

脑电情绪识别的深度学习研究综述*

李锦瑶^{1,2,3}, 杜肖兵^{1,2,3}, 朱志亮^{2,3,4}, 邓小明^{2,3}, 马翠霞^{1,2,3}, 王宏安^{1,2,3}



¹(中国科学院大学 计算机科学与技术学院, 北京 100049)

²(计算机科学国家重点实验室(中国科学院 软件研究所), 北京 100190)

³(中国科学院 软件研究所 人机交互北京市重点实验室, 北京 100190)

⁴(华东交通大学 软件学院, 江西 南昌 330013)

通信作者: 马翠霞, E-mail: cuixia@iscas.ac.cn

摘要: 情绪是情感的外在体现, 影响人类的认知、感知、理性决策等日常活动. 情绪识别作为实现计算机全面智能的一项基础任务, 在情感计算和人机交互领域被深入研究和广泛应用. 相比面部表情、语音或其他生理信号, 利用脑电进行情绪识别具有时间分辨率高、成本低、识别效果好、可靠性高的优势. 近年来, 越来越多的深度学习框架被应用于基于脑电信号的情绪识别, 并取得了比传统机器学习方法更加优异的效果. 基于深度脑电特征的情绪识别是当前的研究热点之一, 也具有一定的挑战性. 目前, 可供参考的针对此研究热点的综述文献较少. 对近年来国内外相关文献进行调研分析, 从模型输入、深度框架、实验设置、实验结果等方面对深度学习在基于脑电的情绪识别中的应用研究做了总结概况, 并在 DEAP 和 SEED 这两个公开的脑电-情绪数据集上对具有代表性的方法进行了定性和定量的多方面对比, 分析和总结这些方法的不足, 同时也对未来可能的研究方向进行了展望.

关键词: 脑电信号; 情绪识别; 深度学习; 特征提取

中图法分类号: TP181

中文引用格式: 李锦瑶, 杜肖兵, 朱志亮, 邓小明, 马翠霞, 王宏安. 脑电情绪识别的深度学习研究综述. 软件学报, 2023, 34(1): 255-276. <http://www.jos.org.cn/1000-9825/6420.htm>

英文引用格式: Li JY, Du XB, Zhu ZL, Deng XM, Ma CX, Wang HA. Deep Learning for EEG-based Emotion Recognition: A Survey. Ruan Jian Xue Bao/Journal of Software, 2023, 34(1): 255-276 (in Chinese). <http://www.jos.org.cn/1000-9825/6420.htm>

Deep Learning for EEG-based Emotion Recognition: A Survey

LI Jin-Yao^{1,2,3}, DU Xiao-Bing^{1,2,3}, ZHU Zhi-Liang^{2,3,4}, DENG Xiao-Ming^{2,3}, MA Cui-Xia^{1,2,3}, WANG Hong-An^{1,2,3}

¹(School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China)

²(State Key Laboratory of Computer Science (Institute of Software, Chinese Academy of Sciences), Beijing 100190, China)

³(Beijing Key Laboratory of Human-computer Interaction, Institute of Software, Chinese Academy of Sciences, Beijing 100190, China)

⁴(School of Software, East China Jiaotong University, Nanchang 330013, China)

Abstract: Emotion is the external expression of affect, which has an influence on cognition, perception, and decision-making of people's daily life. As one of the basic problems in the realization of overall computer intelligence, emotion recognition has been studied in depth and widely applied in fields of affective computing and human-computer interaction. Comparing with facial expression, speech and other physiological signals, using EEG to recognize emotion is attracting more attention for its higher temporal resolution, lower cost, better identification accuracy, and higher reliability. In recent years, more deep learning architectures are applied and have achieved better performance than traditional machine learning methods in this task. Deep learning for EEG-based emotion recognition is one of the research focuses and it remains many challenges to overcome. Considering that there exist few reviews literature to refer to, this study

* 基金项目: 北京市自然科学基金(4212029); 国家自然科学基金(61872346); 2019 年牛顿奖中国奖(NP2PB/100047); 中国博士后科学基金(2020M680697); 江西省青年科学基金(20202BABL212006)

收稿时间: 2020-11-06; 修改时间: 2021-03-21, 2021-05-09; 采用时间: 2021-07-12; jos 在线出版时间: 2021-10-20

CNKI 网络首发时间: 2022-11-15

investigates the implementation of deep learning in EEG-based emotion recognition. Specifically, input formulation, deep learning architecture, experimental setting and results are surveyed. Besides, articles that evaluated their model on the widely used datasets, DEAP and SEED, perform qualitative and quantitative analysis are carefully screened from different aspects and a comparison is accomplished. Finally, the total work is summarized and the prospect of future work is given.

Key words: electroencephalography (EEG); emotion recognition; deep learning; feature extraction

情绪是人在受到外界或自身刺激时产生的心理和生理反应,影响人的认知、感知、理性决策等日常活动^[1]. 识别和表达情绪是交流理解的基础,要实现自然人机交互,就要求计算机具有和人类相似的情绪感知和识别能力. 情绪识别作为实现计算机全面智能的基础任务之一,在人机交互中占据重要地位^[2],在人工智能、医疗、远程教育、军事等领域具有广阔的应用前景.

在众多情绪识别任务中,基于面部表情或语音的情绪识别虽易于分辨,却容易伪装.相比之下,基于生理信号的情绪识别不易伪装且包含更多信息,其中脑电(electroencephalography, EEG)能够反映脑神经细胞在大脑皮层或头皮表面的电生理活动,具有较高的时间分辨率.凭借 EEG 操作简单、成本低(与 PET、fMRI 等其他神经影像技术相比)、识别效果好的优势^[3],基于 EEG 的情绪识别正在得到越来越多的关注.早期识别算法以传统机器学习方法为主,包括支持向量机、K 近邻、随机森林、朴素贝叶斯等^[4-6].这类方法多采用人工提取的特征作为输入,这一操作不仅耗时耗力,而且容易造成信息损失.此外,EEG 信号在数据采集过程中易受噪声干扰,具有较低的信噪比,且信号本身在时间上具有不对称性和不稳定性.这些特性对于相对依赖人工特征提取和先验知识的传统机器学习方法来说是一种挑战.随着大数据和高性能计算技术的发展,深度学习在计算机视觉、自然语言处理、多媒体学习等领域都取得了新的突破.近年来,越来越多的深度学习框架被应用于 EEG 情绪识别.这些方法不仅具有较高的识别准确率,更是在特征提取、领域偏移、数据扩充等方面提出了新的解决方案.

将深度学习应用于基于 EEG 的情绪识别是当前的研究热点,然而,国内外可供参考的综述性文献较少.因此,本文针对基于 EEG 的情绪识别任务,研究深度学习方法在其中的应用,从模型输入、深度框架、实验设置、实验结果等方面总结不同方法,进行更加全面的比较和分析.本文第 1 节介绍相关背景知识;第 2 节从输入特征形式、样本切分方式和数据增强技术 3 个方面详述模型的输入;第 3 节对深度学习框架在 EEG 情绪识别中的应用进行分类总结,详述不同方法的技术细节;第 4 节介绍常用公共数据库、评价指标和实验策略,为进一步比较奠定基础;第 5 节基于 DEAP 和 SEED 数据库,从多个维度对相关工作进行对比,阐述定量与定性分析结果;第 6 节总结全文并从 3 个方面展望未来工作.

1 相关背景

1.1 情绪量化模型

量化情绪是对情绪进行识别的重要前提.常用的情绪量化模型主要有两类:离散模型和维度模型.在离散模型中,复杂情绪由有限的几种基本情绪组合而成.目前,大多数研究倾向于认同 6 种基本情绪^[7],即高兴(happiness)、悲伤(sadness)、惊讶(surprise)、恐惧(fear)、愤怒(anger)和厌恶(disgust).离散模型简洁直观,但能描述的情绪状态有限.维度模型能够连续地反映情绪的变化过程,可以弥补离散模型的不足.维度模型基于认知评价将情绪空间划分为 2 个、3 个或更多维度,不同情绪状态根据维度属性分布在空间中的不同位置.目前广泛应用的维度模型是效价-唤醒度(valence-arousal, VA)二维空间模型^[8],其中效价反映情绪是积极还是消极的,唤醒度反映情绪的强烈程度,如图 1 所示.此外,离散模型和维度模型之间具有一定的映射关系,高兴、悲伤和恐惧 3 种情绪在 VA 模型中的分布可见图 1.

1.2 情感计算与情绪识别

情感计算最初由美国麻省理工学院媒体实验室的 Picard 教授提出,其专著^[9]中定义“情感计算是关于、产生于、或能够对情绪施加影响的计算”.情感计算的目的在于通过赋予计算机识别、理解、表达和适应人的情绪的能力来实现计算机的全面智能,建立和谐的人机环境^[9].情感计算为人工智能发展带来了新的应用前景,让人类从不同的角度认知世界.利用不同模态的数据进行情绪识别则是情感计算的基本问题.

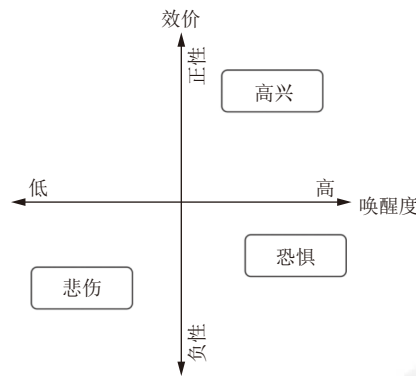


图1 情绪的VA二维空间模型示意图

早期的情绪识别方法通常依靠面部表情、语音、手势或身体姿态^[10,11],例如首先从音频中提取韵律、音质、频谱等特征,再利用模式识别方法进行分类。从这些数据中提取的特征虽易于辨别,但行为上的信号本身容易伪装,识别结果缺乏可靠性。相比之下,生理信号则不易伪装且包含更多信息。常用的生理信号包括EEG、肌电、皮电、心电、皮温、瞳距等。利用这类数据进行情绪识别,在将数据送入模型之前,通常需要进行预处理。数据预处理一般包括伪迹去除、降采样、滤波、成分分解等,具体操作视不同数据模态而有所差异。情绪识别主要包括特征提取和情绪分类两部分。特征提取有多种方式:一种是人工设计并提取特征,例如将时序信号从时域转换到频域并计算频谱特征;另一种是自动特征学习,例如将预处理后的数据直接送入循环神经网络,映射到高层语义空间学习时序特征;也可以结合上述两种方式,将人工提取的特征进一步送入模型学习更有利于分类的特征。情绪分类则根据采用的情绪量化模型对特征进行分类。以EEG信号为例,情绪识别的流程可概括为图2。

