



基于并行可分离卷积和标签平滑正则化的脑电情感识别

张永^{1,2}, 刘纪奎¹, 柯文龙¹

(1. 湖州师范学院信息工程学院, 浙江 湖州 313000;

2. 辽宁师范大学计算机与信息技术学院, 辽宁 大连 116081)

摘要: 近年来, 基于深度学习和脑电图 (EEG) 的情感识别方法取得了较好的效果。然而, 现有方法依然存在脑电情感特征提取不够全面、受人工错误标注的情感标签影响较大等问题。对此, 提出了并行可分离卷积和标签平滑正则化 (PSC-LSR) 网络模型。首先, 通过注意力机制, 赋予 EEG 重要时间点和重要通道更大的权重, 得到 EEG 的浅层情感特征; 其次, 采用并行可分离卷积模块全面提取 EEG 情感信息, 得到深层情感特征; 最后, 在优化模型参数时采用了情感标签平滑正则化方法, 使模型对错误标签有更大的容错概率, 增强了网络模型的泛化性和鲁棒性, 提高了脑电情感识别的准确率。提出的方法在两个数据集进行了验证, 其中, 在 DEAP 数据集中, 唤醒和效价两个维度的平均准确率分别达到了 99.23% 和 99.13%; 在 Dreamer 数据集中, 唤醒和效价两个维度的平均准确率分别达到了 97.33% 和 97.25%。

关键词: 脑电信号; 情感识别; 标签平滑正则化; 注意力机制

中图分类号: TP391

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2023112

EEG emotion recognition based on parallel separable convolution and label smoothing regularization

ZHANG Yong^{1,2}, LIU Jikui¹, KE Wenlong¹

1. School of Information Engineering, Huzhou University, Huzhou 313000, China

2. School of Computer Science and Information Technology, Liaoning Normal University, Dalian 116081, China

Abstract: In recent years, emotion recognition methods based on deep learning and electroencephalogram (EEG) have achieved good results. However, existing methods still have issues such as incomplete extraction of emotional features from EEG and significant impact from artificially mislabeled emotional labels. A parallel separable convolution and label smoothing regularization (PSC-LSR) network model was proposed. Firstly, through the attention me-

收稿日期: 2023-04-06; 修回日期: 2023-05-12

通信作者: 柯文龙, wlke@zjhu.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61772252); 辽宁省教育厅科学研究经费项目 (No.LJKZ0965); 湖州科技计划项目 (No.2022GZ08)

Foundation Items: The National Natural Science Foundation of China (No.61772252), The Scientific Research Foundation of the Education Department of Liaoning Province (No.LJKZ0965), The Huzhou Science and Technology Plan Project (No.2022GZ08)

chanism, EEG important time points and important channels were given greater weight to obtain shallow emotional features of EEG. Secondly, a parallel separable convolution module was used to comprehensively extract EEG emotional information and obtain deep emotional features. Finally, the emotion label smoothing regularization method was used to optimize the model parameters, which increased model's fault tolerance probability for incorrect labels, enhanced the generalization and robustness of the network model, and improved accuracy of EEG emotion recognition. The proposed method has been validated in two datasets, in which the average accuracy rates of arousal and valence dimensions in the DEAP dataset reaches 99.23% and 99.13%, respectively. In the Dreamer dataset, the average accuracy rates for both arousal and valence dimensions reaches 97.33% and 97.25%.

Key words: EEG, emotional recognition, label smoothing regularization, attention mechanism

0 引言

情感是人类一种常见的心理现象,与人们日常的行为和认知密切相关。对于人们自身来说,如果能很好地管理自己的情感,那么对自己的心理健康、为人处世甚至个人事业成功都有帮助。人类的情感产生分为生理信号和非生理信号,生理信号包括呼吸、心律、体温、脑电图(electroencephalogram, EEG)、眼电图(electrooculogram, EOG)、心电图(electrocardiogram, ECG)、肌电图(electromyogram, EMG)、皮肤电反应(galvanic skin response, GSR)等^[1-3]。非生理信号包括面部表情、动作行为、言语或者在社交平台发出的评论等^[4]。面部表情识别通常使用图像识别方法实现^[5-6]。语音识别根据人们在不同的情感状态下产生的不同语音语调或表达方式进行识别^[7-8]。非生理信号虽然更容易采集,但是由于它可以被人类伪装,所以得到的数据不一定真实。生理信号是人体系统对外部刺激更真实、更直接的反应,人们很难隐藏真实的情感状态。研究人员通常选用更加真实的生理信号,其中 EEG 在人类情感识别方面具有良好的时间分辨率,相比其他生理信号更加准确,所以基于 EEG 的情感识别方法受到研究人员的关注。

目前,情感识别有两种研究模型:离散模型和维度模型。在离散模型中,Ekman^[9]提出了快乐、悲伤、惊讶、恐惧、愤怒和厌恶 6 种基本情感;Plutchik^[10]提出了车轮模型,指出情感由恐惧、愤

怒、接受、悲伤、好奇、惊讶、喜悦和厌恶 8 种情感构成。研究发现,人们的情感之间存在很多细微的联系,离散模型表现出的情感是有限的。相比之下,维度模型可以在区分情感之间的联系方面表现得更好,维度模型将情感的表达用空间坐标系的方式表现出来,一般分为 2D 或者 3D 模型。Russel 绕线理论^[11]提出了 2D 情感坐标系,以唤醒(arousal)、效价(valence)为坐标轴。其中,唤醒指情感的强度,从平静到兴奋;效价指情感的状态,从消极到积极。Mehrabian^[12]在 2D 模型的基础上通过增加一个主导度(dominance)维度建立了 3D 模型,主导度指某个人对某方面的控制程度。

基于脑电信号的情感识别方法通常分为两个步骤。第一个步骤是对采集到的脑电信号进行特征提取,由于脑电信号比较微弱,采集时很容易受其他噪声干扰^[13],所以在特征提取前一般要对脑电信号进行预处理,去除采集过程中的伪迹。预处理的常用方法有滤波法和独立成分分析(independent component analysis, ICA)法等。滤波法指用滤波器滤除噪声较大的频段,滤波可以在很大程度上减少生理信号中掺杂的干扰与噪声^[14]。对于滤波器无法滤除的频段,可以采用 ICA 法将多通道的观测信号分解成多个独立成分^[15]。对预处理之后的数据进行特征提取,有 3 种常使用的脑电信号特征:时域、频域和时频域特征,常见的特征提取方法有微分熵(differential entropy,



DE)、功率谱密度(power spectral density, PSD)、傅里叶变换和小波变换等。第二个步骤是对提取到的特征进行情感分类,使用支持向量机(support vector machine, SVM)、 K 近邻(K -nearest neighbor, KNN)和人工神经网络(artificial neural network, ANN)等机器学习方法,得到情感识别的分类结果。传统方法基于人工提取 EEG 情感特征,特征提取不够全面,很多潜在的情感特征信息被忽略,导致情感识别结果不佳。

