基于脑电信号和眼动信号的情感识别

李 澳、张恒千、尹 钟*

上海理工大学光电信息与计算机工程学院,上海

收稿日期: 2024年12月27日; 录用日期: 2025年1月20日; 发布日期: 2025年1月29日

摘要

本文探讨了基于脑电信号和眼动信号的多模态情感识别方法,使用了上海交通大学SEED-IV和SEED-V数据集。通过预处理脑电信号并提取功率谱密度和微分熵特征,结合眼动信号的时域和频域特征,提出了一种基于对比学习的数据融合方法,显著提高了情感识别的准确率。实验结果表明,多模态融合,尤其是脑电信号与眼动信号的结合,能显著提升识别性能;同时,微分熵特征优于功率谱密度特征,具有更好的分类准确性。

关键词

情感识别,对比学习,机器学习,特征融合

Emotion Recognition Based on EEG Signals and Eye Movement Signals

Ao Li, Hengqian Zhang, Zhong Yin*

School of Optical-Electrical Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Dec. 27th, 2024; accepted: Jan. 20th, 2025; published: Jan. 29th, 2025

Abstract

This paper explores a multimodal emotion recognition method based on EEG and eye movement signals, using the SEED-IV and SEED-V datasets from Shanghai Jiao Tong University. After preprocessing the EEG signals, power spectral density and differential entropy features were extracted, while time-domain and frequency-domain features were extracted from the eye movement signals. A contrastive learning-based data fusion method is proposed, which effectively bridges the distribution gap between different modalities and significantly improves emotion recognition accuracy.

*通讯作者。

文章引用: 李澳, 张恒千, 尹钟. 基于脑电信号和眼动信号的情感识别[J]. 建模与仿真, 2025, 14(1): 1244-1252. DOI: 10.12677/mos.2025.141112

Experimental results show that multimodal fusion, particularly the combination of EEG and eye movement signals, significantly enhances recognition performance. Additionally, the differential entropy feature of EEG signals outperforms the power spectral density feature in classification accuracy, indicating better separability in the dataset.

Keywords

Emotional Recognition, Comparative Learning, Machine Learning, Feature Fusion

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

http://creativecommons.org/licenses/by/4.0/



Open Access

1. 引言

情感计算作为一种新兴的跨学科研究领域,它旨在通过计算机分析和处理人类情感信息,实现对人类情感状态的自动识别和理解[1]。随着人工智能技术的逐渐发展,情感计算已经成为了人工智能领域新兴的研究方向,对社会发展的必要性也与日俱增。

目前,情感计算技术主要使用的是脑电信号(Electroencephalogram, EEG)进行情感识别[2]。相较于面部表情、语音等信号,脑电信号反应速度快、准确性高,能更直接反映人的内在情感状态,且不易受到外在表现的干扰。分析脑电信号可以精确判断个体情感,为情感干预提供科学依据。此外,脑电信号与其他多模态生理信号结合,可大幅提升情感识别的综合准确度,展现了其在情感计算技术中不可替代的重要价值[3]。

使用基于脑电的多模态生理信号进行情感识别有着诸多优势。首先,在不同情境下,人体的生理信号会产生不同的变化,如心率、血压、皮肤电阻等。通过对这些多模态生理信号的综合分析,可以更加全面地了解个体的情感状态[4];其次,多模态生理信号的综合分析可以减少单一生理信号的局限性[5]。例如,某些情感状态可能只有在特定的生理信号下才能被识别出来,而单独使用脑电信号时可能无法捕捉到这种情感状态。因此,多模态生理信号的综合分析可以提高情感识别的准确性和稳定性[6]。

本文是基于脑电信号和眼动信号两种模态的信息,使用对比学习的思想搭建两种信号的融合模型。 本文使用上海交通大学公开数据集 SEED-IV和 SEED-V数据集对设计的模型进行验证,并使用多种机器 学习分类模型进行分类。本文贡献如下:

- (1) 针对 SEED-IV和 SEED-V两个数据集,本文提出了可行的预处理过程并提取脑电的功率谱密度 (Power spectral density, PSD)和微分熵(Differential Entropy, DE)特征以及眼动特征。探究不同脑电特征的分类精度以及融合特征与单一特征的分类精度。
- (2) 本文探索了一种新的数据融合方法,使用对比学习的思想,拉近来自被试同一试验的脑电特征和 眼动特征这两个配对特征的分布,并融合不同模态间的特征信息,利用机器学习方法,极大地提高情感 识别的精度。

2. 情感数据集以及特征提取

2.1. SEED-IV和 SEED-V数据集

SEED-IV 数据集[7]和 SEED-V数据集[8]是由上海交通大学类脑计算与机器智能研究中心设计试验并采集的公开数据集。如表 1 所述, SEED-IV 数据使用了 72 段精心挑选的影片片段(SEED-V数据集的受试

者观看 15 段不同类型的情绪刺激材料),这些片段旨在引发快乐、悲伤、恐惧或中性情绪(SEED-V数据集的实验引发快乐、悲伤、恐惧、中性和恶心情绪)。共有 15 名受试者参与了实验(SEED-V数据集共有 16 名受试者),每个受试者分别在不同的日子进行了 3 次实验会话,每次会话包含 24 个试验。(SEED-V数据集的每次实验包含 15 个试验)在每个试验中,受试者观看其中一个影片片段,同时记录他们的脑电信号和眼动信号。这些数据通过 62 通道的 ESI NeuroScan 系统和 SMI 眼动追踪眼镜收集。

Table 1. Dataset information 表 1. 数据集信息

数据参数	SEED-IV	SEED-V
受试者/个	15	16
实验/次	3	3
试验/个	24	15
脑电通道/个	62	62
采样率/Hz	200	200
刺激类型	电影片段	电影片段

2.1.1. 数据预处理

脑电信号预处理是分析脑电信号数据的关键步骤。本文对原始脑电信号数据进行了多个预处理步骤:首先,将脑电信号降采样至 128 Hz,以减少数据中的计算负载和噪声;其次,本文执行了全局平均参考,通过将脑电信号数据减去所有电极当前时刻的平均值来消除共同噪声并提高脑电信号数据的信噪比;第三,本文应用了一个 4~45 Hz 的带通滤波器,以消除脑电信号数据中不需要的低频漂移和高频噪声;第四,本文使用独立成分分析来消除脑电信号数据中的眨眼伪迹;最后,本文将预处理的脑电信号数据分段为持续时间 4 秒的不重叠时段。

2.1.2. 数据特征提取

本文应用连续小波变换来提取不同频带的特征,包括: Theta (4.0~7.0 Hz), Alpha (7.0~13.0 Hz), lower Beta (13.0~20.0 Hz), upper Beta (20.0~30.0 Hz)和 Gamma (30.0~45.0 Hz) [9]共 5 个频段的特征。频率分辨率设置为 0.1 Hz,以捕获 EEG 信号的细粒度谱信息。

脑电信号中的功率谱密度和微分熵特征提取是认知神经科学和情绪识别研究中的关键步骤。以下是对这两种特征提取方法的详细阐述。

PSD 特征[10]用于衡量 EEG 信号在各个频带中的能量强弱。功率谱密度是对随机信号频域特性的分析,特别是脑电信号。尽管随机信号的总能量是无限的,但其平均功率是有限的。因此,分析脑电信号时,通常使用功率谱密度来描述其频域特性。

DE 特征[11]用于评估 EEG 信号的随机性或不确定性,与信号的复杂度相关。微分熵是香农熵在连续信号上的推广,用于量化连续随机变量的概率分布中的不确定性总量。

在 SEED-IV 数据集和 SEED-V数据集中,眼动信息的收集使用了 SMI 眼动追踪眼镜。每次实验中收集了包括眨眼、事件、注视、瞳孔直径和眼跳等不同类型的眼动信息。SEED-IV 数据集和 SEED-V数据集从这些原始数据中提取了多种眼动特征,如瞳孔直径、注视、眼跳和眨眼的持续时间等,以便于进行更深入的情绪识别分析。