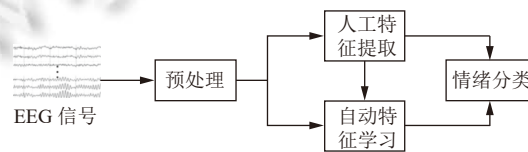


图2 典型的基于EEG的情绪识别方法流程

1.3 基于EEG的情绪识别

在众多生理信号中,EEG是记录头皮电位变化的信号,能够较好地反映人的精神状态和情绪变化。各种研究表明EEG和情绪之间存在一定关联^[12,13]。随着EEG采集设备和相关信号处理方法的迅速发展,基于EEG的情绪识别已成为目前比较常用的方法,并得到越来越多的应用^[14,15]。市面上常见的EEG采集设备包括Neurosky、Emotiv、Neuroelectrics和Biosemi,不同设备支持的时间和空间分辨率不同。时间分辨率和采样频率相关,如512 Hz。空间分辨率与电极数量相关,可以按照10-20、10-10或10-5定位标准在颅骨上放置电极。总体上,EEG信号具有较高的时间分辨率和相对较低的空间分辨率。Palus^[16]指出EEG具有时序上的不对称性和不稳定性,这两种非线性特性增加了数据处理和情绪识别的难度。不对称性指不同认知活动激活的脑叶不同,激活程度也存在差异。不稳定性指EEG信号的均值、方差和协方差等属性会随时间部分或完全地变化。

最早的基于EEG的情绪识别方法由Musha等人^[17]提出。他们定义了情绪矩阵,将EEG特征转换为四元情绪向量,对应4种基本情绪,实现了最大时间分辨率为0.64 s的实时情绪分析。自此,越来越多的机器学习和模式识别方法都被应用到此领域。随着深度学习方法在各类识别任务中取得巨大成功,多种深度学习框架(至少含有两个隐层的神经网络)开始被应用于EEG情绪识别。相比传统机器学习方法,深度学习贴合人脑的学习机制,对问题有更强的描述和拟合能力。此外,深度学习框架的特性能够克服传统EEG情绪识别方法存在的一些不足,并解决任务本身存在的一些挑战。其一,与图像、音频和文本等数据相比,EEG信号的收集相对复杂,导致当前EEG-情绪

数据库中可供使用的数据量较小. 对此, 可以采用生成对抗网络 (generative adversarial networks, GAN) 对数据进行扩充, 从而获得更多数据用于模型的训练^[18]. 其二, 情绪是一种复杂的反应, 涉及时间、空间、环境、语言、文化、种族等因素^[19], 从 EEG 数据中人工提取情绪相关特征不仅耗时耗力, 也会损失原始信号蕴含的信息. 借助卷积神经网络 (convolutional neural networks, CNN) 和循环神经网络 (recurrent neural networks, RNN) 强大的特征检测和提取能力, 可以将原始 EEG 数据作为模型输入, 充分利用深度模型的学习能力^[20,21]. 其三, EEG 信号本身具有不稳定性, 不同被试之间的 EEG 数据存在较大差异, 给不依赖于被试的识别任务和模型投入实际应用带来阻碍. 针对训练集与测试集间存在的领域偏移 (domain shift) 问题, 可以引入领域判别器 (domain discriminator), 帮助提取更具领域不变性的特征, 从而缩减源领域与目标领域之间的分布差距^[22]. 其四, 神经科学研究表明, 人的情绪与各种大脑皮层区域密切相关^[23]. EEG 作为一种多通道信号, 利用空间信息探索其不同电极之间的关联以及不同脑区对情绪识别的贡献有重要意义. 对此, 可以设计邻接矩阵表示电极拓扑关系, 利用图卷积神经网络 (graph convolutional neural networks, GCNN) 学习节点之间的本质关联^[24]. 也可以根据电极的空间位置对它们进行分组, 并分别送入特征提取模块然后进行特征融合, 学习脑区内部和脑区之间的关联性^[25].

2 模型输入

EEG 通过测量离子流在大脑神经元内部产生的电压波动, 按照时间顺序来记录头皮活动. 因此, EEG 设备采集的是一种多通道时间序列, 其中每个通道对应一个 EEG 电极. 鉴于 EEG 信号具有较高的时间分辨率, 蕴含丰富的信息, 如何处理采集到的数据作为模型的输入成为关键问题之一. 针对此问题, 本节从输入特征形式、样本切分方式、数据增强技术 3 个方面进行分析.

2.1 输入特征形式

EEG 情绪识别模型的输入形式可以大致分为 4 类: 人工设计的传统特征, 原始或经过预处理的信号, 编码后的图像特征, 以及其他特征.

大多数研究利用人工设计的 EEG 特征作为模型输入. 这些特征大致包括 4 类, 时域 (time domain)、频域 (frequency domain)、时频域 (time-frequency domain) 和空间域 (space domain)^[26], 如图 3 所示. 时域特征最直观易得, 主要包括事件相关电位、信号统计量、Hjorth 参数特征、分形维数等. 频域特征以离散傅里叶变换为基础计算得到, 典型的特征有功率谱密度 (power spectral density, PSD)、微分熵 (differential entropy, DE) 等. 时频域特征通常借助滑动窗口对信号进行分段, 并对每段子信号进行时频域信号变换, 从而结合时域和频域的优势, 提高对不稳定信号的处理能力. 具体地, 通常将原始 EEG 信号分解为 5 个子频段——delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (13–30 Hz), gamma (31–45 Hz), 并在不同频段和电极上分别提取特征. 最终, 对每个频段生成一个的输入矩阵, 其中 t 代表时间步长, n 代表电极数量, d 代表特征维度. 以脑区的不对称性能够帮助情绪识别为出发点, 空间域特征近年来逐渐受到关注, 其主要分为空频域特征和电极组合特征. 较常用的特征包括不对称差 (differential asymmetry, DASM)、不对称商 (rational asymmetry, RASM) 和微分 caudality (differential caudality, DCAU). Zheng 等人^[4]在 5 个频段下, 选择不同数量的电极, 分别计算 PSD、DE、DASM、RASM、DCAU 特征, 并送入支持向量机或深度神经网络进行学习分类, 探索更利于识别的电极和频段. 随后, 许多研究^[27–29]也利用这五种特征作为输入. 受到上述研究中 DE 特征普遍表现较好的启发, 有一定数量的研究^[18,22,30–36]选用 DE 特征作为深度模型的输入.

少数研究直接将预处理后 (或原始) 的 EEG 信号送入模型进行训练. 一方面实现端到端的情绪识别, 节约人工特征提取的成本; 另一方面尽可能保留原始信息, 给深度学习框架留出更多发挥空间. Alhagry 等人^[37]将分段后的源数据送入 LSTM 进行特征学习, 实现端到端的情绪识别. 除了直接利用原始的排序方式, 也可以按特定规则重新组织数据. Wen 等人^[38]利用皮尔森相关系数度量电极对之间的相关性, 并按照最大化全局信息和最大化局部信息两种算法重新排列电极顺序. Wang 等人^[21]根据 EEG 的电极分布图, 将电极序列映射到表示空间拓扑结构的二维平面, 在时间上叠加后形成一种三维输入结构, 有利于在一种输入结构上综合考虑时间和空间特征. 为了适应

Inception 模型的输入形式, Cimtay 等人^[39]对分段后的原始数据进行重塑 (reshape), 引入噪声将数据从二维扩充到三维. Yang 等人^[20]同时利用预处理后的时间序列和二维拓扑结构作为模型输入, 分别学习特征向量, 并进行特征层融合从而实现分类.

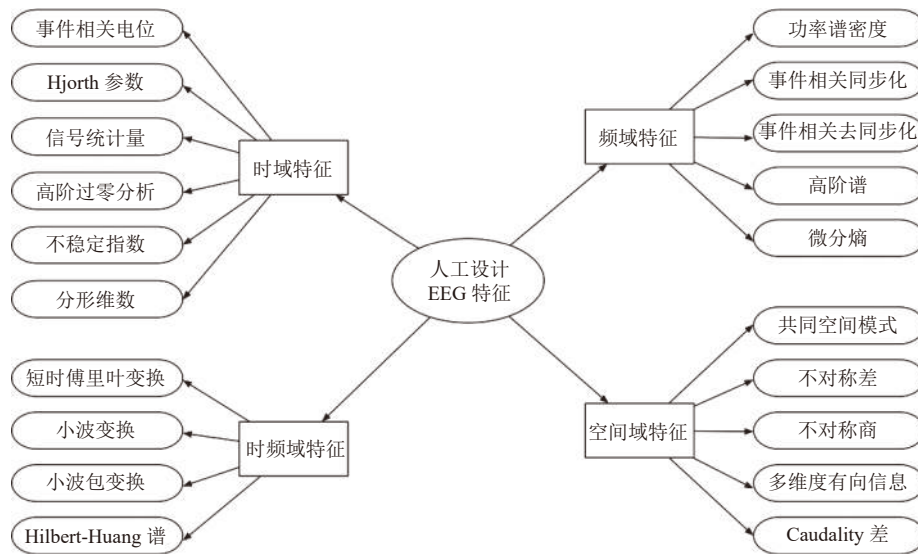


图3 人工设计的 EEG 特征的分类及各类别下的常用特征

此外, 有研究将 EEG 编码为图像作为模型输入. Li 等人^[40]对原始数据提取 PSD 特征后, 将电极序列按拓扑结构转化为二维方阵, 再根据方阵元素值的大小生成热图. Li 等人^[41]将经过连续小波变换的信号进一步转化为量值图 (scalogram), 获得网格状图像. 通过快速傅里叶变换 (short-time Fourier transform, STFT), Donmez 等人^[42]将 EEG 信号转化为频谱图 (spectrogram) 作为网络输入, Wang 等人^[43]将这一图像输入表述为电极频率分布图 (electrode-frequency distribution maps, EFDm), 该图能够同时提供频域和位置信息. Wang 等人^[44]分别度量 6 组对称电极经过 STFT 变换后的图像的差值, 并加权融合 6 组差值, 获得单张输入图像.

还有研究采用其他方式对 EEG 进行编码作为模型输入. 例如, 为了探索不同脑区对于情绪识别的贡献, Li 等人^[25]将 62 个电极划分入 16 类, 每类对应一个脑部区域, 从而探索不同脑区与情绪的关联. 在 16 组 DE 特征输入矩阵的基础上, 相应深度模型实现时间和空间, 局部和全局的特征学习. 考虑到电极之间的位置关系不能代表功能关系, Song 等^[24]将 EEG 特征用图结构表示, 训练 GCNN 动态更新邻接矩阵, 学习节点之间的本质关联.

2.2 样本切分方式

EEG 数据库包含的样本量通常较小, 然而深度学习模型一般需要大量数据进行训练. 在这种矛盾下, 如何切分数据成为一个值得考虑的问题. 在此, 样本切分指的是对一个试次 (trial) 的实验数据 (一个被试观看一个视频片段的数据), 利用多大的滑动窗口进行切分, 作为一个样本. 一方面, 我们希望获得尽可能多的样本用于训练, 避免模型过拟合, 提高泛化能力. 这就意味着要使用尽可能小的滑动窗对被试观看一个视频片段的实验数据进行切分, 在片段层次 (segment-level) 进行情绪识别. 另一方面, 情绪是一种变化的过程. 在实验采集中, 我们无法保证被试在每个时刻都被激发了相应的情绪, 那么按照小窗口切分出的某些片段可能并不包含期望的信息, 这就要求滑动窗口尽可能大. 在不同的研究中, 使用的滑动窗口大小从 1 s^[4,24,45], 2 s^[24], 3 s^[46], 4 s^[21,47], 5 s^[37], 6 s^[48], 9 s^[22,49]到 60 s^[29], 62 s^[33]不等. Candra 等人^[50]探索了不同窗口大小对 EEG 情绪分类的影响, 发现更加有效的窗口大小为 3–12 s. 此外, 在某些应用场景下, 更加关注情绪的演变过程和长期情绪监控, 此时利用切分后的片段训练模型则不能满足需求, 需要进行试次层次 (trial-level) 的情绪识别^[40,41,51], 即将一个被试观看一个视频片段的数据作为一个样本, 也可以视为滑动窗口大小等于单个试次长度.