近年来,深度学习在情感识别领域表现出良好的效果。然而,随着网络模型深度的增加,网络模型训练参数量越来越大,网络运行速度减慢。并且,由于没有全面提取 EEG 的情感特征信息,识别精度并不理想。因此,设计一种新的网络架构,减少网络模型参数量并提高情感分类准确率是很有必要的。首先,在传统的 EEG 情感识别任务中,研究者过分相信被试者人工标注的情感标签。但人类产生的情感往往是主观的,对每一种情感的理解不尽相同^[16],对不同的事物也存在不同的偏见,所以不同被试者在评级中就会产生主观错误^[17-18]。其次,经过长时间的实验,被试者的注意力和精力可能会下降,也会导致错误的分类标签^[19]。最后,采集 EEG 的实验环境中存在的噪声和其他因素都会导致产生错误的分类标签。而传统方法不考虑错误标签的情况,直接计算网络模型损失,这会使模型优化的参数不是最佳值,出现过拟合问题。标签平滑在深度学习中对分类任务有良好的表现,Szegedy 等^[20]在深度学习中使用了标签平滑正则化(label smoothing regularization, LSR)方法,LSR 通过设置优化软目标方式,容许标签存在一定的误差,可以提高网络模型的鲁棒性与泛化性。

为了解决 EEG 特征提取不充分以及网络模型参数量大的问题,本文提出了一种基于并行可分离卷积和标签平滑正则化(parallel separable convolution and label smoothing regularization,

PSC-LSR)的网络模型。由于不同时间点和不同通道对情感识别的重要性是不一致的,PSC-LSR 采用空间注意力机制,对重要通道和关键时间点赋予较大权重,充分提取脑电信号的情感特征。模型使用了并行的可分离卷积模块,用空间可分离卷积将卷积核分为两个更小的卷积核,对不同大小的卷积核分别提取较近位置和较远位置的特征,在不降低精度的情况下减少网络模型参数量。考虑在基于 EEG 的情感识别任务中人工标注标签存在误差,PSC-LSR 采用了标签平滑正则化策略,使用软目标优化网络模型参数,提高了情感识别分类准确率。在 DEAP 和 Dreamer 两个公开数据集上进行了大量实验,相比近些年的其他深度学习方法都取得了更好的效果。本文的主要贡献如下。

(1)设计了一种基于卷积神经网络(convolutional neural network, CNN)的网络模型 PSC-LSR,使用并行可分离卷积更全面地提取脑电信号特征,在降低网络模型参数量、提高运行速度的同时,获得了更好的分类准确率。

(2)设计了网络参数的优化方法,通过标签平滑正则化策略增强了网络模型对错误标签的容错能力。

(3)运用空间注意力机制,赋予 EEG 关键时间点和重要通道更大的权重,改善了 EEG 情感特征提取不够全面的问题。

1 相关工作

传统情感识别任务大多采用基于机器学习的方法。首先对 EEG 进行特征提取,Khalili 等^[21]将脑电信号的均值、峰值、标准差、原始脑电信号首次差异的平均绝对值、归一化后脑电信号首次差异的平均绝对值等作为脑电信号的时域特征。由于脑电信号同时具有时域和频域特征,且是非平稳非线性的时变信号,只提取一个方面的特征是不够全面的,更常使用时频域特征作为情

感识别的特征。Lin 等^[22]将采集的脑电信号用短时傅里叶变换映射到 delta 频段 (1~4 Hz)、theta 频段 (4~8 Hz)、alpha 频段 (8~13 Hz)、beta 频段 (13~30 Hz) 和 gamma 频段 (36~44 Hz) 5 种频段上, 得到每个频段的 PSD, 计算电极间的对称关系组合出 4 种特征用于后续的情感识别。其次对提取的特征进行情感分类, 在过去的研究中, 很多研究者提出了多种不同的模型。Murugappan 等^[23]将喜悦、恐惧、厌恶和惊讶 4 种不同的情感样本融合在一起, 采用模糊 C 均值聚类算法, 实现了对 4 种情感的划分。Khosrowabadi 等^[24]采用自组织映射的方法取得了比较好的情感分类效果。Liu 等^[25]将提取的脑电信号与眼电信号融合后, 使用线性 SVM 进行情感分类。Hassan 等^[26]将提取后的特征使用高斯核进行情感分类。Vijayakumar 等^[27]将提取的 DEAP 数据库中 8 个外周生理信号的统计特征输入 8 个分类器中, 对情感进行二分类。Mert 等^[28]使用 KNN 和 ANN 对情感进行二分类, 平均识别率分别为 59% 和 73.93%。Wu 等^[29]通过梯度提升决策树 (gradient boosting decision tree, GBDT) 对情感进行分类, 情感平均识别率达 75.18%。然而, 传统方法仅靠人工选择性地提取脑电信号特征, 对脑电情感特征提取不够充分, 导致分类器精度和泛化性较差。

近几年, 深度学习被广泛应用于情感识别中, 采用端到端的方法, 直接将数据处理后的信号作为网络模型的输入, 最后直接输出分类结果。常用的用于 EEG 的深度学习方法有自动编码器 (auto encoder, AE)、CNN、深度信念网络 (deep belief network, DBN) 和递归神经网络 (recurrent neural network, RNN) 等。基于 CNN 的脑电情感识别是目前的热点之一。Cheng 等^[30]使用简单的三层网络算法——卷积层、池化层、全连接层, 在 DEAP 数据集上的总体准确率为 65.10%。Kwon 等^[31]创建了包括 4 个卷积层和 7 个全连接层的二维 CNN, 在 DEAP 数据集上效价和唤醒维度的准确率达到

80.46% 和 76.56%。Cui 等^[32]将区域非对称 CNN 方法用于基于 EEG 的情感识别任务, 在 DEAP 和 Dreamer 数据集上获得了 95% 的识别准确率。Li 等^[33]采用扩张全卷积网络, 在 DEAP 数据集上的唤醒和效价维度实现了 94.59% 和 95.32% 的准确率。Liu 等^[34]采用多模态情感识别方法, 在 DEAP 数据集上的唤醒和效价维度实现了 90.60% 和 89.00% 的准确率。Li 等^[35]使用监督对比学习方法, 更大尺度地分离情感标签, 在 DEAP 数据集上的唤醒、效价和主导度维度分别取得了 98.35%、98.51% 和 98.55% 的准确率。

上述情感识别方法虽然取得了很好的效果, 但是没有考虑如何更全面地提取脑电信号特征, 也忽视了人工标注的情感标签可能存在误差, 并导致模型分类准确率不高以及泛化性较低。本文提出的 PSC-LSR 网络模型有效地解决了上述问题。

2 方法

2.1 基础模型框架

本文提出的 PSC-LSR 网络模型框架如图 1 所示。首先, 将原始 EEG 信号转换为二维特征图 X , 将其输入空间注意力模块 (spatial attention module, SAM) 进行重要时间点和通道权重重分配, 得到浅层特征 X' ; 然后, 使用并行可分离卷积模块作为主要特征提取器, 输入浅层特征 X' , 提取到 EEG 的深层特征 D ; 最后, 将 EEG 的深层情感特征降维输入全连接网络, 并使用本文设计的混合函数优化模型参数, 得到每个维度的二分类 (消极/积极) 结果。其中, 主干网络并行可分离卷积模块有 3 个不同于传统卷积网络的地方。首先, 使用可分离卷积代替传统卷积, 把 3×3 的卷积核拆分为 3×1 和 1×3 的卷积核, 把 5×5 的卷积核拆分为 5×1 和 1×5 的卷积核, 可以大大减少网络参数量。其次, 使用并行网络, 在上层使用 3×3 的卷积核拆分, 提取 EEG 特征图较近位置的信息, 在下层

