文后续部分,默认 $T=2$ 。

表 1 时长 T 对分类准确率的影响(ACC±STD(%))

Table 1 Effect of duration T on classification accuracy(ACC±STD(%)) 单位: %			
T/s	SEED	DEAP-Arousal	DEAP-Valence
1	95.11±0.21	95.05±1.23	94.64±1.35
2	95.18±0.28	95.45±1.76	94.86±1.85
3	93.72±2.58	95.23±2.23	94.82±2.31
4	94.82±0.43	95.04±2.54	94.71±2.59

值得注意的是,不同长度之间的准确率差距非常小,SEED 数据集上的极差为 1.46%,在 DEAP 数据集上分别为 0.41% 和 0.22%,因此 4D-INN 在 EEG 片段时长方面具有一定的鲁棒性。

3.2 4D-INN 的整体性能

为探究 4D-INN 在 SEED 和 DEAP 数据集上的整体性能,分析了两个数据集中每个被试的分类情况。

4D-INN 在 SEED 数据集上的表现如图 6 所示,15 名被试的分类准确率均在 90% 以上,其中最大值和最小值分别为 99.47% 和 90.05%。整体来看,平均分类准确率为 95.18,标准差为

0.28%, 极差为 9.42%, 有 8 名被试的准确率超过平均值。

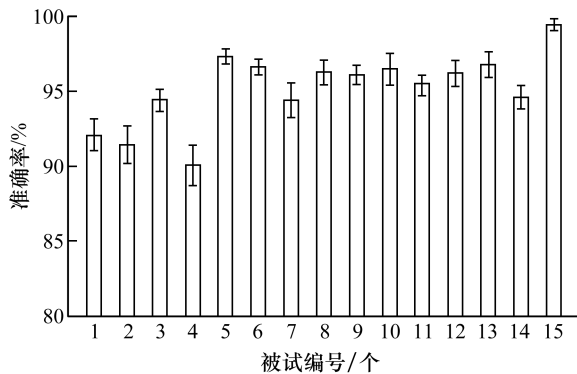


图 6 4D-INN 在 SEED 数据集上的表现

Fig. 6 Performance of 4D-INN on SEED dataset

4D-INN 在 DEAP 数据集上的表现如图 7 所示。在 Arousal 分类方面,有 31 名被试的准确率超过 90%, 其中最大值和最小值分别为 99.5% 和 78.41%。从整体来看, 平均分类准确率为 95.45%, 标准差为 1.76%, 极差为 21.09%, 有 18 名被试的准确率超过平均值。在 Valence 分类方面, 有 30 名被试的准确率在 90% 以上, 最大值和最小值分别为 99.33% 和 86.75%。整体来看, 平均分类准确率为 94.86%, 标准差为 1.85%, 极差为 12.58%, 有 18 名被试的准确率超过平均值。特别注意的是, 在 Arousal 和 Valence 两个方面, 第 22 名被试都大幅低于平均值, 分别为 78.41% 和 86.75%, 这是在两个数据集上都未出现过的情况。推测这是由该被试在观看视频刺激时不认真或者实验结束后主观自我评估出现了错误导致。

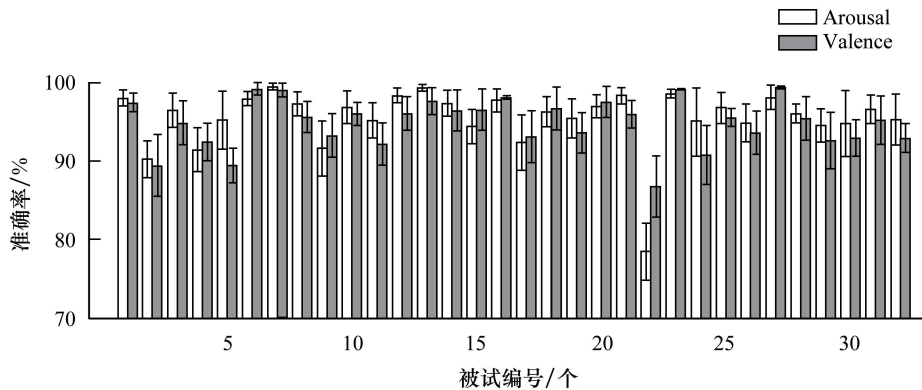


图 7 4D-INN 在 DEAP 数据集上的表现

Fig. 7 Performance of 4D-INN on DEAP dataset

总体来看, 4D-INN 在 SEED 和 DEAP 两个数据集上均取得了较为稳定的结果, 标准差和极差都在可控范围以内, 这表明 4D-INN 不会因为被试和情绪量化模型的不同而产生较大的误差。