3. 多模态生理数据融合

本文设计了一个对比学习[12]模型,使得每位被试来自同一实验的脑电特征向量和眼动特征向量具

有最大的相似性。然后,将两种特征向量使用多头注意力机制进行融合,生成融合特征向量。并使用融 合特征向量进行分类。

3.1. 特征向量提取

如图 1 所示,本研究对脑电信号和眼动信号进行了彻底的预处理和特征提取,进而实现了这两种信号的有效融合。具体来说,脑电信号首先通过预处理流程,从预处理后的脑电信号 $x(t) \in \mathbb{R}^{n\times d}$ 中提取出 PSD 和 DE 两种特征 $x_{PSD}, x_{DE} \in \mathbb{R}^{5\times n}$ 。其中,n 表示脑电的导联数,d 代表每个导联的数据点个数,5 表示对应的 5 个频段。以 PSD 特征为例,该特征被送入定制的脑电编码器中,从而转换为具有表征性的脑电特征向量

$$\boldsymbol{x}_{eeg} = \boldsymbol{\varphi}^{\text{ReLU}} \left(\boldsymbol{W}_{2}^{\text{Liner}} \left(\boldsymbol{\varphi}^{\text{ReLU}} \left(\boldsymbol{W}^{\text{Liner}} \boldsymbol{x}_{PSD} \right) \right) \right)$$
 (1)

其中 $\mathbf{x}_{eeg} \in \mathbf{R}^c$, c 代表了向量的长度。同理,眼动信息在预处理后得到相应的眼动特征 $\mathbf{z} \in \mathbf{R}^{m \times k}$ 。其中 \mathbf{m} 和 k 代表了眼动的特征种类和对应的特征数量。该特征经过专用的眼动编码器处理,转化为眼动特征向量

$$\mathbf{x}_{eye} = \varphi^{\text{ReLU}} \left(\mathbf{W}_{2}^{\text{Liner}} \left(\varphi^{\text{ReLU}} \left(\mathbf{W}^{\text{Liner}} \mathbf{z} \right) \right) \right)$$
 (2)

其中, $\mathbf{x}_{eve} \in \mathbf{R}^c$ 。

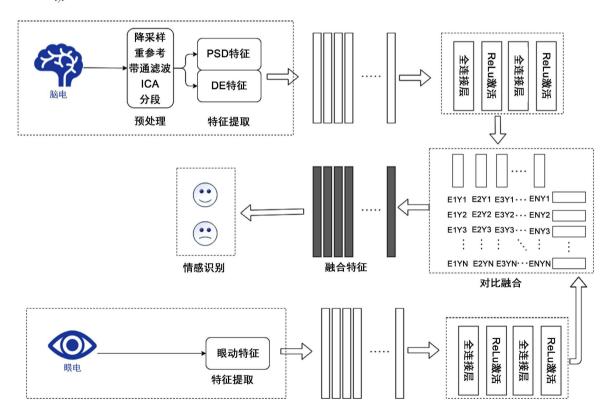


Figure 1. Multi modal data fusion method 图 1. 多模态数据融合方法

3.2. 对比融合

对比学习作为一种强大的无监督学习范式,通过最大化正样本间的相似度同时最小化负样本间的相

似度,可有效地学习到数据的深层次表示。这种方法的核心优势在于其高效的数据利用能力,能够从未标记的大规模数据集中学习到鲁棒性高且具有泛化能力的特征表示。此外,对比学习减少了对昂贵标签数据的依赖,提高了模型在实际应用中的可行性和效率。