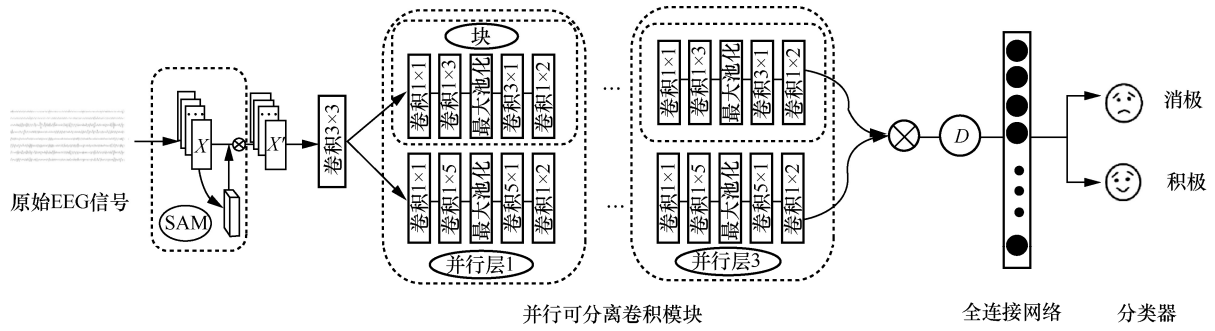


图1 PSC-LSR 网络模型框架

使用了 5×5 的卷积核拆分, 提取较远的位置信息。最后, 实验证明了并行可分离卷积最合适的组合: 对于上层网络, 较小的卷积核可以使用更深的网络; 对于下层较大的卷积核, 对具体层数分别做了具体实验以确定大小。

接下来对模型的每个部分进行详细介绍。

2.2 空间注意力机制

注意力机制在一些基于深度学习的任务(如图像识别和情感计算)中取得了巨大成功^[36], 在模型中适当加入注意力机制可以很好地提高模型性能。注意力机制可以使网络模型选择重要的可用信息, 舍弃不重要的信息, 具体实现方式是赋予重要信息较大权重, 从而使模型更关注特定的地方。

在情感识别中, 构建脑电特征图的每个像素点代表了不同时间点和不同通道的信息, 空间注意力机制可以找到脑电特征图空间中更重要位置的信息, 它代表 EEG 中更重要的时间点和重要通道, 能更全面地帮助提取 EEG 时域特征和发现更重要的通道。

在本文使用的空间注意力机制中, 对于预处理后的原始 EEG, 使用平均池化函数和最大池化函数, 可更好地提高注意力机制的表现力。将经过两种池化处理后的特征图拼接, 输入一个卷积核大小为 7×7 的卷积层, 采用 sigmoid 激活函数, 得到经过空间注意力机制的新特征, 上述步骤表示为:

$$X' = s(f^{7 \times 7}([\text{AvgPool}(X), \text{MaxPool}(X)])) \quad (1)$$

其中, $X' \in R^{H \times W}$ 表示经过空间注意力机制后被赋予不同权重的新特征图, H 和 W 表示特征图的高和宽; AvgPool 表示平均池化函数; MaxPool 表示最大池化函数; $f^{7 \times 7}$ 表示卷积核大小为 7×7 的卷积层; s 为激活函数。具体处理过程如图 2 所示。

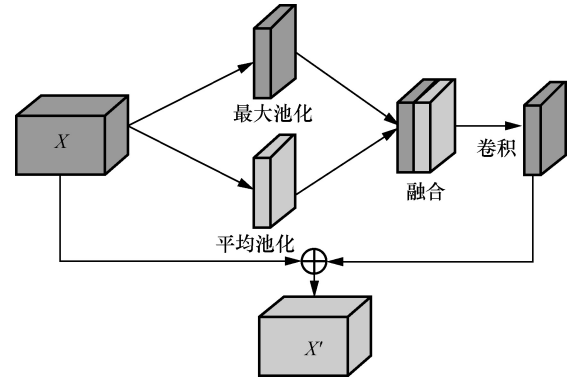


图2 具体处理过程

2.3 可分离卷积

随着网络模型深度的增加, 网络参数量越来越大, 模型运行速度越来越慢, 需要设计一种对脑电情感识别效率高的网络模型, 以更好地处理脑电情感识别任务。为了缓解传统 CNN 参数量大的缺点, 相比传统标准卷积, 可分离卷积在提取相同特征的情况下, 网络训练参数量更小^[37]。

可分离卷积通常分为空间可分离卷积和深度可分离卷积。空间可分离卷积将一个卷积核分成两个较小的卷积核分别进行卷积, 这可以使网络模型参数大大降低。具体来说, 对于一个 $m \times m$ 的

卷积核, 将其拆分为 $m \times 1$ 和 $1 \times m$ 的两个卷积核分别卷积。空间可分离卷积如图 3 所示, 将 3×3 的卷积核拆分为 3×1 和 1×3 的卷积核先后卷积。深度可分离卷积将卷积拆分为两个步骤进行, 分别进行逐通道卷积和逐点卷积。

PSC-LSR 采用了并行空间可分离卷积, 该模块一共搭建了 3 个并行层, 更多的并行层使模型能提取到 EEG 更丰富的深层情感特征。每个并行层包含了上下并行的两个块, 上层采用 3×3 的卷积核提取 EEG 各通道周围的情感信息, 下层采用 5×5 的卷积核提取 EEG 较远通道之间的情感特征。通过上下层结合来提取 EEG 的深层特征, 从而提取情感信息的全面表征。对于每个可分离卷积块, 采用了空间可分离卷积, 将大小为 $k \times k$ 的卷积核拆分为 $k \times 1$ 和 $1 \times k$ 的卷积核分别进行卷积, 并在中间加入了最大池化层, 大小为 2×1 , 避免了重叠池化问题, 可分离卷积块的内部结构如图 4 所示, 其中 H 、 W 、 L 分别表示块的高、宽、长。

2.4 损失函数

PSC-LSR 模型使用了 3 种损失函数——交叉熵损失、对比损失和标签平滑正则化损失, 并根据以上 3 种损失函数针对 EEG 情感识别提出了混合损失函数。

(1) 交叉熵损失

交叉熵损失能够表示情感标签真实概率分布与预测概率分布的异同程度, 交叉熵损失越小, 说明模型预测的情感与真实情感越接近, 可以表示为:

$$l_e = -\frac{1}{M} \sum_{k \in K} [y_k \ln(p_k) + (1 - y_k) \ln(1 - p_k)] \quad (2)$$

其中, l_e 表示交叉熵损失; K 表示所有的 EEG 情感样本; 对于第 k 个情感标签样本, 当它被分为高类别时, y_k 为 1, 当它被分为低类别时, y_k 为 0; p_k 表示模型预测情感样本的概率; M 表示情感标签的数量。