2.3 数据增强技术

鉴于 EEG 的采集步骤相对繁琐, 现有数据库的数据量一般较小. 在这种情况下, 即使采用 1 s 的窗口大小切分样本用于训练, 依旧容易产生过拟合现象. 对此, 一些研究利用数据增强技术, 对样本进行了扩充, 试图提高模型的泛化能力.

在深度学习框架中, 生成对抗网络 (generative adversarial networks, GAN)^[52]常被用于数据增强. 其包含两个模块, 生成器 (generator) 和判别器 (discriminator). 生成器的目标是生成看似真实、与原始数据相似的实例. 判别器的任务则是尽可能区分实例来源于数据库还是生成器. 两种模型在互相博弈的过程中得到提升, 达到均衡. Luo 等人^[18]利用条件 Wasserstein GAN (conditional Wasserstein GAN, CWGAN) 生成 EEG 的 DE 特征, 并选用 3 种指标衡量生成数据的质量. 在 SVM 分类器下, 利用增强后数据进行训练, 获得的分类精度较原始数据库有明显提升. 此外, 一些工作采用了图像处理领域常用的数据增强策略. Li 等人^[41]通过向处理后的网格状图像中加入 5 dB 的高斯白噪声, 将数据扩充了 250 倍, 帮助提高模型的泛化能力. Li 等人^[40]利用椒盐噪声, 对 EEG 编码后的图像样本集扩充了 20 倍. Wang 等人^[33]通过向 DE 特征加入高斯噪声获得更多训练样本, 并比较了不同增强倍数对结果的影响.

3 深度学习框架分类概述

近年来, 受到深度学习在各类识别任务中取得巨大成功的影响, 如图 4 所示, 多种深度学习框架也被应用于 EEG 情绪识别. 将深度学习框架限定为至少含有两个隐层的神经网络, 本节对相关工作涉及的方法进行简单介绍, 并总结它们在具体任务中的实现细节.

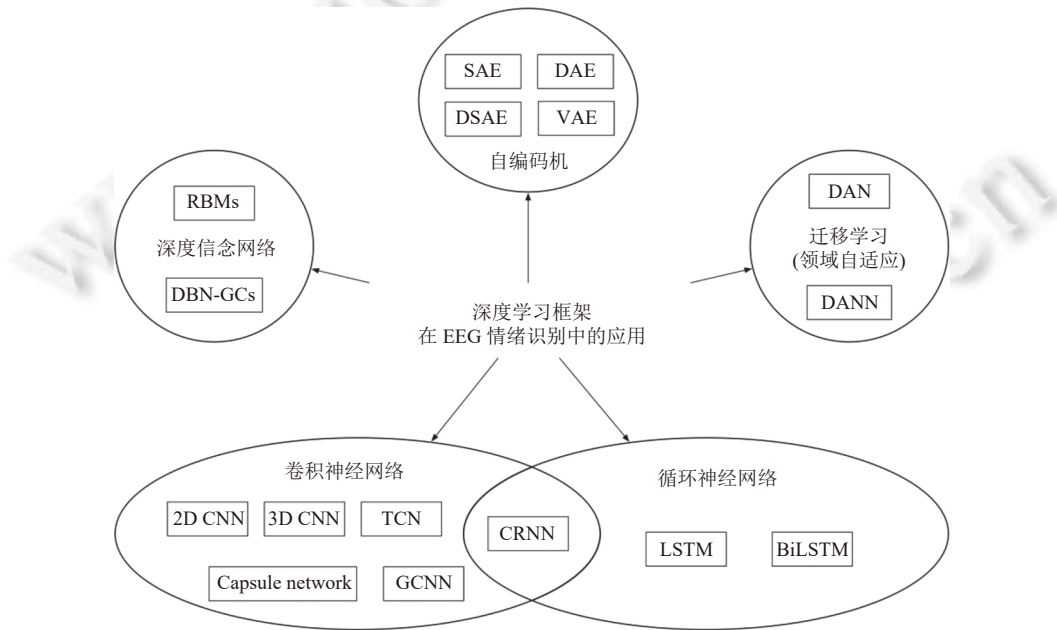


图 4 用于 EEG 情绪识别的深度学习框架

3.1 自编码器与深度信念网络

自编码器 (autoencoder, AE)^[53]是一种无监督学习神经网络, 含 3 层神经元, 如图 5 所示. 由输入层到隐层是编码过程, 由隐层到输出层是解码重构过程. 训练目标是使得重构误差尽可能小, 从而在隐层学习更好的表征. Chai 等人^[54]提出一种结合 AE 和子空间对齐的框架, 提取具有领域不变性的特征, 解决 EEG 数据存在的不稳定性问题. 该框架可以嫁接到任何监督性学习方法上进行情绪分类. 堆栈自编码器 (stacked autoencoder, SAE) 提取一个

训练好的 AE 的隐层, 作为另一个 AE 的输入, 并训练这个新的 AE, 如此堆叠多个 AE, 实现多层级的特征学习. Jirayucharoensak 等人^[45]和 Yang 等人^[34]分别采取 2 或 3 个 AE 堆叠实现 SAE, 并用 Softmax 对最终的 EEG 特征进行分类. 另一种 AE 的变体——降噪自编码器 (denoising autoencoder, DAE)^[55]——通过向输入层加入噪声干扰, 训练网络能够对含噪数据进行编码、解码, 从而提高隐层特征的表达能力. Xu 等人^[47]利用堆栈 DAE (stacked DAE, SDAE) 进行特征学习并利用 Softmax 分类. 此外, Li 等人^[56]利用变分 AE (variational AE, VAE) 无监督地学习 EEG 信号的低维表示.

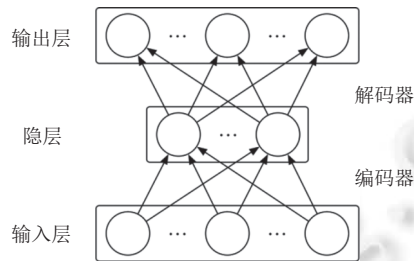


图5 自编码器框架

深度信念网络 (deep belief networks, DBN)^[57]是一种概率生成模型. 它由若干层受限玻尔兹曼机 (restricted Boltzmann machines, RBM) 组成, 其深度结构能够有效学习高层特征表示. 一个 RBM 只有两层神经元: 可视层和隐层, 分别用于输入数据和特征检测, 两层之间双向连接, 层内神经元之间无连接. 将多个 RBM 堆叠则构成 DBN, 其中底层 RBM 的输出作为高层 RBM 的输入. DBN 既可用于无监督学习, 也可用于有监督学习. 如图 6 所示, 通常采用含有两个隐层的 DBN 进行 EEG 情绪识别^[4,47], 其训练主要包括 3 个步骤: 首先, 对每层进行无监督的预训练; 其次, 将 RBM 展开为编码器和解码器, 利用反向传播算法对各层进行无监督的微调, 学习使得输入和重构向量尽可能相似的网络权重和偏置; 最后, 在顶端 RBM 的可视层加入表示类别的神经元, 反向传播更新参数, 实现有监督的微调. 然而 DBN 的隐层缺少约束, 使得其很难挖掘 EEG 电极间和频段间的相关信息, Chao 等人^[58]引入胶质细胞链 (glia chain), 利用 DBN-GCs 网络学习隐层神经元的关联. 胶质细胞能够调整隐层神经元的激活概率, 并向相邻胶质细胞传递信号. 此外, 文章利用集成学习, 将特征子集分别送入 5 个 DBN-GCs, 拼接各模块的隐层输出, 作为判别 RBM 的输入.

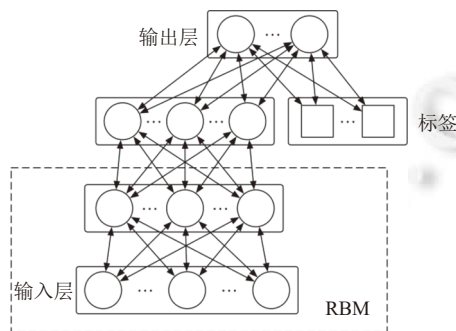


图6 深度信念网络框架

3.2 卷积神经网络与循环神经网络

CNN^[59]是一种前馈神经网络, 利用反向传播算法进行训练, 其中一个典型层包含 3 级: 卷积、非线性变换、池化. 卷积运算具有局部连接、权值共享、等变表示的特性. 通常在同一卷积层内使用多个不同的卷积核以学习不同特征. 非线性变换, 也称为激活函数, 能够给模型加入非线性因素, 提高表达能力. 目前大部分 CNN 采用线性整流 (rectified linear unit, ReLU) 函数^[60]作为激活函数. 池化是一种过滤细节的方法, 使用某一位置周围的总体特

征代替网络在该位置的输出. 常用的池化函数包括最大池化、平均池化等. 现有研究设计或应用不同的 CNN 结构进行 EEG 情绪识别. Wang 等人^[33]将 LeNet^[61]和 ResNet^[62]这两种典型网络架构分别用于情绪识别. Cimtay 等人^[39]在预训练的 InceptionResNetV2 模型的基础上增加 5 个全连接层, 并利用中值滤波处理网络的输出. Moon 等^[46]设计了 3 种不同的 CNN 框架(分别包含 2 层、5 层和 10 层)用于 EEG 情绪识别, 卷积核大小均为 3×3, 附加全连接层输出分类结果. 实验结果显示 CNN 能够取得比 SVM 更好的分类精度, 但 CNN 结构过于复杂时会产生过拟合现象. Tripathi 等人^[48]采用的 CNN 框架含有 2 个卷积层, 1 个最大池化层和 1 个全连接层, 实验结果表明选择能够持续激发神经元的激活函数尤为重要. Tripathi 等认为最佳选择为在第一个卷积层后, 为代价分类模型选择 tanh 函数, 为唤醒度分类模型选择 ReLU 函数; 在第 2 个卷积层后均采用 tanh 函数. 类似上述几种 CNN 的框架在其他研究^[38,42,44,51]中也被应用. 鉴于常规的 2D CNN 仅能提取空间信息, Wang 等人^[21]利用 3D CNN 同时捕获时间和空间特征, 并配合批标准化(batch normalization)和密集预测(dense prediction)解决数据分布偏移和参考标签(ground truth)不可靠问题. 相关消融实验证实了上述模块的有效性. 近年来涌现出许多对 CNN 进行结构创新的框架, 其中时间卷积网络(temporal convolutional networks, TCN)^[63]利用因果卷积、空洞卷积和残差结构, 兼具时域建模能力和少量参数下的特征提取能力, 在多种任务上都取得了优于 RNN 的表现, Yang 等人^[30]将 TCN 应用到 EEG 情绪识别. 考虑到传统 CNN 无法表示对象之间的空间层级关系, 且最大池化容易造成信息损失, 胶囊网络(capsule network, CapsNet)和囊间动态路由算法^[64]被提出. CapsNet 能够表示局部和全局之间的相对位置关系, 其核心单元是胶囊(Capsule). 随着 CapsNet 在各个领域的广泛应用, Liu 等人^[65]也利用多层特征引导的 CapsNet 实现端到端的 EEG 情绪识别. 图卷积网络(graph convolutional neural networks, GCNN)将 CNN 与图谱理论结合, 适用于处理非欧式结构数据(non-euclidean structure data). Song 等人^[24]用图表示 EEG 数据, 借助 GCNN 探索不同电极在情绪识别任务上的关联, 考虑到传统 GCNN 预先定义图的邻接矩阵不利于学习电极在功能上的联系, 文章提出以自适应的方式动态更新邻接矩阵. Wang 等人^[35]将 GCNN 与宽度学习系统(broad learning system, BLS)结合得到 BDGLS(broad dynamical graph learning system), 可以随机地将学习的图空间生成为可能的低维空间, 然后再扩展到一个宽度随机空间, 其增强节点能够帮助搜索更适合的情绪分类特征. Wang 等人^[29]向 GCNN 引入 EEG 的锁相值(phase-locking value, PLV)连接性, 以同时提取时空特性和功能连接上的固有信息.