具体而言,使用对比学习可以有效地对齐不同模态信号的特征,使得在经过对比学习优化的特征空间中,相似模态的样本彼此更为接近,而不相似模态的样本则彼此远离。通过这种方式,不同模态的特征被投射到一个共享的特征空间中,从而能够增强这些特征之间的语义一致性。在进行特征融合时,这种对齐效果显著提高了融合后特征的分类效果。融合后的特征不仅保留了各自模态的丰富信息,还通过对比学习实现了模态间的有效对齐,从而使得分类器能够更好地利用融合特征进行精确的分类。

在本文中,我们利用对比学习的思想,将被试在一次试验中的脑电特征向量与眼动特征向量定义为 正样本,将其他向量之间的关系定义为负样本。我们希望被试在一次试验中的脑电特征向量和眼动特征 向量在同一空间内的表征尽可能的相似,同时将脑电的特征向量与其他眼动特征向量尽可能的推远。这 种方法促使来自两个模态的特征向量在高维空间中更加密切地聚集在一起,从而提升了其在情感识别等 任务中的表现。

具体来说,在得到对应的脑电特征向量 \mathbf{x}_{eeg} 和眼动特征向量 \mathbf{x}_{eye} 后,由公式(3)计算对应特征的相似度:

$$f_{\text{SIM}}\left(\boldsymbol{x}_{eeg}, \boldsymbol{x}_{eye}\right) = \frac{\boldsymbol{x}_{eeg}^{\text{T}} \cdot \boldsymbol{x}_{eye}}{\left\|\boldsymbol{x}_{eeg}\right\|_{2} \left\|\boldsymbol{x}_{eye}\right\|_{2}} \tag{3}$$

其中"·"表示向量的点乘, $\|\cdot\|_2$ 表示计算 L2 范数。这里使用 $x/\|x\|_2$ 范数限制,使得计算的相似程度被限制在[0,1]之间,可以为后续的分类提供更好、更鲁棒的特征表征。对比损失旨在最大化对角线位置表征的相似性,使用如下损失函数:

$$L_{1} = -\frac{1}{2} \left[\log \frac{\exp\left(f_{\text{SIM}}\left(\boldsymbol{x}_{eeg}^{i}, \boldsymbol{x}_{eye}^{i}\right)/\tau\right)}{\sum_{j=1}^{N} \exp\left(f_{\text{SIM}}\left(\boldsymbol{x}_{eeg}^{i}, \boldsymbol{x}_{eye}^{j}\right)/\tau\right)} + \log \frac{\exp\left(f_{\text{SIM}}\left(\boldsymbol{x}_{eeg}^{i}, \boldsymbol{x}_{eye}^{i}\right)/\tau\right)}{\sum_{j=1}^{N} \exp\left(f_{\text{SIM}}\left(\boldsymbol{x}_{eeg}^{j}, \boldsymbol{x}_{eye}^{i}\right)/\tau\right)} \right)$$

$$(4)$$

其中,N 为批量大小, \mathbf{x}_{eeg}^{i} 是第 i 个脑电特征向量, \mathbf{x}_{eeg}^{i} 是与 \mathbf{x}_{eye}^{i} 配对的眼动特征向量, τ 是温度参数,用于控制软化概率分布的程度。本文使用的损失函数旨在确保两种不同模态的编码器能够在处理相同情感数据时,提取出尽可能相似的情感向量。

3.3. 融合特征

在经过对比学习后的两种特征表征如果直接进行拼接,会损失一部分模态间的互补信息,导致分类精度的下降。为了有效地整合这些特征并保留模态间的互补信息,多头自注意力机制提供了一个灵活的框架,允许模型在不同的特征子空间中自适应地关注不同的信息。

具体来说,每个头在处理输入特征时可以关注到不同的信息维度,从而捕获不同的依赖关系和互补性质,这是单一注意力机制难以实现的。通过这种方式,我们能够在保持特征独立性的同时,有效地融合不同模态之间的信息。通过引入多头自注意力机制,我们的模型不仅能够捕捉到复杂的、跨模态的特征依赖关系,还能够通过自适应地调整注意力分配来增强模型的泛化能力。最终,在多个数据集上的实