(2) 对比损失

对比损失在基于 EEG 的情感识别任务中有良好的表现^[35]。将相同的情感标签样本定义为正样本组, 不同的情感标签定义为负样本组, 使用对比损失可以减少正样本组之间的距离, 使负样本组之间间隔更远, 从而提高情感识别任务的分类准确率, 计算式为:

$$l_c = -\sum_{k \in K} \frac{1}{i} \ln \frac{\sum_{x^+ \in I} \exp\left(x_k \times \frac{x^+}{\alpha}\right)}{\sum_{x^- \in J} \exp\left(x_k \times \frac{x^-}{\alpha}\right)} \quad (3)$$

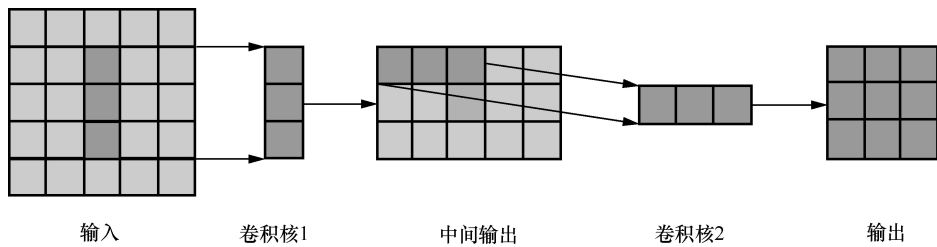


图3 空间可分离卷积

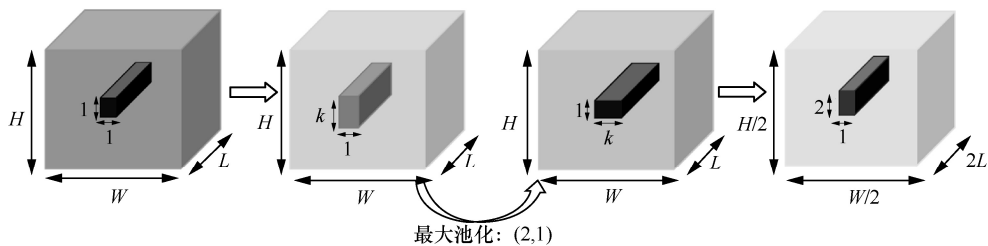


图4 可分离卷积块的内部结构



其中, l_c 表示对比损失, I 表示正样本集, J 表示负样本集, i 表示正样本集的数量, x_k 表示从每批量 EEG 样本集中随机选择的第 k 个样本, x^+ 表示与 x_k 具有相同标签的情感样本, x^- 表示与 x_k 具有不同标签的情感样本, K 表示所有的 EEG 情感样本, α 是超参数。

(3) 标签平滑正则化损失

标签平滑是一种标签优化策略, 它不同于传统分类任务中采用确定的标签作为硬目标, 而是采用硬目标的加权平均和标签上的均匀分布作为软目标。在传统任务分类中, 使用交叉熵损失优化目标时, 常将预测结果用独热 (one-hot) 向量编码形式表示, one-hot 向量只包含 0 和 1, 网络模型完全依赖标签的标注, 只考虑预测正确标签的损失, 而不考虑错误标签的信息, 网络模型通常在训练集上有很好的结果, 而在测试集上表现不好。标签平滑对分类错误的标签有一定的容错率, 它允许模型在一定范围内存在错误。对于错误的标签并不完全舍弃, 而是以较小的概率选择错误的标签。因此, 可以在很大程度上提高网络模型的泛化性和鲁棒性, 缓解网络模型的过拟合问题, 对分类任务有出色的表现。

在传统的基于 EEG 的情感识别任务中, 只采用交叉熵损失训练网络参数时, 由于它只能考虑情感标签样本中预测正确位置 (one-hot 标签中为 1 的位置) 的损失, 忽略了预测错误位置 (one-hot 标签中为 0 的位置) 的损失, 这是一种硬目标优化策略, 会使得模型完全信任人工标注情感标签的信息。但是, 由于被试者主观性或者长时间测试产生困倦等各种因素, 标注的情感标签不一定是完全正确的。如果完全相信人工标注的情感标签, 可能导致模型在训练集上有出色的表现, 但是在测试集上的表现不是很好。因此, 本文提出采用情感标签平滑正则化来解决上述问题, 可以使网络模型容许预测的标签有一定的容错能力, 标签平滑正则化损失表示为:

$$l_s = -\frac{1}{M} \sum_{k \in K} [y'_k \ln(p_k) + (1 - y'_k) \ln(1 - p_k)] \quad (4)$$

假设情感标签有 σ 的概率出错, 那么情感标签为真实值的概率为 $1 - \sigma$ 。其中, 对于第 k 个情感标签样本, 当它被分为高类别时, y'_k 为 $1 - \sigma$, 当它被分为低类别时, y'_k 为 σ ; p_k 表示模型预测情感样本的概率; M 表示情感标签的数量。

(4) 混合损失函数

使用交叉熵损失函数来度量预测的概率分布与情感标签标注的概率分布的距离; 对比损失函数使相同的情感标签样本分布更近, 不同的情感标签样本分布更远; 标签平滑正则化损失函数用来解决人工标注的情感样本可能出错的问题, 增强网络模型的泛化性。本文根据上述 3 种损失函数, 在基于 EEG 的情感识别任务中提出了混合损失函数, 表示为:

$$L = \mu \cdot l_s + \frac{1 - \mu}{2} \cdot l_c + \frac{1 - \mu}{2} \cdot l_e \quad (5)$$

其中, L 是模型要优化的混合损失函数值, l_s 表示情感标签平滑正则化损失, l_c 表示对比损失, l_e 表示交叉熵损失, μ 表示训练的超参数。

3 实验与分析

3.1 数据集及数据处理

本文使用 DEAP^[38]和 Dreamer^[39]两个公开的 EEG 数据集进行实验。DEAP 数据集中采集的 EEG 已下采样至 128 Hz, 并且使用 4~45 Hz 的带通滤波器滤除较高和较低波段的数据, 对 32 名被试者分别进行了 40 次实验, 每次实验给予被试者一定刺激, 记录了 63 s 内被试者的 32 个电极通道数据, 其中前 3 s 是基线信号, 被试者根据唤醒、效价、主导度打分 (分值为 1~9)。Dreamer 数据集是一个多模态数据库, 记录了 23 名被试者观看了 18 个电影片段的 EEG 和 EOG 数据, 这些片段时长从 63 s 到 393 s 不等, 然后被试者从唤醒、效价、主导度 3 个维度在评级表上打分 (分值为 1~5), EEG 数据记录了 14 个不同通道的电极信息。两个数据集的

数据结构见表1, 其中 M 表示平均值。

表1 两个数据集的数据结构

数据集	数组名称	数组形状	数组内容
DEAP	数据	$40 \times 40 \times 8\ 064$	实验次数 $\times$ 通道 $\times$ 数据
	标签	40×4	实验次数 $\times$ 标签
Dreamer	数据	$18 \times 14 \times 25\ 472 (M)$	实验次数 $\times$ 通道 $\times$ 数据
	标签	18×3	实验次数 $\times$ 标签