RNN 是包含反馈连接的神经网络, 适用于序列建模, 利用随时间反向传播算法进行训练. 在输入为较长的时间序列时容易出现梯度消失或爆炸问题. 长短期记忆网络(long short term memory network, LSTM)^[66]则是一种改进模型, 可以解决传统 RNN 存在的长距离依赖问题. LSTM 定义了细胞状态变量, 核心思想是让信息选择性通过, 不断去除或增加信息到细胞状态. 这一目标通过输入门、输出门和遗忘门这 3 种门控单元实现. 顾名思义, 输入门确定存放什么新信息到细胞状态中, 忘记门决定需要从细胞状态中忘掉的信息, 输出门确定输出的信息. LSTM 在 EEG 情绪识别中常被作为特征提取器. Alhagry 等人^[37]利用包含双层 LSTM 的架构实现端到端的情绪识别. Li 等人^[22]利用 LSTM 对大脑半球数据学习时序信息并输出高层特征. 除了利用 LSTM 学习时序特征以外, Du 等人^[67]利用 LSTM 提取不同 EEG 电极之间的依赖关系, 并引入注意力机制对不同电极的权重进行学习. 在 RNN 的变体中, 双向 RNN 由前向 RNN 和后向 RNN 组成, 能够更好地捕捉双向的语义依赖. Zhang 等人^[49]提出时空 RNN(spatial-temporal RNN, STRNN), 包含空间和时间两层 RNN 结构. 在空间 RNN 中, 打破 RNN 的输入为时间序列的传统, 用 4 种预先定义的空间顺序遍历数据并分别送入 4 个 RNN, 提取空间特征. 4 个 RNN 形成互补关系, 能够尽可能消除噪声干扰. 在时间 RNN 中, 拼接空间 RNN 的输出作为输入, 利用双向 RNN 提取时序特征. 每层 RNN 结构后都附加线性投影避免稀疏化, 最终利用 Softmax 进行分类. Li 等人^[25]提出局部到全局的时空神经网络(regional to global spatial-temporal neural network, R2G-STNN), 将数据以多种组织方式分别送入双向 LSTM 模块, 提取时间和空间, 局部和全局的特征.

为了综合利用 CNN 和 RNN 的优势, 出现了一批两者的混合模型. Yang 等人^[20]将数据并行地送入 CNN 和 LSTM, 对提取的特征进行融合后利用 Softmax 分类, 如图 7(a) 所示. Li 等人^[40]和 Li 等人^[41]将数据依次送入 CNN 和 LSTM, 先后学习电极之间的关联和时序上的信息, 如图 7(b) 所示.

3.3 迁移学习中的领域自适应方法

迁移学习^[68]旨在将从某个领域学习到的知识或模式应用到其他不同但相关的领域中, 从而摆脱传统机器学

习需要大量标注数据,且数据需要满足独立同分布假设的限制.这一产生背景较好得吻合了 EEG 数据存在的问题,即数据量较小和领域偏移 (domain shift) 显著.领域偏移指训练集和测试集数据之间的分布差异. EEG 数据不稳定性较强且不同被试间存在个体差异,使得基于 EEG 的情绪识别中领域偏移问题显著,当已训练的模型被应用于新被试的数据或新环境下采集的数据时,模型识别能力下降. Wang 等人^[43]借助迁移学习的思想,进行跨数据集的 EEG 情绪识别.具体方法是将某个数据集上训练的模型迁移到另一个数据集,进行微调后验证模型的表现,从而解决数据集样本较少和 EEG 数据不稳定性较强的问题.研究利用了两种微调策略,一种是对网络的各层参数都进行更新,另一种是只对最后的全连接层进行调整,实验结果表明后一种策略的效果更佳.领域自适应是迁移学习中的代表性方法,它通过尽可能缩减两个领域的分布差异,提取具有领域不变性的特征用于识别.领域对抗神经网络 (domain-adversarial neural network, DANN^[69]) 通常包含 3 个模块,特征提取器、标签分类器和领域判别器 (domain discriminator, DD). 如图 8 所示, X_S 和 X_T 分别表示源领域 (source domain, 即训练集) 和目标领域 (target domain, 即测试集) 的输入数据, H_S 和 H_T 是相应的高层特征. DD 的作用是区分数据来自源领域还是目标领域. DANN 的训练目标是使 DD 的区分能力尽可能弱.其对抗性体现在最小化分类器误差的同时,最大化 DD 误差.具体可以通过在前馈模型中加入梯度反转层 (gradient reverse layer, GRL) 实现对抗学习. Jin 等人^[32]将 DANN 应用于基于 EEG 的情绪识别.考虑到左、右大脑半球在结构和功能上都存在不对称性, Li 等人^[22]引入 2 个局部 DD 分别缩减左、右大脑半球内部训练集和测试集数据的分布差异,此外设置 1 个全局 DD 约束整体数据库的一致性. Li 等人^[25]也在网络中利用对抗训练缓解领域偏移问题.基于提取特征的非迁移性会随网络深度增加而增强的研究发现, Long 等人^[70]提出深度自适应网络 (deep adaptation network, DAN), 将深度网络的全连接层设置为自适应层,引入多核最大均值差异 (multi-kernel maximum mean discrepancies, MK-MMD) 方法度量领域差异,提高训练模型在测试集上的迁移性. Li 等人^[31]利用 3 个普通全连接层, 2 个自适应全连接层和 1 个输出层,实现 DAN 在 EEG 情绪识别上的应用. Luo 等人^[36]将领域自适应算法与 Wasserstein GAN 结合,提出 WGANDA (Wasserstein generative adversarial network domain adaptation) 模型,在预训练阶段将源领域和目标领域映射到一个共同的空间,在对抗训练阶段在上述空间缩小两个领域之间的差距.

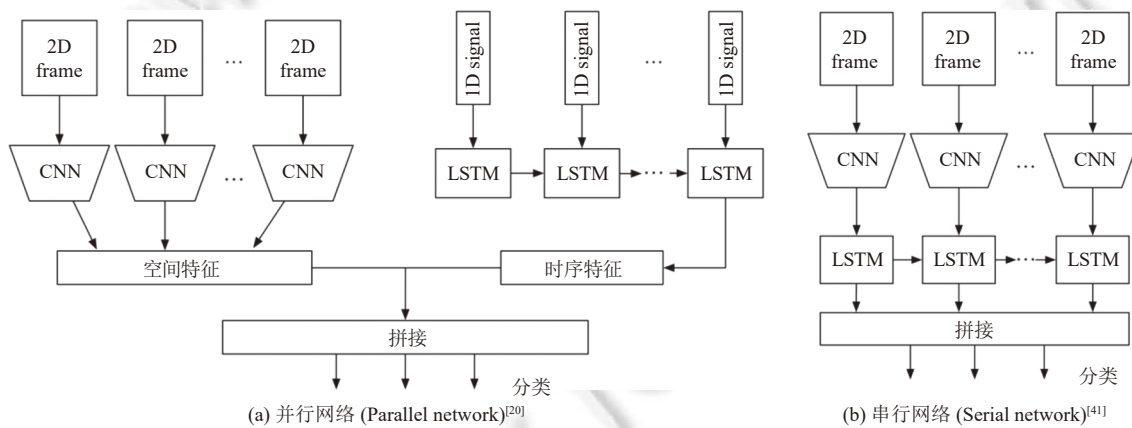


图 7 两种 CRNN 的网络框架

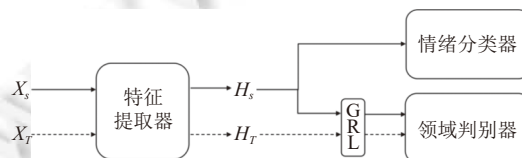


图 8 领域对抗神经网络框架

4 实验设置

外部刺激和内部响应是主要的情绪诱发方式. 其中, 通过图片、音乐和视频等外部刺激诱发不同的情绪是最普遍使用的方法. 而视频结合视觉和听觉体验, 更接近现实生活中情绪的产生环境, 在诱发具体情绪方面较为常用. 例如, Philippot^[71]和 Gross 等人^[72]分别筛选了诱发 6 种或 8 种离散情绪的电影片段, 并均以主观报告的方式证明了电影片段的诱发效果. 本文也将重点放在视频诱发的 EEG-情绪数据库. 本节首先介绍常用的公共数据库, 再对情绪识别评价指标和实验策略进行梳理.

4.1 常用公开数据库

DEAP 数据库^[73]由伦敦玛丽女王大学 (Queen Mary University of London) 发布, 初衷是用于构建音乐视频 (music video, MV) 推荐系统, 因此采用 MV 诱发情绪. 在实验中, 32 名被试 (50% 女性, 平均年龄 26.9) 分别被要求观看 40 段时长为 1 分钟的 MV. 用 Biosemi ActiveTwo 系统采集 32 导 EEG 信号和 8 导外周生理信号, 采样率为 512 Hz. 每观看完一段视频, 被试根据 Russell 的二维情绪空间模型对 MV 的效价 (valence)、唤醒度 (arousal)、喜爱度 (liking) 和支配度 (dominance) 等分别进行自我评估, 分数从 1 至 9. 基于此, 可以设置多种识别任务. 一种是对效价和唤醒度等分别进行二分类, 具体地, 大多数研究利用一个阈值, 判定分数大于阈值为高效价/唤醒度, 分数小于等于阈值为低效价/唤醒度, 阈值通常为 5. 考虑到不同被试的阈值可能存在一定差异, 少数研究^[30]利用 k-means 聚类方法, 对每个被试的 40 次试验进行聚类, 以获得两类标签. 在这种情况下, 阈值也在 5 上下浮动. 另一种识别任务是结合效价和唤醒度进行四分类, 即在二维空间中, 在两个坐标轴上均以阈值为界限将空间划分为 4 块, 每块代表一个类别, 则 4 类分别是: 高效价高唤醒度 (HVHA)、高效价低唤醒度 (HVLA)、低效价高唤醒度 (LVHA)、低效价低唤醒度 (LVLA).