验结果验证了我们方法的有效性,相较于直接拼接特征的基线模型,采用多头自注意力机制显著提高了分类任务的精度。

根据特征向量,自主意力融合模块会生成对应的线性变阵矩阵,分别为 $W_Q, W_K, W_V \in \mathbb{R}^{n \times 2}$,将两种对应的特征表征拼接为一个特征向量并与这三个矩阵分别相乘,得到对应的 Q 矩阵,K 矩阵和 V 矩阵, $O.K.V \in \mathbb{R}^{2 \times 2}$ 。之后将这三个矩阵分别经过全连接层,并进行自注意力操作得到矩阵 $\mathbf{x}_{\cdot} \in \mathbb{R}^{n \times 2}$,其中

$$\boldsymbol{x}_{c} = \boldsymbol{\varphi}^{\text{softmax}} \left(\frac{\boldsymbol{W}^{\text{Liner}} \left(\boldsymbol{Q} \right) \boldsymbol{W}^{\text{Liner}} \left(\boldsymbol{K} \right)^{\text{T}}}{\sqrt{d_{k}}} \right) \boldsymbol{W}^{\text{Liner}} \left(\boldsymbol{V} \right)$$
 (5)

 d_k 为 Q、K 矩阵的列数,用以防止内积过大。将 Q、K、V 矩阵分别乘以 m 个线性层后,会得到 m 个结果,将这 m 个结果进行拼接后在

经过一层全连接层,恢复为输入的形状得到结果。最后的输出为:

$$\boldsymbol{x}_{out} = \boldsymbol{\varphi}^{\text{Dropout}} \left[\boldsymbol{W}^{\text{ReLU}} \left(\boldsymbol{W}^{\text{Liner}} \left[\boldsymbol{\varphi}^{\text{concat}} \left(\boldsymbol{x}_{c}^{1}, \boldsymbol{x}_{c}^{2}, \dots, \boldsymbol{x}_{c}^{m} \right) \right] \right) \right]$$
 (6)

之后我们使用交叉熵函数作为分类的损失函数:

$$L_{2} = \frac{1}{N} \sum_{i} \sum_{c=1}^{M} Y_{c}^{i} \log(p_{c}^{i})$$
 (7)

其中M 是类别的数量, Y_c^i 为符号函数(0 或 1),如果样本i 的真实类别与预测类别相等则 Y_c^i 等于 1,否则等于 0, p_c^i 为模型预测样本属于c 类的概率。

因此本文的最终损失函数为:

$$L = L_1 + L_2 \tag{8}$$

4. 实验

4.1. 实验范式

本研究采用了被试依赖的实验设计,对每位参与者进行单独的训练和测试。具体而言,每项实验中,选取每个试验前 2/3 的数据(按时间顺序排列)作为训练集,余下的 1/3 作为测试集。本文通过计算分类精度和宏 F1 得分来评估模型性能。

为了探索数据融合对分类结果的影响,本文采用了逻辑回归(Logistic Regression, LR) [13]、随机森林 (Random Forest, RM) [14]、支持向量机(Support Vector Machine, SVM) [15]、K 近邻(K Neighbors, KN) [16] 和朴素贝叶斯(Naive Bayes, NB) [17]等五种分类器对融合后的数据进行分析。通过这种方法,本文旨在评估不同分类模型在处理融合数据时的效能,并探讨哪种方法在情感识别方面表现最佳。

4.2. 实验结果与分析

表 2 记录了 SEED-IV数据集和 SEED-V数据集中仅使用脑电特征的分类精度和宏 F1 分数。可以看出在两个数据集中,脑电信号的 DE 特征的分类精度比 PSD 特征更高,说明在这两个数据集中 DE 特征的可分性优于 PSD 特征。