3.3 不同模型对比

为了探究 4D-INN 的优越性, 选择了以下几种常用的分类模型进行对比:

CCNN (continuous convolutional neural network)^[13]: 该方法利用不同频段中 DE 特征所携带的信息, 利用 CNN 挖掘脑电信号中频率和空间信息并进行分类。

PCRNN (parallel convolutional recurrent neural network)^[22]: 该方法考虑原始脑电信号中蕴含的信息, 利用 LSTM 和 CNN 模块挖掘时空信息, 之后将两个模块的输出特征融合送入全连接层

分类。

HCNN (hierarchical convolutional neural networks)^[23]: 该方法使用 β 频段的 DE 特征作为输入, 只考虑脑电信号的空间信息, 使用分层的 CNN 架构进行脑电情绪分类。

4D-CRNN (4 dimensional convolutional recurrent neural network)^[16]: 该模型使用了与 4D-BIN 相似的方法, 构建四维数据后利用卷积层提取空间信息, 利用循环神经网络提取时间信息。

EmotionNet (eeg-specific three dimensional Convolutional Neural Network)^[14]: 该方法对二维和三维卷积核分别提取和同时提取空间和时间特征的有效性进行了详细的比较, 该模型挖掘了脑电信号的时空信息。

MCNN (multi-dimensional convolutional neural

networks)^[24];该方法利用脑电信号每个时间点的空间和频带特征构建 3D 矩阵,利用多维卷积神经网络进行分类。

对比结果如表 2 所示,4D-INN 在 SEED 和 DEAP 两个数据集上均取得了最优秀的表现。相比较于 4D-INN,CCNN 和 RGNN 缺少了时间维度的信息,HCNN、PCRNN 和 Emotion-Net 缺少了频段特征,而 MCNN 虽然融合了时间、频段和空间信息,但是其神经网络模型并没有循环神经网络层,这就使得时间特征并没有得到充分的利用。而对于 4D-CRNN 准确率低于 4D-INN 的原因会在后面解释。

表 2 其他模型与 4D-INN 准确率对比 (ACC±STD(%))

Method	SEED	DEAP-Arousal	DEAP-Valence
CCNN	—	90.24	89.45
PCRNN	—	90.80	91.03
HCNN	88.20	—	—
4D-CRNN	94.74±2.32	94.58±3.69	94.22±2.61
Emotion Net	—	73.3	72.1
MCNN	—	85.88	87.32
4D-INN	95.68±0.28	95.65±1.76	95.46±1.85

为了验证 Inception-BiLSTM 模块在时间信息学习方面的有效性,对 CCNN 和 HCNN 的原始模型不变,在 CNN 模块后添加 Inception-BiLSTM 模块,对于 PCRNN 和 4D-CRNN,使用 Inception-BiLSTM 模块代替 LSTM。

改进结果如表 3 所示,相比较于与原模型,CCNN 在 DEAP-Arousal 和 DEAP-Valence 两个维度分别提高了 1.64% 和 1.08%,而 PCRNN 在 DEAP-Arousal 和 DEAP-Valence 两个维度分别提高了 3.24% 和 2.23%,HCNN 在 SEED 上提高了 2%,4D-CRNN 在 SEED、DEAP-Arousal 和 DEAP-Valence 三个维度分别提高了 0.25%、0.39% 和 0.48%。因此,Inception-BiLSTM 能够有效的从脑电信号中提取时序信息,而且相比单层 LSTM,Inception-BiLSTM 能够提取挖掘更多的时序信息,这也是为什么 4D-CRNN 的准确率低于 4D-INN 的原因。

表 3 Inception-BiLSTM 改进模型准确率 (ACC±STD(%))

Method	SEED	DEAP-Arousal	DEAP-Valence
CCNN	—	91.88	90.53
PCRNN	—	94.04	93.26
HCNN	90.20	—	—
4D-CRNN	94.99±1.82	94.97±1.82	94.70±1.97

4 结 论

1) 为提高脑电信号情绪识别的准确率,应充分考量脑电信号自身特点,保证神经网络模型的输入尽可能多的包含频域、空间和时域信息,综合了频率、空间和时间信息的 4D 特征结构要优于不包含这些信息的 2D 和 3D 特征结构。同时包含 CNN 和 LSTM 的混合神经网络分类性能要优于单一的神经网络,CNN 和 LSTM 的纵向融合比 CNN 和 LSTM 的并行连接更适合提取空间和时间信息。