SEED 数据库^[4]由上海交通大学发布, 考虑到被试的文化背景, 用于诱发情绪的 15 个视频片段均选自华语电影, 每个片段 4 分钟左右. 数据库涉及 3 种情绪——正性、中性、负性效价, 每种情绪对应 5 个片段. 采用 ESI NeuroScan 系统采集 15 名被试 (7 名男性, 8 名女性, 平均年龄 23.3) 的 62 导 EEG 信号, 采样率为 1000 Hz. 此外, 每个被试都进行了 3 个时段 (session) 的实验, 每个时段都观看内容相同 (顺序不同) 的 15 个视频片段, 相邻时段间隔两周, 即共 45 个试次. SEED 数据库的这一设置为验证模型在时间上的稳定性提供可能.

DEAP 和 SEED 是应用最广泛的两个 EEG-情绪数据库, 下文也以这两个数据库为主, 进行实验策略汇总与实验结果对比分析. 类似的数据库还有 MAHNOB-HCI^[74]、DREAMER^[75]、CAS-THU^[14]等, 在此不展开介绍. 上述公共数据库的关键信息对比如表 1 所示.

表 1 常用 EEG-情绪数据库对比

数据库	被试数量	视频片段数量	电极数量	情绪标签
DEAP ^[73]	32	40	32	唤醒度, 效价, 喜爱度, 支配度
SEED ^[4]	15	15	62	正性, 中性, 负性
MAHNOB-HCI ^[74]	27	20	32	唤醒度, 效价, 支配度
DREAMER ^[75]	23	18	14	9种离散情绪
CAS-THU ^[14]	30	16	14	8种离散情绪

4.2 评价指标

对于情绪识别这一分类任务, 研究者普遍采用准确率 (accuracy, ACC) 和标准差 (standard deviation, STD) 作为首要的评价指标. 对于二分类任务, ACC 的计算如公式 (1) 所示, 其中 TP、TN、FP、FN 分别代表真实为正预测为正、真实为负预测为负、真实为正预测为正、真实为负预测为负的样本数量.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

此外, 精准度 (precision)、召回率 (recall)、F1 分数 (F1 score) 和混淆矩阵 (confusion matrix) 也可以用于对情

绪识别模型的表现进行评价. $F1$ 分数被定义为精准率和召回率的调和平均数, 能够兼顾精准率和召回率这两种指标.

$$recall = \frac{TP}{TP + FN} \quad (2)$$

$$precision = \frac{TP}{TP + FP} \quad (3)$$

$$F1 \text{ score} = \frac{2 \times recall \times precision}{recall + precision} \quad (4)$$

混淆矩阵是一个 $N \times N$ 的矩阵, N 为分类类别数. 每一行代表预测值, 每一列代表真实值. 对角线上的值越大, 对应的模型表现更佳.

4.3 实验策略

为了验证模型在不同方面的表现, 满足不同评估标准的需求, 多种实验策略相继被提出. 不同策略在应用场景、识别难度上均存在差异. 对于不同的深度学习框架, 本文基于相同的实验策略对它们的表现进行比较, 以取得更加有意义的对比与分析. 表 2 对 DEAP 与 SEED 数据库上常见的实验策略进行了汇总. 其中, 以字母 d 为首的策略代表 DEAP 数据库上的实验策略, 相应的, SEED 数据库的实验策略以字母 s 为首.

表 2 基于 DEAP 与 SEED 数据库的实验策略

编号	数据库	实验策略	分类标准
d1	DEAP	被试依赖实验, 训练集:测试集=24:16	(1) 唤醒度, 效价, 喜爱度, 支配度分别二分类(高、低)或三分类(高、中、低) (2) 四分类(高效价高唤醒度、高效价低唤醒度、低效价高唤醒度、低效价低唤醒度)
d2		被试依赖实验, 4折交叉验证	
d3		被试依赖实验, 5折交叉验证	
d4		被试依赖实验, 10折交叉验证	
d5		片段依赖实验, 试次留一交叉验证	
d6		被试独立实验, 5折交叉验证	
d7		被试独立实验, 被试留一交叉验证	
s1	SEED	被试依赖实验, 训练集:测试集=9:6	三分类(正性、中性、负性)
s2		被试独立实验, 被试留一交叉验证	
s3		被试一对一(subject-to-subject)实验, 每次选取一个被试的数据作为训练集, 另一个被试的数据作为测试集	
s4		被试独立实验, 5折交叉验证	
s5		时段一对一(session-to-session)实验, 对每个被试选取一个时段的数据作为训练集, 另一个时段的数据作为测试集	
s6		被试独立实验, 训练集:测试集=9:6	
s7		被试依赖实验, 片段留一交叉验证	
c1	SEED & DEAP	训练集: SEED, 测试集: DEAP	同 DEAP 的分类准则
c2		训练集: DEAP的一部分和SEED, 测试集: DEAP	

最常见的是被试依赖 (subject-dependent) 策略, 指的是利用同一个被试的 EEG 数据进行训练和测试, 具体可以采用比例法、交叉验证法 (包含留一法) 划分训练集与测试集. 例如, 表 2 中的策略 d1 表示在 DEAP 数据库上, 对同一被试观看 40 个视频片段的数据, 选取其中 24 个片段对应的样本作为训练集, 剩余 16 个片段对应的样本作为测试集. 对每个被试分别建模, 然后取所有被试的识别准确率的平均作为最终评价标准. 由于 EEG 数据具有被试之间、时间上的不稳定性, 这种被试依赖的策略可以一定程度上避免数据分布偏移问题, 易于获得较高的识别精度. 相似的还有片段依赖 (clip-dependent) 策略, 即训练和测试集来源于不同被试观看同一个视频片段的数据. 例

如,表 2 的策略 d5,表示对于同一个视频片段对应的 40 个被试的 EEG 记录,每次选取 1 个被试的数据作为测试集,剩余 39 个被试的数据作为训练集,如此循环 40 次进行交叉验证,最终取 40×32 次实验的准确率的平均值作为最终结果。

相比以上两种策略,被试独立 (subject-independent) 和跨时段 (cross-session) 策略更具挑战性,但也更符合实际应用场景。在被试独立策略中,训练集和测试集的数据来源于不同的被试,可用于验证模型对不同个体的泛化能力。最常用的是留一交叉验证法,即每轮选取一个被试的所有数据用于测试,剩余所有被试的数据用于训练,循环被试数量轮,例如表 2 的策略 d7 和 s2。鉴于 SEED 数据库包含 3 个时段的实验数据的特殊性,研究者们设计了跨时段的实验策略,验证模型在时间上的稳定性,包括对 3 个时段进行留一交叉验证,或一对一交叉验证(也属于被试依赖的实验策略)。例如,表 2 的策略 s7 代表在 SEED 数据库下,对于同一个被试的 3 个时段的 45 次试验,每次选取同一个时段的 15 次试验对应的样本作为测试集,剩余 30 次试验的数据作为训练集,循环 3 轮,对每个被试分别建模,因此最终对 3×15 次识别结果取平均。

随着迁移学习的兴起,跨数据集的实验策略也相继被提出。与分别在 DEAP 和 SEED 上验证模型的研究相比,跨数据集的研究明显数量较少。因此我们只收集了两种策略。c1 代表利用 SEED 的数据进行训练,并在 DEAP 数据集上进行测试。c2 在 c1 的基础上,拿出 DEAP 的少量数据用于微调,从而提高模型在测试阶段的识别准确率。

5 方法对比与分析

为了进一步对深度学习框架进行定性和定量分析,本文选定 DEAP 和 SEED 这两个常用数据库,对基于它们开展实验的代表性工作整理。本文对文献的筛选标准包括:(1) 文章研究基于 EEG 的情绪识别任务;(2) 文章只利用 EEG 这一种模态的数据;(3) 文章发表时间介于 2014–2020 年之间;(4) 文章使用的方法包含深度学习框架(至少含有两个隐层的神经网络);(5) 文章在 DEAP 或 SEED 数据库上对方法进行验证。表 3 从模型输入、任务层次(分段层次或试验层次)、深度学习框架、实验策略、实验结果这 5 个方面对不同工作进行了总结概括(其中带有*标记的 SVM 方法作为传统机器学习方法代表而出现,为深度学习方法提供参考),我们可以从中观察并统计得到如下结论。

在输入特征形式上,如图 9(a),大多数深度学习方法还是以人工设计的特征作为输入,占比 64.7%,其中涉及 DE 特征的占 82.6%,可见 DE 是多数研究的首选特征;此外,分别有 17.6% 和 14.7% 的研究使用预处理后的原始信号或图像作为输入,这两类输入在近两年逐渐兴起。

表 3 从多个方面总结基于 DEAP 和 SEED 数据库的文献

参考文献	模型输入	任务层次	深度模型	实验策略	实验结果 ACC (%) / STD (%)
[4]	DE, PSD, DASM, RASM, DCAU 特征	片段层次	SVM* DBN	s1 s1	83.99/09.72 86.08/08.34
[24]	DE, PSD, DASM, RASM, DCAU 特征	片段层次	Dynamic graph convolutional neural network (DGCNN)	s1 s2	90.40/08.49 79.95/09.02
[43]	电极-频率分布映射	片段层次	Residual block-based CNN and transfer learning	s1 c2	90.59 三分类 82.84/10.74
[76]	DE 特征	片段层次	Channel-fused dense convolutional network	s1	90.63/04.34
[77]	DE 特征	片段层次	Bi-hemispheric discrepancy model (BiHDM)	s1 s2	93.12/06.06 85.40/07.53
[22]	DE 特征	片段层次	Bi-directional domain discriminative neural network (BiDANN)	s1 s2	92.38/07.04 83.28/09.60

表3 从多个方面总结基于 DEAP 和 SEED 数据库的文献 (续)