表 3 是脑电信号和眼动信号预处理后经过各自对应的特征提取器后,直接将两种特征向量进行拼接后直接进行融合,最后输入到分类器的结果。可以看出,在加入眼动信号后,数据的分类精度有了的提升,但这种直接拼接的方法没有考虑到不同模态之间的互补信息,忽略了从一种模态到另一种模态的多个特征域的异构性质。这会导致模型难以发现不同模态之间的互补能力来区分情绪变化。

Table 2. EEG feature accuracy/macro F1 score (%) 表 2. 脑电特征精度/宏 F1 分数(%)

分类器 -	PSD 特征			DE 特征			
	1	2	3	1	2	3	
			SEED-IV				
LR	32.5/32.2	33.6/31.8	32.0/31.3	36.9/33.3	38.4/34.5	35.8/34.2	
RM	33.4/32.5	36.8/32.5	29.2/29.1	34.3/32.0	38.4/36.2	32.4/31.5	
SVM	34.5/30.2	36.8/33.0	29.3/28.4	36.9/35.4	38.4/38.1	32.9/32.3	
KN	33.7/32.5	32.5/30.8	32.7/32.1	36.9/36.1	36.3/34.3	30.6/30.8	
NB	28.5/28.5	32.6/31.9	33.0/32.4	27.9/26.3	31.6/32.0	32.4/31.5	
			SEED-V				
LR	30.2/29.8	28.3/28.5	28.1/28.6	29.2/29.3	29.7/28.8	32.0/31.9	
RM	30.2/30.4	30.3/28.9	29.4/28.3	29.2/28.8	29.7/28.0	33.0/32.1	
SVM	30.2/29.7	25.3/25.1	26.1/27.0	29.6/29.5	29.7/29.2	33.4/32.0	
KN	28.3/29.1	27.9/28.1	28.1/28.3	28.8/28.9	25.5/25.1	31.3/30.1	
NB	20.3/20.3	21.9/21.3	27.7/25.9	23.8/24.0	25.7/25.8	28.5/26.4	

Table 3. Splicing feature accuracy/macro F1 score (%) 表 3. 拼接特征精度/宏 F1 分数(%)

模型	PSD 特征			DE 特征			
	1	2	3	1	2	3	
			SEED-IV				
LR	75.1/75.2	76.5/75.8	76.9/76.6	77.2/77.3	79.2/78.0	78.9/78.5	
RM	66.2/65.9	68.7/66.7	71.4/70.3	69.4/69.1	71.9/70.4	72.8/71.8	
SVM	65.8/65.2	68.4/66.2	68.9/67.8	71.2/70.7	73.1/71.8	74.7/73.8	
KN	58.5/57.8	58.6/56.4	60.8/59.9	62.6/62.2	64.1/62.1	66.1/65.3	
NB	43.2/41.0	46.6/43.7	46.3/43.6	50.7/49.1	55.0/52.9	54.6/53.3	
			SEED-V				
LR	81.7/81.1	85.3/85.1	85.5/84.6	85.8/85.0	88.6/88.6	87.9/87.1	
RM	74.8/73.9	76.3/75.7	77.2/75.7	78.5/77.5	79.9/79.6	79.2/77.8	
SVM	73.0/71.8	76.7/76.6	78.1/76.6	75.5/74.2	79.7/79.3	79.1/77.5	
KN	65.1/64.2	67.7/67.4	69.1/68.0	72.2/71.1	71.9/71.3	74.8/73.7	
NB	51.2/49.3	55.3/54.6	56.1/54.7	47.0/44.4	55.3/53.8	52.2/49.6	

表 4 展示的是经过本文的模型之后的分类精度,可以看出,经过对比融合的脑电特征和眼动特征具有很高的可分性。这表示经过对比学习模型的特征提取器在不同模态的特征中提取到对应情感的最大表征向量,极大的提高了情感识别的准确率。

Table 4. Fusion feature accuracy/macro F1 score (%) 表 4. 融合特征精度/宏 F1 分数(%)