2) 4D-INN 模型在 SEED 数据集和 DEAP 数据集的效价和唤醒维度上分别取得了 (95.68±0.28)%、(95.65±1.76)% 和 (95.46±1.85)% 的准确率,均高于现有的分类模型,结果证明了所提模型的可行性与有效性。Inception-BiLSTM 在脑电信号时序信息提取方面具有一定的优势,相比较与 CCNN、PCRNN、HCNN 和 4D-CRNN 的分类准确率,使用 Inception-BiLSTM 优化后分别提升了 1.64%、3.24%、2% 和 0.25%。

3) 4D-INN 算法深度挖掘脑电信号信息的时增加了运行时间,如何降低运行时间将作进一步研究。

参考文献:

- [1] 李锦瑶,杜肖兵,朱志亮,等. 脑电情绪识别的深度学习研究综述[J]. 软件学报,2023,34(1):255-276.
- [2] 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.
- [3] 秦天鹏,生慧,岳路,等. 脑电信号情绪识别研究综述[J]. 计算机工程与应用,2023,59(15):38-54.
- [4] 潘家辉,何志鹏,李自娜,等. 多模态情绪识别研究综述[J]. 智能系统学报,2020,15(4):633-645.
- [5] 宋静,张雪英,孙颖,等. 基于 PAD 情绪模型的情感语音识别[J]. 微电子学与计算机,2016,33(9):128-131.

- [6] 许芬, 闰文彬, 张晓平. 姿态情绪识别研究综述[J]. 计算机应用研究, 2021, 38(12): 3521-3526.
- [7] 王跃飞, 马伟丽, 王文康, 等. 基于生理特征映射的驾驶员情绪在线识别模型构建方法[J]. 机械工程学报, 2022, 58(20): 379-390.
- [8] AVITAN L, TEICHER M, ABELES M. EEG generator-a model of potentials in a volume conductor[J]. Journal of neurophysiology, 2009, 102(5): 3046-3059.
- [9] 王发旺, 陈睿, 伏云发. 基于 DBN 和 RF 的跨被试情绪识别研究[J]. 南京大学学报(自然科学), 2021, 57(4): 617-626.
- [10] 李贤哲, 暴伟, 谢能刚. 基于脑电信号的情绪识别[J]. 北京生物医学工程, 2022, 41(1): 8-16.
- [11] 郝笑弘, 薛保菊. 基于目标域自适应 SVM 分类器的微博情绪分类[J]. 工程数学学报, 2021, 38(5): 627-636.
- [12] THAMMASAN N, FUKUI K I, NUMAO M. et al. Application of deep belief networks in eeg-based dynamic music-emotion recognition [C]//2016 International Joint Conference on Neural Networks. Vancouver, Canada: IJCNN, 2016: 881-888.
- [13] YANG Y X, GAO Z K, WANG X M, et al. A recurrence quantification analysis-based channel-frequency convolutional neural network for emotion recognition from EEG[J]. Chaos, 2018, 28(8): 085724.
- [14] WANG Y, HUANG Z Y, MCCANE B, et al. EmotioNet: A 3-D Convolutional Neural Network for EEG-based Emotion Recognition[C]//2018 International Joint Conference on Neural Networks. Rio de Janeiro, Brazil: IJCNN, 2018: 1151-1157.
- [15] ALHAGRY S, FAHMY A A, ELKHORIBI R A, et al. Emotion Recognition based on EEG using LSTM recurrent neural network[J]. International journal of advanced computer science and applications, 2017, 8(10): 355-358.
- [16] SHEN F Y, DAI G J, LIN G, et al. EEG-based emotion recognition using 4D convolutional recurrent neural network [J]. Cognitive neurodynamics, 2020, 14(6): 815-828.
- [17] 林景栋, 吴欣怡, 柴毅, 等. 卷积神经网络结构优化综述[J]. 自动化学报, 2020, 46(1): 24-37.
- [18] 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.
- [19] CHAO H, DONG L, LIU Y L, et al. Emotion recognition from multiband EEG signals using CapsNet[J]. Sensors, 2019, 19(9): 2212.
- [20] ZHENG W L, ZHU J Y, LU B L. Identifying stable patterns over time for emotion recognition from EEG[J]. IEEE transactions on affective computing, 2019, 10(3): 417-429.
- [21] YANG Y L, WU Q F, FU Y Z, et al. Continuous convolutional neural network with 3D input for EEG-Based emotion recognition[C]//Neural Information Processing: 25th International Conference. Siem Reap, Cambodia: ICONIP, 2018: 433-443.
- [22] 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. Rio de Janeiro, Brazil: IJCNN, 2018: 1-7.
- [23] 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.
- [24] 陈景霞, 闵重丹, 林文涛. 基于 3D 矩阵特征与多维卷积网络的脑电信号情感识别[J]. 陕西科技大学学报, 2022, 40(2): 178-186.