参考文献	模型输入	任务层次	深度模型	实验策略	实验结果 ACC (%) / STD (%)
[67]	DE特征	试次层次	Attention-based LSTM with domain discriminator (ATDD-LSTM)	s1	91.08/05.34
				s2	90.92/01.05
				s7	79.26/12.79
				d5	二分类 效价: 90.91/12.95 唤醒度: 90.87/11.32
				d7	二分类 效价: 69.06/06.37 唤醒度: 72.97/06.57
[25]	16个区域的DE特征矩阵	片段层次	R2G-STNN	s1	93.38/05.96
				s2	84.16/07.63
[18]	DE特征	片段层次	Conditional Wasserstein GAN (CWGAN) and linear SVM	s1	86.96/12.72
				d1	效价: 73.89/10.82 唤醒度: 78.17/09.58
[49]	DE特征	片段层次	STNN	s1	89.50
[35]	DE特征	片段层次	BDGLS	s1	93.66/06.11
[54]	DE特征	片段层次	Sub-space alignment autoencoder (SAAE)	s2	77.88
				s3	62.50
				s5	81.81
[32]	DE特征	片段层次	DANN	s2	79.19/13.14
[31]	DE特征	片段层次	DAN	s2	83.81/08.56
[39]	重组源数据	segment level	InceptionResnetV2 and median filter	s2	78.34/06.11
				d7	二分类 效价: 72.81/05.07
				c1	二分类 效价: 58.10/09.51
[36]	DE特征	片段层次	WGANDA	s2	87.07/07.14
				d7	二分类 效价: 67.99/06.56 唤醒度: 66.85/05.52
[56]	未知	未知	VAE and LSTM	s2	84.29
				d7	二分类 效价: 71.76 唤醒度: 72.43
[34]	DE特征	片段层次	SAE	s4	85.5
[27]	DE, PSD, DASM, RASM, DCAU特征	片段层次	Graph regularized extreme learning machine (GELM)	s4	91.07/07.54
				s5	79.28
				d6	69.67
[33]	数据扩充的DE特征	片段层次	LeNet ResNet	s6	74.3
				s6	75.0
[28]	DE, PSD, DASM, RASM, DCAU特征	片段层次	Hierarchical network structure with subnetwork nodes	s6	93.26
				s6	85.71
				s7	80.84
[21]	经过预处理的源数据3D张量	片段层次	3D CNN	d2	二分类 效价: 72.1 唤醒度: 73.1
[37]	预处理的源数据	片段层次	LSTM	d2	二分类 效价: 85.45 唤醒度: 85.65 喜爱度: 87.99

表 3 从多个方面总结基于 DEAP 和 SEED 数据库的文献 (续)

参考文献	模型输入	任务层次	深度模型	实验策略	实验结果 ACC (%) / STD (%)
[41]	通过小波和量值变换的 网格状帧	试次层次	CNN and LSTM	d3	二分类 效价: 72.06 唤醒度: 74.12
[30]	DE 特征	片段层次	TCN	d3	二分类 效价: 74.4 唤醒度: 71.4
[40]	数据扩充的 EEG 多维特征图像 序列	试次层次	CNN and LSTM	d3	4 classes 75.21
[47]	PSD 特征	片段层次	SDAE DBN	d3	二分类 效价: 84.77 唤醒度: 85.86 喜爱度: 87.11 二分类 效价: 88.59 唤醒度: 88.33 喜爱度: 89.20
[38]	重组的预处理源数据	片段层次	CNN	d3	二分类 效价: 77.98 唤醒度: 72.98
[20]	预处理源数据的 2D 帧序列和电 极序列	片段层次	Parallel convolutional recurrent neural network (CRNN)	d4	二分类 效价: 90.80/03.08 唤醒度: 91.03/02.99
[58]	时域, 频域和时频域的 4 种特征	试次层次	DBN-GC-based ensemble deep learning	d4	二分类 效价: 76.83 唤醒度: 75.92
[44]	融合的 EEG 图	片段层次	CNN	d4	二分类 效价: 82.88
[65]	预处理的源数据	片段层次	Multi-level features guided capsule network	d4	二分类 效价: 97.97 唤醒度: 98.31 支配度: 98.32
[48]	时域统计特征	片段层次	DNN CNN	d5	二分类 效价: 75.78 唤醒度: 73.125 3 classes 效价: 58.44 唤醒度: 55.70 二分类 效价: 81.406 唤醒度: 73.36 3 classes 效价: 66.79 唤醒度: 57.58
[46]	大脑连通特征	片段层次	CNN	d6	二分类 效价: 99.72
[45]	PSD 特征和所有对称电极对的 功率谱差异	片段层次	SAE and covariate shift adaptation of principal components	d7	三分类 效价: 53.42 唤醒度: 52.03

关于样本的切分方式, 如图 9(b), 鉴于深度模型训练需要尽可能多的样本, 88.6% 的研究将一次试验的数据切分为多个样本, 进行片段层次的识别任务, 其中滑动窗口大小为 1 s 的占 48.6%; 只有 11.4% 的研究尝试试验层次

的情绪识别, 2016、2017、2018 和 2020 年各一篇. 基于此, 我们认为实验层次的情绪识别是一个潜在的研究方向, 研究者们可以在此投入更多精力.

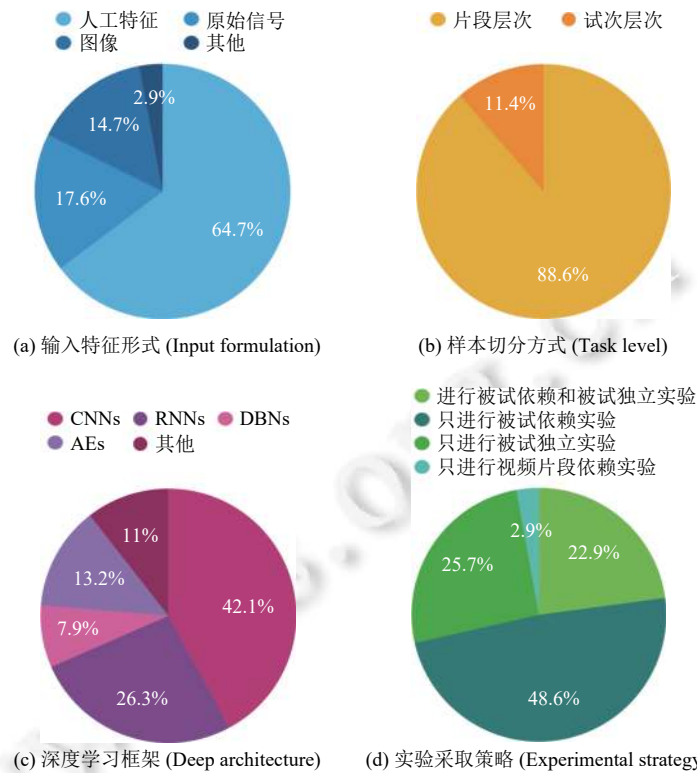


图 9 4 种衡量指标下不同方法的占比

从深度学习框架类型的角度, 如图 9(c), 排名前 3 的分别是 CNN 及其变体 (42.1%), RNN 及其变体 (26.3%), AE 及其变体 (13.2%). 由此可见, 从 EEG 中提取空间和时间上的特征仍是研究热点, 无监督的特征提取也受到越来越多的关注. 更加多样化的框架有待被应用.

在实验策略方面, 如图 9(d), 所有工作中, 涉及被试依赖策略的研究占 71.4%, 涉及被试独立策略的研究占 48.6%. 单篇工作只涉及被试依赖策略的占 48.6%, 只涉及被试独立策略的占 25.7%, 只涉及片段依赖策略的占 2.9%. 可见更多研究倾向于只在一类策略上开展实验, 其中选择被试依赖策略的居多, 但这样并不能全面地验证模型的表现. 意识到这一点, 有 22.9% 的研究同时应用被试依赖和被试独立的实验策略, 对模型进行更加系统的评估. 此外, 有 5.7% 的研究涉及跨数据集的实验策略.

本文进一步针对 3 种常用的实验策略 (s1, s2, d3), 对比不同模型基于此策略的实验结果, 并作出更具体的分析.

● SEED 数据库下的被试依赖实验 (s1 策略)

实验策略为对于同一被试的所有数据, 选取前 9 个视频片段对应的数据作为训练集, 剩余 6 个视频片段对应的数据作为测试集. 对 15 个被试分别建模, 对 15 次识别准确率取平均 (部分研究对 3 个时段分别实验取平均) 作为最终准确率. 图 10 展示了 8 种深度学习框架在 s1 策略下的表现. 除了 DBN 和 DGCNN 采用 DE、PSD、DASM、RASM、DCAU 这 5 种特征分别作为输入, 其他 6 种方法都只用 DE 特征作为输入, 但 8 种方法都在采用 DE 特征时获得最高准确率. 此外, DBN、CWGAN、DGCNN、BDGLS 采用大小为 1 s 的滑动窗口切分样本, STNN、BiDANN、R2G-STNN 采用 9 s 的滑动窗口. 理论上, 9 s 的样本包含更多信息, 应取得更佳的识别准确率, 但 1 s 的滑动窗口能够且分出更多样本. 关于哪一种滑动窗口大小更佳, 我们并未从实验结果中找到明显的规律. 值得注意的是, ATDD-LSTM 进行的是试验层次的实验, 相比其他切分方式, 在一定程度上识别难度更大. 识别准确率高于 90% 的 5 个模型可以被分为两类, 基于 LSTM 和 DD 的模型, 以及基于 GCNN 的模型. 其一, R2G-STNN、

BiDANN 和 ATDD-LSTM 利用 BiLSTM 或 LSTM 提取特征, 引入 DD 帮助提取更具领域不变性的特征. 此外, 三者都同时关注局部和全局信息: ATDD-LSTM 运用注意力机制模拟不同电极对于各种情绪类别的贡献权重, BiDANN 考虑了左右半脑的差异性, R2G-STNN 将电极划分到 16 个不同的脑区分别提取特征并整合. R2G-STNN 更是在此基础上融合了时间和空间信息, 可见, 综合考虑多方面特征能够帮助情绪识别. 其二, DGCNN 和 BDGLS 都采用 GCNN 对 EEG 的拓扑结构与功能联系进行建模, 可见, 学习电极之间的本质关联十分重要. 值得一提的是, CWGAN 虽然对数据进行了扩充, 但其精度是在以线性 SVM 为分类器的情况下取得的, 我们推测 SVM 可能未能充分利用扩充后的数据, 从而限制了最终结果.

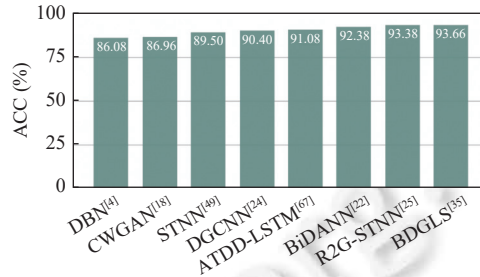


图 10 8 种深度学习框架在实验策略 s1 下的识别准确率

● SEED 数据库下的被试独立实验 (s2 策略)

在实验中, 对同一时段的数据, 每轮选 1 个被试的所有数据作为测试集, 剩余 14 个被试的所有数据作为训练集, 循环 15 次 (部分研究循环 3 个时段的 45 次) 进行交叉验证, 并取平均识别准确率作为最终结果. 图 11 对比了 8 种深度学习框架在 s2 策略下的识别准确率. 与实验策略 s1 相比, 识别精度普遍下降, 这是由于被试独立实验的训练数据与测试数据之间存在分布差异, 使得此任务更具挑战性. 因此, 更多的领域自适应方法被应用于此实验设置, 包括引入 DD, 应用 DAN, 采用子空间对齐等. 其中, R2G-STNN 与其他同样采用 DD 缩减领域偏移的方法相比, 在特征提取器的设计上更加复杂, 同时兼顾时间和空间, 局部和全局的特征学习. WGAN 包含预训练和对抗训练 2 个阶段, 将训练集和测试集映射到一个共同空间下, 并缩小它们之间的分布差异, WGAN 保证了模型的稳定性和收敛能力. 准确率最高的 ATDD-LSTM 在利用 DD 约束特征提取的基础上, 引入注意力机制学习不同电极对不同情感类别的贡献程度, 此外, 其试验层次的样本切分方式可能包含了更多信息, 提高了特征的代表性.