模型	PSD 特征			DE 特征		
	1	2	3	1	2	3
			SEED-IV			
LR	93.1/93.2	93.4/93.0	94.8/94.9	92.8/92.8	93.5/93.0	94.3/94.7
RM	96.8/96.8	96.7/.96.7	96.6/96.6	96.8/96.8	96.8/96.7	96.5/96.5
SVM	84.8/84.7	85.8/84.9	88.9/88.9	87.0/86.9	88.3/87.6	90.4/90.4
KN	77.6/77.5	77.1/76.2	83.5/88.9	78.9/78.8	79.3/78.5	84.0/84.1
NB	54.4/53.5	60.6/58.5	57.8/56.6	62.2/61.5	68.9/67.1	65.4/64.5
			SEED-V			
LR	95.5/95.5	96.4/96.5	96.3/96.1	96.4/96.3	97.5/97.6	97.5/97.2
RM	96.8/96.8	96.3/96.9	96.7/96.7	96.8/96.8	96.8/96.8	96.9/96.8
SVM	92.7/92.5	92.2/92.4	93.1/92.6	92.5/92.4	93.3/93.3	92.4/91.8
KN	85.8/85.6	85.6/85.9	86.7/86.5	89.3/89.0	89.7/89.9	89.9/89.5
NB	73.5/73.3	68.9/68.6	70.4/68.9	69.9/69.9	68.5/67.6	64.5/62.7

5. 结束语

本研究围绕使用多模态生理数据进行情绪识别展开,特别关注脑电信号和眼动信号的结合应用。通过深入分析 SEED-IV 和 SEED-V 数据集,本文提出了一种新颖的脑电和眼动信号融合方法,并基于对比学习思想进行优化。研究结果表明,融合后的多模态特征在情绪识别任务中表现出较高的准确率,尤其是脑电信号的 DE 特征在分类性能上优于 PSD 特征。这一发现不仅验证了多模态数据融合在情绪识别领域的应用价值,也为未来相关技术的发展提供了新的视角。

尽管本研究取得了显著的成果,但在未来的研究中仍有进一步的发展空间。例如,可以探索更多种类的生理信号,以及进一步优化数据处理和特征提取方法,从而提高系统的普适性和准确性。此外,对模型的泛化能力和实时处理能力进行优化,将是未来研究的重要方向。在实际应用中,如何平衡模型的复杂度和效能,以及如何更好地处理实时数据,也是需要重点关注的问题。通过这些努力,本文期望未来能够开发出更为高效、准确且实用的情绪识别系统,以服务于更广泛的应用场景。

参考文献

- [1] Maithri, M., Raghavendra, U., Gudigar, A., Samanth, J., Prabal Datta Barua, Murugappan, M., *et al.* (2022) Automated Emotion Recognition: Current Trends and Future Perspectives. *Computer Methods and Programs in Biomedicine*, **215**, Article ID: 106646. https://doi.org/10.1016/j.cmpb.2022.106646
- [2] Jafari, M., Shoeibi, A., Khodatars, M., Bagherzadeh, S., Shalbaf, A., García, D.L., et al. (2023) Emotion Recognition in EEG Signals Using Deep Learning Methods: A Review. Computers in Biology and Medicine, 165, Article ID: 107450. https://doi.org/10.1016/j.compbiomed.2023.107450
- [3] Zhang, X., Liu, J., Shen, J., Li, S., Hou, K., Hu, B., et al. (2021) Emotion Recognition from Multimodal Physiological Signals Using a Regularized Deep Fusion of Kernel Machine. *IEEE Transactions on Cybernetics*, 51, 4386-4399. https://doi.org/10.1109/tcyb.2020.2987575
- [4] Wang, Q., Wang, M., Yang, Y. and Zhang, X. (2022) Multi-Modal Emotion Recognition Using EEG and Speech Signals. *Computers in Biology and Medicine*, **149**, Article ID: 105907. https://doi.org/10.1016/j.compbiomed.2022.105907
- [5] Zhu, Q., Lu, G. and Yan, J. (2020) Valence-Arousal Model Based Emotion Recognition Using EEG, Peripheral