以 DEAP 数据集为例, 数据处理过程如图5所示, 将每个视频的 EEG 信号分为 n 段, DEAP 数据集 $n=60$, Dreamer 数据集 $n=63 \sim 393$, 每一段时间是 T (单位: s), 这里 $T=1$ s, 由于采集的 EEG 已下采样至 128 Hz, 所以每秒得到 128 个数据, 作为 EEG 情感信息矩阵的宽, 电极通道数量 E 作为 EEG 情感信息矩阵的长, DEAP 数据集 $E=32$, Dreamer 数据集 $E=14$ 。

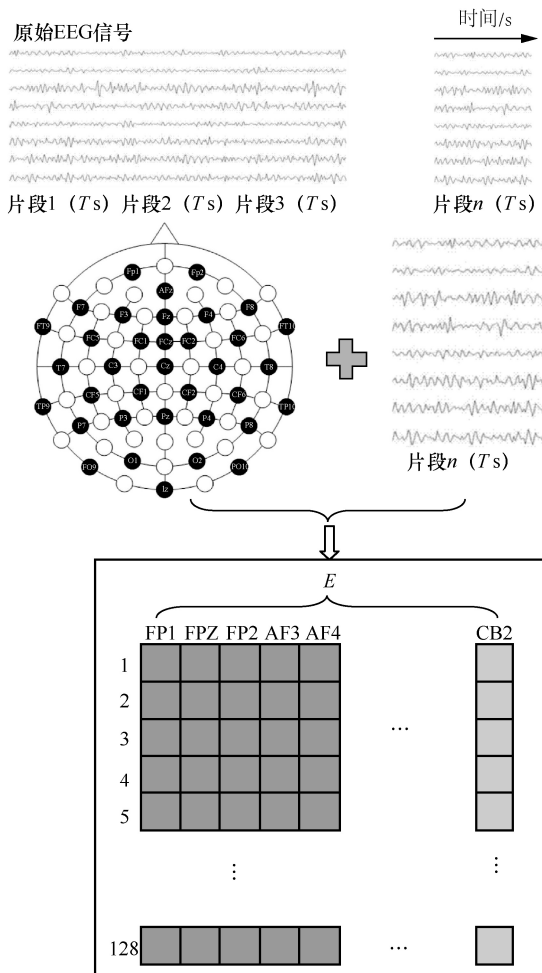


图5 数据处理过程

3.2 实验参数

DEAP 数据集包含 32 名被试者的数据, 每名被试者进行了 40 次实验, 每个实验持续 60 s, 因此每个被试者有 2 400 s 的数据, 每秒的数据生成的矩阵大小为 128×32 。Dreamer 数据集有 14 个电极通道, 每秒的数据生成的矩阵大小为 128×14 。DEAP 数据集中每个被试者填写自测表, 按照 1~9 打分, 阈值设置为 5, 分值为 5 以下则认为低唤醒/效价, 5 以及 5 分以上则认为高唤醒/效价。在 Dreamer 数据集中阈值为 3, 分值为 3 以下则认为低唤醒/效价, 3 以及 3 分以上则认为高唤醒/效价。本文将 σ 设置为 0.1, 通过实验发现, 在损失函数的占比 μ 为 0.2 时, LSR 可以取得最好的效果。采用学习率衰减策略, 起初在优化器中设定较大的学习率, 为 0.05, 以确保模型加速收敛, 并设置衰减步长为 0.000 01, 批处理大小设置为 128, 每轮训练 60 次, 采用十折交叉验证 (10-fold cross validation) 来验证每个被试者的表现。实验采用 Linux 操作系统, 使用 NVIDIA TITAN V100 的 GPU 进行模型训练, 内存为 32 GB, Python 版本为 Python3.7, 深度学习框架使用的是 PyTorch。网络模型具体参数见表2。

表2 网络模型具体参数

名称	DEAP 输出维度 (卷积核)	Dreamer 输出维度 (卷积核)
输入	128×32	128×14
空间注意力模块	128×32	128×14
卷积	$64 \times 128 \times 32 (3)$	$64 \times 128 \times 14 (3)$
并行层 1	$128 \times 64 \times 16 (3,5)$	$128 \times 64 \times 7 (3,5)$
并行层 2	$256 \times 32 \times 8 (3,5)$	$128 \times 32 \times 3 (3,5)$
并行层 3	$512 \times 16 \times 4 (3,5)$	$512 \times 16 \times 1 (3,5)$
全连接网络	32 768	8 192

3.3 实验结果

较深的网络更适合小的卷积核, 因为这样会减少参数量, 所以上层卷积结构采用的卷积核为 3×3 , 设置了 3 个块结构。对于下层网络结构, 采用的卷积核为 5×5 。对于具体层数, 在 DEAP 和 Dreamer 数据集上做了相关实验, 发现下层结构为 3 层时可



以取得最好的效果,在 DEAP 数据集的唤醒和效价的准确率分别为 99.23%和 99.13%。当下层网络卷积层数为 0,即不采用并行网络,采用单层网络提取特征时,取得的结果最差,采用并行可分离卷积网络相比单层网络在 DEAP 数据集的唤醒和效价维度分别提升了 0.51%和 0.17%。下层架构中选取不同卷积层数在 DEAP 数据集上的结果见表 3。

表 3 下层架构中选取不同卷积层数在 DEAP 数据集上的结果 (平均值±标准差)

层数	唤醒	效价
0	98.72%±1.52%	98.96%±1.05%
1	99.14%±0.83%	98.86%±1.32%
2	99.01%±1.31%	98.98%±1.13%
3	99.23%±0.91%	99.13%±0.76%

对损失函数中超参数 μ 的确认,经过了大量的实验验证。 μ 代表加入的情感标签平滑正则化损失在总损失函数的占比, μ 的大小在一定程度上反映了模型对预测错误标签的容忍概率,分别令 μ 为 0、0.1、0.2、0.3 进行实验,不同 μ 值在 DEAP 数据集上的识别准确率见表 4。当 $\mu=0$,即不采用 LSR 损失时,情感识别准确率最低;当 $\mu=0.2$ 时,识别准确率最高,在 DEAP 数据集的唤醒和效价的识别准确率分别为 99.23%和 99.13%。相比不采用 LSR 损失,其性能分别提升了 0.3%和 0.51%,这说明采用 LSR 损失是很有必要的,因为人工标注的情感数据集在一些因素的影响下有可能出现错误。LSR 损失提高了模型对预测错误标签的容忍性,使网络模型在测试集上表现得很好,提高了识别准确率。

表 4 不同 μ 值在 DEAP 数据集上的识别准确率 (平均值±标准差)

μ	唤醒	效价
0	98.93%±1.14%	98.62%±1.24%
0.1	99.06%±1.04%	99.00%±0.87%
0.2	99.23%±0.91%	99.13%±0.76%
0.3	98.97%±1.14%	98.94%±0.88%

对于空间注意力机制,分别在模型中测试了使用和不使用 SAM 的结果,实验结果对比见表 5。

采用空间注意力机制在 DEAP 数据集的唤醒和效价维度的识别准确率分别提升了 0.48%和 0.25%,SAM 初步提取了浅层的情感特征,对 EEG 信号重要时间点和重要通道进行权重再分配是有必要的。