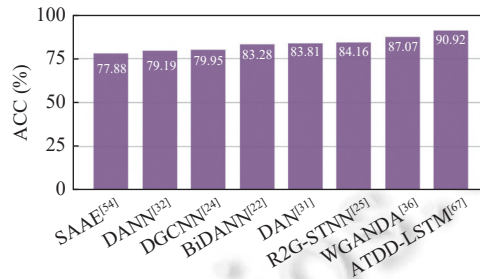


图 11 8 种深度学习框架在实验策略 s2 下的识别准确率

● DEAP 数据库下的被试依赖实验 (d3 策略)

在 d3 实验策略下, 一个被试的 40 次试验数据被平分为 5 份, 每次取一份作为测试集, 剩余 4 份作为训练集, 交叉验证, 取 5×32 个模型的平均准确率作为最终结果. 图 12 对比了对效价和唤醒度分别进行二分类的 5 种深度学习方法的表现. 与 SEED 数据库上的实验结果相比, DEAP 数据库上的识别准确率普遍较低. 在 CRNN 中, 研究者将 EEG 编码为图像, 并选择了试验层次的样本切分方法, 虽然进行了数据扩充, 但仍未取得较高的精度. 用于训练 TCN 的监督信息是经过 k-means 聚类的二分类标签, 这种方法考虑了不同被试之间的阈值差异, 但从识别结果上观察, 是否考虑此差异的影响并不大. 同样是采用了 CNN 进行特征提取, CNN^[38]的输入是根据皮尔森相关系数重新排列电极顺序的预处理后的原始信号, 其 CNN 结构简单, 精度却相比结构相对复杂的 TCN^[30]与 CRNN^[41]有

些许提升,可见输入形式对结果有一定影响.排名较高的 SDAE 和 DBN 出自同一篇工作,它们只利用了每次 60 s 试验中后 40 s 的 EEG 信号,并以 4 s 的滑动窗口切分数据获得 PSD 特征作为输入,这可能是促成它们较高准确率的原因之一.

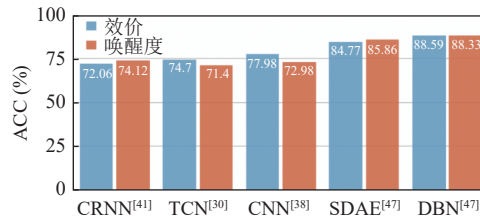


图 12 5 种深度学习框架在实验策略 d3 下的识别准确率

由上述分析可见,在不同实验策略下占据主导地位的深度学习框架不尽相同.针对不同深度模型,本文总结其应用在 EEG 情绪识别任务的优劣如下.

- 自编码器与深度信念网络

自编码器与深度信念网络的结构相对简单,复杂度较低,能够从一定程度上提取出与情绪分类相关的 EEG 特征.但是单纯利用自编码器或深度信念网络的模型的特征提取能力有限,拟合效果欠佳,导致其分类准确率无法进一步提高,因此常被用作基本的对比模型.但编码器的编码器与解码器思想能够指导更复杂的模型设计,间接辅助 EEG 情绪识别性能提升.

- 卷积神经网络与循环神经网络

相比卷积操作,以双向 LSTM 为代表的 RNN 结构在 EEG 情绪识别上被更加频繁地用于特征提取,本文分析这得益于 EEG 本身是一种时序信号,而 LSTM 能够较好地捕获 EEG 样本在不同采样点上的时间依赖关系.此外,通过改变 LSTM 的输入可以实现对不同脑区的 EEG 信号的表征学习,从而间接引入空间信息.对于 CNN 而言,它能够较好地提取空间信息,但是简单地将脑电帽电极分布映射到二维空间,进而重排列 EEG 信号作为网络输入只考虑了电极间的位置关系,忽略了不同脑区间的功能联系.应用 GCNN 能够打破这一瓶颈,以图结构为媒介学习电极之间的本质关联.相应地,预先处理得到时域特征用于邻接矩阵的计算,可以一定程度上引入时间信息.

- 领域自适应方法

领域偏移问题是伴随 EEG 信号的固有挑战,缩减训练集与测试集之间的数据分布差异对于提高情绪识别模型的泛化性能至关重要.应用领域自适应方法是目前的主流方法,最常见的操作是在网络中加入领域判别模块和梯度逆转层,简便且能够保留网络的端到端属性,对相应特征进行可视化能够证明此方法的有效性.由此形成了一种包括特征提取器、领域判别器和情绪分类器 3 个模块的典型 EEG 情绪识别网络框架.

6 总结与展望

情绪识别作为人机交互的重要组成部分,具有重要研究意义,并已在人工智能、医疗、远程教育、军事等领域发挥作用.结合 EEG 数据和深度学习方法的情绪识别研究,近年来得到越来越多的关注.本文从模型输入、深度框架、实验设置、实验结果等方面,对深度学习在基于脑电的情绪识别中的应用进行了总结.尤其针对基于 DEAP 和 SEED 开展实验验证的代表性工作进行了概括(如表 3 所示),并进行了定性和定量分析.

图 13 以时间线的形式,对近 7 年内深度框架在 EEG 情绪识别中的应用脉络进行了整理.从研究数量上看,从 2014 年到 2020 年,越来越多的深度学习方法被应用于 EEG 情绪识别,说明此研究正在得到越来越多的关注.从方法组成上看,从最开始的 DBN, SAE 到后来的 CNN、RNN 等,模型的多样性正在增加,许多在其他领域被提出的新模型能够很快地被应用于此任务,更多领域自适应方法被用于解决数据偏移问题.此外,模型的深度也不断增加,与深度学习的发展趋势相一致.

尽管已经取得了一定成果,当前研究仍面临很多挑战.对此,本文从数据库建立、研究任务、特征设计和算法优化 4 个方面,进一步探究和梳理未来的可行研究思路.

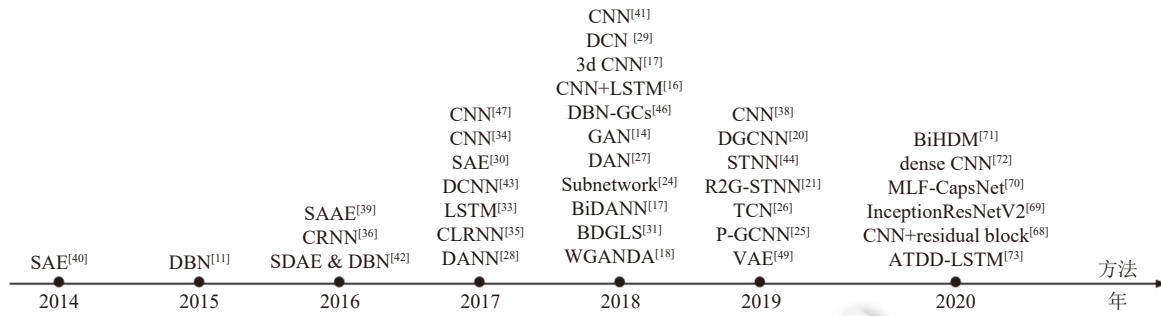


图 13 深度学习框架在 EEG 情绪识别中应用的时间线

● 在数据库建立方面,利用自然交互技术建立更加自然的情绪诱发材料库,服务于更加高质量的数据采集;建立规模更大、情绪粒度更细致的 EEG-情绪数据库,满足深度学习对于训练数据的大量需求,支撑更多实际应用。

传统的情绪刺激材料包括音乐、图片和视频等,存在刺激强度不足、沉浸感差的问题。由此诱发的被试情绪可能无法与刺激材料所传达的情绪类别保持一致,导致数据质量较低,从而在模型验证阶段影响性能。随着自然交互技术的发展,增强现实、虚拟现实和混合现实逐渐进入公众视野。依托上述技术,未来可以围绕如何构建沉浸感更强、用户友好度更高的情绪诱发材料库,更加有效、精准地诱发情绪展开讨论。此外,由于 EEG 数据采集较为繁琐,目前的 EEG-情绪数据库规模相对较小。以 DEAP 数据库为例,其收集了 32 个被试各自观看 40 段 MV 的数据,且每段 MV 的时长为 1 min。如果进行试次层次的情绪识别,仅能获得 1 280 个样本。如果进行片段层次的情绪识别且以 1 s 为滑动窗口对样本进行切分,也仅能获得 76 800 个样本。这些量级对于需要大量样本进行训练的深度学习方法来说显得有些吝啬。因此,有必要建立更大规模的 EEG-情绪数据库,为理论研究与算法验证提供坚实基础。另外,目前大多数 EEG-情绪数据库基于 VA 二维空间模型或单纯以正性、中性和负性作为情绪类别,探讨并建立包含更加细粒度情绪的 EEG-情绪数据库更有实际意义。例如:对于辅助型人工智能系统,如果能够精准识别用户的细粒度、生活化情绪类别将推动自然性人机交互的实现。

● 在研究任务方面,开展更加细粒度、更具实用意义的情绪识别与分析,实现更加精准的情感计算智能。

目前的脑电情绪识别大多数进行多类别单标签预测,且是基于维度模型的二分类问题。这简化了人类情绪是混合且复杂的实际情况。为了适应实际应用,有必要将识别任务扩展到基于离散情绪模型的多标签学习领域。另一方面,在某些情景下,比起预测用户在情绪认知过程中体验到的离散情绪类别,我们更关心用户在情绪认知过程中对不同情绪体验的强度。此外,利用离散情绪模型表示的部分情绪类别之间的高度相关性给离散情绪识别带来较大挑战。针对以上问题,可以引入多标签情绪分布学习技术进行解决。

● 在特征设计方面,引入认知心理学和神经科学领域的理论机制,从多学科交叉角度来设计与情绪最相关的脑电特征。

鉴于 EEG 信号本身信噪比较低,提取人工设计的特征仍然是组成 EEG 情绪识别的重要部分和影响模型表现的重要因素。目前的特征多来源于数字信号处理领域,代表性特征包括 DE、PSD 等,以这些特征为输入模型也已取得了一定成果。但这些特征缺乏直觉,EEG 信号蕴含的丰富信息还未被充分挖掘。一种新的思路是融入认知心理学和神经科学领域的研究发现,将大脑与情绪的相互作用作为先验知识,设计与情绪本身具有较高关联性的特征。例如,人类情绪认知相关研究表明,在情绪表达过程中大脑空间临近区和功能相关区之间存在联系和通路,所以在基于 EEG 的情绪识别过程中引入情感机制和原理提取空间相邻区域和功能相关区域的特征能够提高模型性能。

● 在算法优化方面,设计多模态融合和多策略联合的深度情绪识别算法,或从 EEG 特性出发设计针对性网络结构,实现多个数据库和多种实验策略上识别的高精度度和高稳定性。

目前的研究大多针对单个数据库或单个实验策略验证模型的表现,尤其是在被试依赖的实验策略上。这种做法易于取得较高的识别准确率,但在实际应用中,预先收集个体的大量 EEG 数据且为每个个体建立识别模型显然是不可取的。因此,能够兼顾被试独立和被试依赖策略,并且在不同数据库上均有较优表现的模型有待提出。在模