- Physiological Signals and Facial Expression. *Proceedings of the 4th International Conference on Machine Learning and Soft Computing*, Haiphong City, 17-19 January 2020, 81-85. https://doi.org/10.1145/3380688.3380694
- [6] Zhang, J., Yin, Z., Chen, P. and Nichele, S. (2020) Emotion Recognition Using Multi-Modal Data and Machine Learning Techniques: A Tutorial and Review. *Information Fusion*, **59**, 103-126. https://doi.org/10.1016/j.inffus.2020.01.011
- [7] Zheng, W., Liu, W., Lu, Y., Lu, B. and Cichocki, A. (2019) Emotionmeter: A Multimodal Framework for Recognizing Human Emotions. *IEEE Transactions on Cybernetics*, **49**, 1110-1122. https://doi.org/10.1109/tcyb.2018.2797176
- [8] Liu, W., Qiu, J., Zheng, W. and Lu, B. (2022) Comparing Recognition Performance and Robustness of Multimodal Deep Learning Models for Multimodal Emotion Recognition. *IEEE Transactions on Cognitive and Developmental Systems*, **14**, 715-729. https://doi.org/10.1109/tcds.2021.3071170
- [9] Liu, H., Zhang, J., Liu, Q. and Cao, J. (2022) Minimum Spanning Tree Based Graph Neural Network for Emotion Classification Using Eeg. *Neural Networks*, **145**, 308-318. https://doi.org/10.1016/j.neunet.2021.10.023
- [10] Bouazizi, S., Benmohamed, E. and Ltifi, H. (2023) Enhancing EEG-Based Emotion Recognition Using PSD-Grouped Deep Echo State Network. *JUCS—Journal of Universal Computer Science*, 29, 1116-1138. https://doi.org/10.3897/jucs.98789
- [11] Hwang, S., Hong, K., Son, G. and Byun, H. (2019) Learning CNN Features from DE Features for EEG-Based Emotion Recognition. *Pattern Analysis and Applications*, 23, 1323-1335. https://doi.org/10.1007/s10044-019-00860-w
- [12] Li, M., Yang, B., Levy, J., Stolcke, A., Rozgic, V., Matsoukas, S., et al. (2021) Contrastive Unsupervised Learning for Speech Emotion Recognition. ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, 6-11 June 2021, 6329-6333. https://doi.org/10.1109/icassp39728.2021.9413910
- [13] Pan, C., Shi, C., Mu, H., Li, J. and Gao, X. (2020) EEG-Based Emotion Recognition Using Logistic Regression with Gaussian Kernel and Laplacian Prior and Investigation of Critical Frequency Bands. *Applied Sciences*, **10**, Article 1619. https://doi.org/10.3390/app10051619
- [14] Chen, L., Su, W., Feng, Y., Wu, M., She, J. and Hirota, K. (2020) Two-Layer Fuzzy Multiple Random Forest for Speech Emotion Recognition in Human-Robot Interaction. *Information Sciences*, **509**, 150-163. https://doi.org/10.1016/j.ins.2019.09.005
- [15] Shen, P., Changjun, Z. and Chen, X. (2011) Automatic Speech Emotion Recognition Using Support Vector Machine. Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology, Harbin, 12-14 August 2011, 621-625. https://doi.org/10.1109/emeit.2011.6023178
- [16] Al Dujaili, M.J., Ebrahimi-Moghadam, A. and Fatlawi, A. (2021) Speech Emotion Recognition Based on SVM and KNN Classifications Fusion. *International Journal of Electrical and Computer Engineering (IJECE)*, 11, 1259-1264. https://doi.org/10.11591/ijece.v11i2.pp1259-1264
- [17] Oktavia, N.Y., Wibawa, A.D., Pane, E.S. and Purnomo, M.H. (2019) Human Emotion Classification Based on EEG Signals Using Naïve Bayes Method. 2019 International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, 21-22 September 2019, 319-324. https://doi.org/10.1109/isemantic.2019.8884224