表 5 实验结果对比 (平均值±标准差)

对比项	唤醒	效价
使用 SAM	99.23%±0.91%	99.13%±0.76%
不使用 SAM	98.75%±1.36%	98.88%±1.01%

为了进一步评估模型的性能,在 DEAP 和 DREAMER 数据集上将 PSC-LSR 和最近 3 年其他方法进行对比,不同方法在 DEAP、DREAMER 数据集上的识别准确率分别见表 6 和表 7。在 DEAP 数据集上,PSC-LSR 相比 DGCNN 在唤醒和效价维度的识别准确率分别提高了 5.73%、6.58%,相比 ECNN-C 在唤醒和效价维度的识别准确率分别提高了 0.72%、0.78%。与其他深度学习方法对比,PSC-LSR 相比 MTCA-Capsnet 在唤醒和效价维度的识别准确率分别提高了 1.82%、1.89%,相比 gcForest 在唤醒和效价维度的识别准确率分别提高了 1.70%、1.44%。类似地,在 DREAMER 数据集上,与其他方法对比,PSC-LSR 也取得了最高的识别准确率。

表 6 不同方法在 DEAP 数据集上的识别准确率

方法	唤醒	效价
DGCNN ^[40]	93.50%	92.55%
MTCA-Capsnet ^[41]	97.41%	97.24%
gcForest ^[42]	97.53%	97.69%
ECNN-C ^[35]	98.51%	98.35%
PSC-LSR	99.23%	99.13%

表 7 不同方法在 DREAMER 数据集上的识别准确率

方法	唤醒	效价
DGCNN ^[40]	88.93%	89.59%
MTCA-Capsnet ^[41]	95.54%	94.96%
gcForest ^[42]	90.41%	89.03%
ECNN-C ^[35]	97.03%	96.89%
PSC-LSR	97.39%	97.22%

在 DEAP 数据集上对 32 名被试者在唤醒和效价维度上分别进行实验,不同被试者在 DEAP 数

据集上的识别准确率如图6所示。2号和4号被试者准确率比较低,之后5~10号基本一致,在11号出现拐点,准确率降低,12~16号基本稳定,17号准确率较低,18~22号准确率逐渐上升,25号之后除27、29、30号准确率较高外,大致呈下降趋势,其中31、32号两名被试者准确率最低。

32名被试者唤醒和效价的准确率走势与ECNN-C、MTCA-Capsnet等深度学习方法基本一致,证明了所提方法在DEAP数据集上的有效性。

同样,在Dreamer数据集上对23名被试者在唤醒和效价维度上分别进行实验,不同被试者在Dreamer数据集上的识别准确率如图7所示。4号和

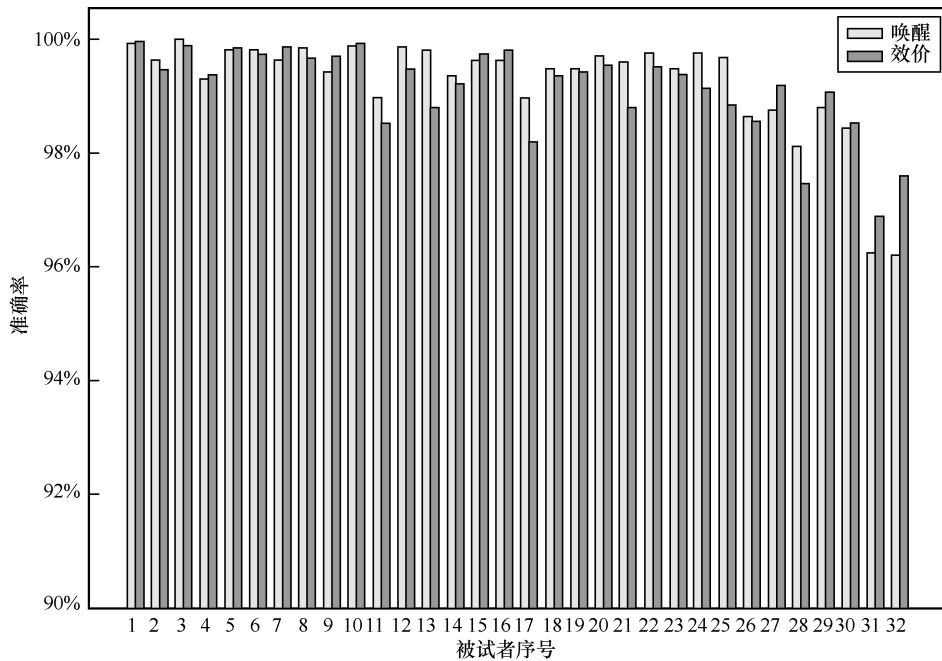


图6 不同被试者在DEAP数据集上的识别准确率

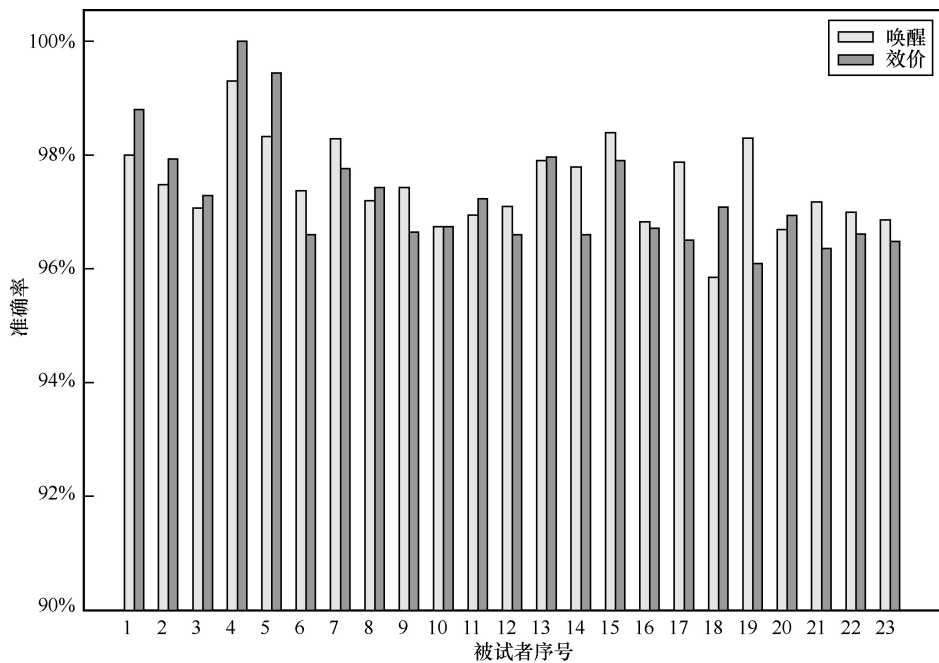


图7 不同被试者在Dreamer数据集上的识别准确率



5 号被试者准确率都非常高, 之后准确率上下起伏。此外, 1~5 号被试者效价准确率高于唤醒, 6~23 号被试者总体来看唤醒准确率大于效价。23 名被试者唤醒和效价的准确率走势与 ECNN-C、MTCA-Capsnet 等深度学习方法基本一致, 证明了所提方法在 Dreamer 数据集上的有效性。

为了充分展示 PSC-LSR 提取的 EEG 情感特征, 本文使用 t 分布—随机邻近嵌入 (t -distributed stochastic neighbor embedding, t -SNE) 方法, 对模型在每个模块提取到的情感特征进行可视化展示。DEAP 测试集的情感标签在 PSC-LSR 不同阶

段的特征分布如图 8 所示。

图 8 中, 圆形代表标签为积极, 星形代表标签为消极。起初, 原始特征分布非常散乱, 经过 SAM 初步提取情感浅层特征后, 两类标签进行了初步分离。之后经过并行可分离卷积模块提取 EEG 情感的深层特征后, 积极与消极明显分离出来, 分布在不同区域。最后, 为了实现分类, 将特征输入全连接网络, 积极与消极分别紧凑聚集在一起, 且彼此间距较远, 表明 PSC-LSR 在 EEG 的情感识别中分类准确率高, 鲁棒性好。

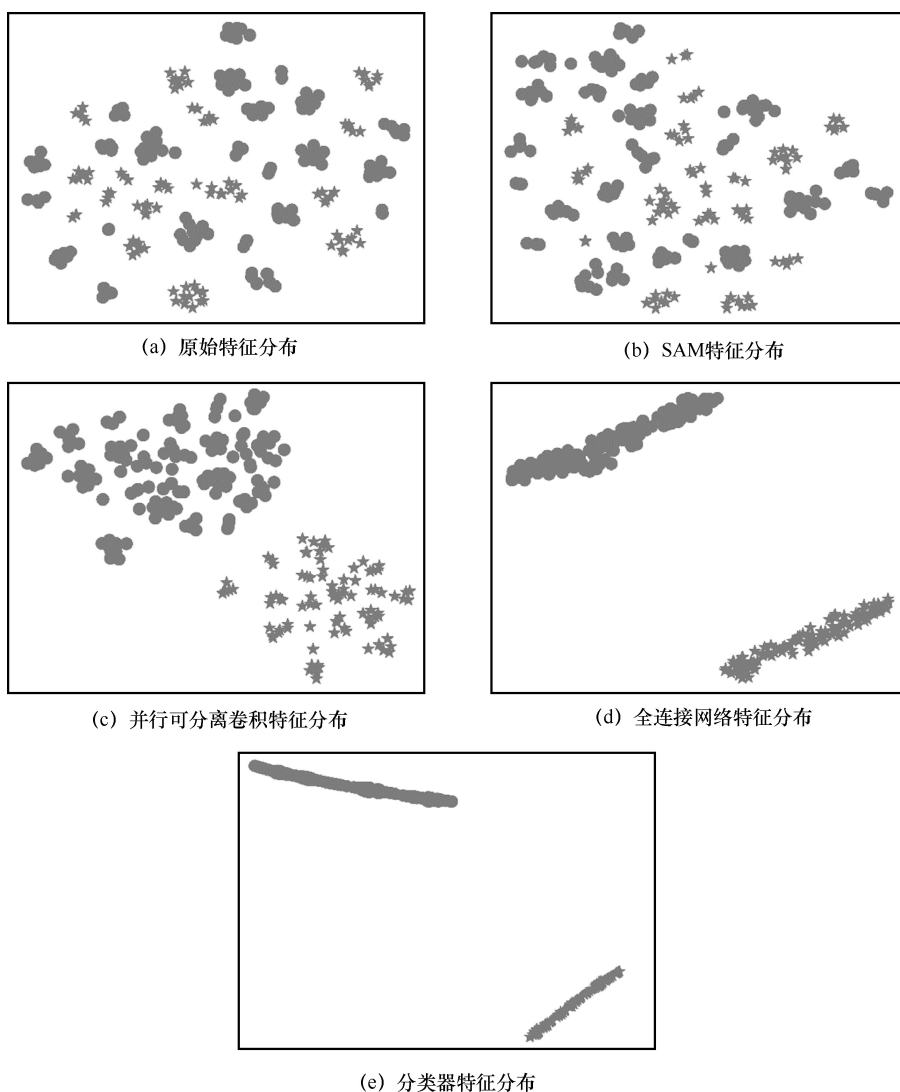


图 8 DEAP 测试集的情感标签在 PSC-LSR 不同阶段的特征分布

4 结束语

本文提出了一种新的深度学习方法——PSC-LSR, 用于 EEG 情感识别。PSC-LSR 使用空间注意力机制与并行可分离卷积网络对 EEG 进行全面的情感特征提取, 采用标签平滑正则化方法提高对人工标注情感标签的容错概率, 设计新的混合损失函数提高模型的泛化性能。本文提出的方法在 DEAP 和 Dreamer 数据集上取得了较好的分类效果。尽管脑电信号很好地描述了情感特征, 但个体间情感特征的差异性导致情感识别模型难以有效地进行迁移学习, 未来将进一步探究域适配策略来解决情感识别模型的跨域迁移问题。

参考文献:

- [1] GAO Z K, LI Y L, YANG Y X, et al. A GPSO-optimized convolutional neural networks for EEG-based emotion recognition[J]. *Neurocomputing*, 2020(380): 225-235.
- [2] CHEN G J, ZHANG X Y, SUN Y, et al. Emotion feature analysis and recognition based on reconstructed EEG sources[J]. *IEEE Access*, 2020(8): 11907-11916.
- [3] WANG Z M, ZHOU X X, WANG W L, et al. Emotion recognition using multimodal deep learning in multiple psychophysiological signals and video[J]. *International Journal of Machine Learning and Cybernetics*, 2020, 11(4): 923-934.
- [4] KESSOUS L, CASTELLANO G, CARIDAKIS G. Multimodal emotion recognition in speech-based interaction using facial expression, body gesture and acoustic analysis[J]. *Journal on Multimodal User Interfaces*, 2010, 3(1): 33-48.
- [5] ANDERSON K, MCOWAN P W. A real-time automated system for the recognition of human facial expressions[J]. *IEEE Transactions on Systems, Man, and Cybernetics Part B, Cybernetics: a Publication of the IEEE Systems, Man, and Cybernetics Society*, 2006, 36(1): 96-105.
- [6] 刘晓旻, 谭华春, 章毓晋. 人脸表情识别研究的新进展[J]. *中国图象图形学报*, 2006, 11(10): 1359-1368.
LIU X M, TAN H C, ZHANG Y J. New research advances in facial expression recognition[J]. *Journal of Image and Graphics*, 2006, 11(10): 1359-1368.
- [7] 刘玉娟, 方富熹. 情绪的语音交流[J]. *中国行为医学科学*, 2007, 16(4): 374-376.
LIU J Y, FANG F X. Voice communication of emotions[J]. *Chinese Journal of Behavioral Medical and Brain Science*, 2007, 16(4): 374-376.
- [8] ANG J, DHILLON R, KRUPSKI A, et al. Prosody-based automatic detection of annoyance and frustration in human-computer dialog[C]//*Proceedings of 7th International Conference on Spoken Language Processing (ICSLP 2002)*. [S.l.:s.n.], 2002: 2037-2040.
- [9] EKMAN P. An argument for basic emotions[J]. *Cognition and Emotion*, 1992, 6(3/4): 169-200.
- [10] PLUTCHIK R. The nature of emotions: human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice[J]. *American Scientist*, 2001, 89(4): 344-350.
- [11] RUSSELL J A. A circumplex model of affect[J]. *Journal of Personality and Social Psychology*, 1980, 39(6): 1161-1178.
- [12] MEHRABIAN A. Comparison of the PAD and PANAS as models for describing emotions and for differentiating anxiety from depression[J]. *Journal of Psychopathology and Behavioral Assessment*, 1997, 19(4): 331-357.
- [13] SHU L, XIE J Y, YANG M Y, et al. A review of emotion recognition using physiological signals[J]. *Sensors (Basel, Switzerland)*, 2018, 18(7): 2074.
- [14] DAI Y X, WANG X, ZHANG P B, et al. Sparsity constrained differential evolution enabled feature-channel-sample hybrid selection for daily-life EEG emotion recognition[J]. *Multimedia Tools and Applications*, 2018, 77(17): 21967-21994.
- [15] YU X Y, CHUM P, SIM K B. Analysis the effect of PCA for feature reduction in non-stationary EEG based motor imagery of BCI system[J]. *Optik*, 2014, 125(3): 1498-1502.
- [16] FAYEK H M, LECH M, CAVEDON L. Modeling subjectiveness in emotion recognition with deep neural networks: ensembles vs soft labels[C]//*Proceedings of 2016 International Joint Conference on Neural Networks (IJCNN)*. Piscataway: IEEE Press, 2016: 566-570.
- [17] RINGEVAL F, EYBEN F, KROUPI E, et al. Prediction of asynchronous dimensional emotion ratings from audiovisual and physiological data[J]. *Pattern Recognition Letters*, 2015(66): 22-30.
- [18] ZHONG P X, WANG D, MIAO C Y. EEG-based emotion recognition using regularized graph neural networks[J]. *IEEE Transactions on Affective Computing*, 2022, 13(3): 1290-1301.
- [19] PORBADNIGK A K, GÖRNITZ N, SANNELLI C, et al. When brain and behavior disagree: tackling systematic label noise in EEG data with machine learning[C]//*Proceedings of 2014 International Winter Workshop on Brain-Computer Interface (BCI)*. Piscataway: IEEE Press, 2014: 1-4.
- [20] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway: IEEE Press, 2016: 2818-2826.
- [21] KHALILI Z, MORADI M H. Emotion recognition system using brain and peripheral signals: using correlation dimension to improve the results of EEG[C]//*Proceedings of 2009 International Joint Conference on Neural Networks*. Piscataway: IEEE Press, 2009: 1571-1575.



- [22] 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.
- [23] MURUGAPPAN M, RIZON M, NAGARAJAN R, et al. Time-frequency analysis of EEG signals for human emotion detection[C]//Proceedings of 4th Kuala Lumpur International Conference on Biomedical Engineering 2008. Heidelberg: Springer, 2008: 262-265.
- [24] KHOSROWABADI R, QUEK H C, WAHAB A, et al. EEG-based emotion recognition using self-organizing map for boundary detection[C]//Proceedings of 2010 20th International Conference on Pattern Recognition. Piscataway: IEEE Press, 2010: 4242-4245.
- [25] LIU W, ZHENG W L, LU B L. Emotion recognition using multimodal deep learning[C]//Proceedings of the 23rd International Conference on Neural Information Processing Springer International Publishing. Heidelberg: Springer, 2016: 521-529.
- [26] HASSAN M M, ALAM M G R, UDDIN M Z, et al. Human emotion recognition using deep belief network architecture[J]. Information Fusion, 2019(51): 10-18.
- [27] VIJAYAKUMAR S, FLYNN R, MURRAY N. A comparative study of machine learning techniques for emotion recognition from peripheral physiological signals[C]//Proceedings of 2020 31st Irish Signals and Systems Conference (ISSC). Piscataway: IEEE Press, 2020: 1-6.
- [28] MERT A, AKAN A. Emotion recognition from EEG signals by using multivariate empirical mode decomposition[J]. Pattern Analysis and Applications, 2018, 21(1): 81-89.
- [29] WU S Y, XU X M, SHU L, et al. Estimation of valence of emotion using two frontal EEG channels[C]//Proceedings of 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Piscataway: IEEE Press, 2017: 1127-1130.
- [30] CHENG C L, WEI X W, JIAN Z. Emotion recognition algorithm based on convolution neural network[C]//Proceedings of 2017 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE). Piscataway: IEEE Press, 2018: 1-5.
- [31] KWON Y H, SHIN S B, KIM S D. Electroencephalography based fusion two-dimensional (2D)-convolution neural networks (CNN) model for emotion recognition system[J]. Sensors, 2018, 18(5): 1383.
- [32] CUI H, LIU A P, ZHANG X, et al. EEG-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network[J]. Knowledge-Based Systems, 2020(205): 106243.
- [33] LI D D, CHAI B, WANG Z, et al. EEG emotion recognition based on 3-D feature representation and dilated fully convolutional networks[J]. IEEE Transactions on Cognitive and Developmental Systems, 2021, 13(4): 885-897.
- [34] LIU W, QIU J L, ZHENG W L, et al. Comparing recognition performance and robustness of multimodal deep learning models for multimodal emotion recognition[J]. IEEE Transactions on Cognitive and Developmental Systems, 2022, 14(2): 715-729.
- [35] LI C, LIN X J, LIU Y, et al. EEG-based emotion recognition via efficient convolutional neural network and contrastive learning[J]. IEEE Sensors Journal, 2022, 22(20): 19608-19619.
- [36] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM Press, 2017: 6000-6010.
- [37] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint, 2017, arXiv: 1704.04861.
- [38] 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.
- [39] KATSIKIANNIS S, RAMZAN N. DREAMER: a database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices[J]. IEEE Journal of Biomedical and Health Informatics, 2018, 22(1): 98-107.
- [40] SONG T F, ZHENG W M, SONG P, et al. EEG emotion recognition using dynamical graph convolutional neural networks[J]. IEEE Transactions on Affective Computing, 2020, 11(3): 532-541.
- [41] CHENG J, CHEN X. Emotion recognition from EEG based on multi-task learning with capsule network and attention mechanism[J]. Computers in Biology and Medicine, 2022(143): 105303.
- [42] CHENG J, CHEN M Y, LI C, et al. Emotion recognition from multi-channel EEG via deep forest[J]. IEEE Journal of Biomedical and Health Informatics, 2021, 25(2): 453-464.

[作者简介]



张永 (1975—), 男, 博士, 湖州师范学院教授、博士生导师, 辽宁师范大学计算机与信息技术学院教授, 主要研究方向为数据挖掘、智能计算、情感计算。



刘纪奎 (2000—), 男, 湖州师范学院在读, 主要研究方向为数据挖掘、情感计算。



柯文龙 (1989—), 男, 博士, 湖州师范学院讲师、硕士生导师, 主要研究方向为机器学习、软件定义网络。