型优化上,首先,从多模态信息融合角度出发可以将 EEG 与其他模态数据结合实现多模态情绪识别,其他模态可以是外周生理信号、视频、音频等。模态融合可以实现优势互补,提高模型的鲁棒性,解决 EEG 易受噪声干扰,数据丢失等问题。其次,联合多种策略提升分类器性能是一种不错的选择。据本文分析,具备多维度特征提取能力的模型往往能取得更佳的识别效果。EEG 信号是一种多通道时序信号,蕴含丰富的时域、频域和空间域信息,实现不同维度特征的有效融合是情绪识别的关键之一。可行思路包括:1)对同样的输入数据,利用不同的子网络分别提取不同类型特征;2)将输入数据划分为不同部分,利用相同的网络框架分别提取多个局部特征表达。最后,专门设计针对 EEG 特性的深度学习框架。目前的网络结构大多从计算机视觉和自然语言处理领域借鉴。未来网络设计可以从 EEG 信号的高分辨率、时序不对称性和不稳定性或其他先验知识出发。例如将 EEG 电极位置按照脑区功能划分为 5 个区域:额叶 (frontal lobe),颞叶 (temporal lobe),顶叶 (parietal lobe),枕叶 (occipital lobe) 和中央沟 (central sulcus),探索脑区之间的显著双向联系对情绪识别精度和模型效率的影响。依据额叶不对称性与情绪唤醒和效价相关的研究发现,利用共同注意力机制捕捉成对脑区之间的相互作用。

References:

- [1] Panksepp J. *Affective Neuroscience: The Foundations of Human and Animal Emotions*. New York: Oxford University Press, 2004.
- [2] Cowie R, Douglas-Cowie E, Tsapatsoulis N, Votsis G, Kollias S, Fellenz W, Taylor JG. Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*, 2001, 18(1): 32–80. [doi: [10.1109/79.911197](https://doi.org/10.1109/79.911197)]
- [3] Alarcão SM, Fonseca MJ. Emotions recognition using EEG signals: A survey. *IEEE Trans. on Affective Computing*, 2019, 10(3): 374–393. [doi: [10.1109/TAFFC.2017.2714671](https://doi.org/10.1109/TAFFC.2017.2714671)]
- [4] Zheng WL, Lu BL. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Trans. on Autonomous Mental Development*, 2015, 7(3): 162–175. [doi: [10.1109/TAMD.2015.2431497](https://doi.org/10.1109/TAMD.2015.2431497)]
- [5] Huang D, Guan CT, Ang KK, Zhang HH, Pan YZ. Asymmetric spatial pattern for EEG-based emotion detection. In: *Proc. of the 2012 Int'l Joint Conf. on Neural Networks (IJCNN)*. Brisbane: IEEE, 2012. 1–7. [doi: [10.1109/IJCNN.2012.6252390](https://doi.org/10.1109/IJCNN.2012.6252390)]
- [6] Piho L, Tjahjadi T. A mutual information based adaptive windowing of informative EEG for emotion recognition. *IEEE Trans. on Affective Computing*, 2020, 11(4): 722–735. [doi: [10.1109/TAFFC.2018.2840973](https://doi.org/10.1109/TAFFC.2018.2840973)]
- [7] Van den Broek EL. Ubiquitous emotion-aware computing. *Personal and Ubiquitous Computing*, 2013, 17(1): 53–67. [doi: [10.1007/s00779-011-0479-9](https://doi.org/10.1007/s00779-011-0479-9)]
- [8] Russell JA. A circumplex model of affect. *Journal of Personality and Social Psychology*, 1980, 39(6): 1161–1178. [doi: [10.1037/h0077714](https://doi.org/10.1037/h0077714)]
- [9] Picard RW. *Affective Computing*. Cambridge: MIT Press, 2000.
- [10] Kleinsmith A, Bianchi-Berthouze N. Affective body expression perception and recognition: A survey. *IEEE Trans. on Affective Computing*, 2013, 4(1): 15–33. [doi: [10.1109/T-AFFC.2012.16](https://doi.org/10.1109/T-AFFC.2012.16)]
- [11] Adolphs R. Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, 2002, 1(1): 21–62. [doi: [10.1177/1534582302001001003](https://doi.org/10.1177/1534582302001001003)]
- [12] Mathersul D, Williams LM, Hopkinson PJ, Kemp AH. Investigating models of affect: Relationships among EEG alpha asymmetry, depression, and anxiety. *Emotion*, 2008, 8(4): 560–572. [doi: [10.1037/a0012811](https://doi.org/10.1037/a0012811)]
- [13] Sammler D, Grigutsch M, Fritz T, Koelsch S. Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology*, 2007, 44(2): 293–304. [doi: [10.1111/j.1469-8986.2007.00497.x](https://doi.org/10.1111/j.1469-8986.2007.00497.x)]
- [14] Liu YJ, Yu MJ, Zhao GZ, Song JJ, Ge Y, Shi YC. Real-time movie-induced discrete emotion recognition from EEG signals. *IEEE Trans. on Affective Computing*, 2018, 9(4): 550–562. [doi: [10.1109/TAFFC.2017.2660485](https://doi.org/10.1109/TAFFC.2017.2660485)]
- [15] Huang YJ, Wu CY, Wong AMK, Lin BS. Novel active comb-shaped dry electrode for EEG measurement in hairy site. *IEEE Trans. on Biomedical Engineering*, 2015, 62(1): 256–263. [doi: [10.1109/TBME.2014.2347318](https://doi.org/10.1109/TBME.2014.2347318)]
- [16] Paluš M. Nonlinearity in normal human EEG: Cycles, temporal asymmetry, nonstationarity and randomness, not chaos. *Biological Cybernetics*, 1996, 75(5): 389–396. [doi: [10.1007/s004220050304](https://doi.org/10.1007/s004220050304)]
- [17] Musha T, Terasaki Y, Haque HA, Ivamitsky GA. Feature extraction from EEGs associated with emotions. *Artificial Life and Robotics*, 1997, 1(1): 15–19. [doi: [10.1007/BF02471106](https://doi.org/10.1007/BF02471106)]
- [18] Luo Y, Lu BL. EEG data augmentation for emotion recognition using a conditional Wasserstein GAN. In: *Proc. of the 40th Annual Int'l Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Honolulu: IEEE, 2018. 2535–2538. [doi: [10.1109/EMBC.2018.8512865](https://doi.org/10.1109/EMBC.2018.8512865)]

- [19] Kim J, André E. Emotion recognition based on physiological changes in music listening. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2008, 30(12): 2067–2083. [doi: [10.1109/TPAMI.2008.26](https://doi.org/10.1109/TPAMI.2008.26)]
- [20] Yang YL, Wu QF, Qiu M, Wang YD, Chen XW. Emotion recognition from multi-channel EEG through parallel convolutional recurrent neural network. In: *Proc. of the 2018 Int'l Joint Conf. on Neural Networks (IJCNN)*. Rio de Janeiro: IEEE, 2018. 1–7. [doi: [10.1109/IJCNN.2018.8489331](https://doi.org/10.1109/IJCNN.2018.8489331)]
- [21] Wang Y, Huang ZY, McCane B, Neo P. EmotioNet: A 3-D convolutional neural network for EEG-based emotion recognition. In: *Proc. of 2018 Int'l Joint Conf. on Neural Networks (IJCNN)*. Rio de Janeiro: IEEE, 2018. 1–7. [doi: [10.1109/IJCNN.2018.8489715](https://doi.org/10.1109/IJCNN.2018.8489715)]
- [22] Li Y, Zheng WN, Cui Z, Zhang T, Zong Y. A novel neural network model based on cerebral hemispheric asymmetry for EEG emotion recognition. In: *Proc. of the 27th Int'l Joint Conf. on Artificial Intelligence*. Stockholm: ACM, 2018. 1561–1567.
- [23] Lotfi E, Akbarzadeh-T MR. Practical emotional neural networks. *Neural Networks*, 2014, 59: 61–72. [doi: [10.1016/j.neunet.2014.06.012](https://doi.org/10.1016/j.neunet.2014.06.012)]
- [24] Song TF, Zheng WM, Song P, Cui Z. EEG emotion recognition using dynamical graph convolutional neural networks. *IEEE Trans. on Affective Computing*, 2020, 11(3): 532–541. [doi: [10.1109/TAFFC.2018.2817622](https://doi.org/10.1109/TAFFC.2018.2817622)]
- [25] Li Y, Zheng WM, Wang L, Zong Y, Cui Z. From regional to global brain: A novel hierarchical spatial-temporal neural network model for EEG emotion recognition. *IEEE Trans. on Affective Computing*, 2019. [doi: [10.1109/TAFFC.2019.2922912](https://doi.org/10.1109/TAFFC.2019.2922912)]
- [26] Zhang GH, Yu MJ, Chen G, Han YH, Zhang D, Zhao GZ, Liu YJ. A review of EEG features for emotion recognition. *Scientia Sinica Informationis*, 2019, 49(9): 1097–1118 (in Chinese with English abstract). [doi: [10.1360/N112018-00337](https://doi.org/10.1360/N112018-00337)]
- [27] Zheng WL, Zhu JY, Lu BL. Identifying stable patterns over time for emotion recognition from EEG. *IEEE Trans. on Affective Computing*, 2019, 10(3): 417–429. [doi: [10.1109/TAFFC.2017.2712143](https://doi.org/10.1109/TAFFC.2017.2712143)]
- [28] Yang YM, Wu QMJ, Zheng WL, Lu BL. EEG-based emotion recognition using hierarchical network with subnetwork nodes. *IEEE Trans. on Cognitive and Developmental Systems*, 2018, 10(2): 408–419. [doi: [10.1109/TCDS.2017.2685338](https://doi.org/10.1109/TCDS.2017.2685338)]
- [29] Wang ZM, Tong Y, Heng X. Phase-locking value based graph convolutional neural networks for emotion recognition. *IEEE Access*, 2019, 7: 93711–93722. [doi: [10.1109/ACCESS.2019.2927768](https://doi.org/10.1109/ACCESS.2019.2927768)]
- [30] Yang LQ, Liu JW. EEG-based emotion recognition using temporal convolutional network. In: *Proc. of the 8th IEEE Data Driven Control and Learning Systems Conf. (DDCLS)*. Dali: IEEE, 2019. 437–442. [doi: [10.1109/DDCLS.2019.8908839](https://doi.org/10.1109/DDCLS.2019.8908839)]
- [31] Li H, Jin YM, Zheng WL, Lu BL. Cross-subject emotion recognition using deep adaptation networks. In: *Proc. of the 25th Int'l Conf. on Neural Information Processing*. Siem Reap: Springer, 2018. 403–413. [doi: [10.1007/978-3-030-04221-9_36](https://doi.org/10.1007/978-3-030-04221-9_36)]
- [32] Jin YM, Luo YD, Zheng WL, Lu BL. EEG-based emotion recognition using domain adaptation network. In: *Proc. of 2017 Int'l Conf. on Orange Technologies (ICOT)*. Singapore: IEEE, 2017. 222–225. [doi: [10.1109/ICOT.2017.8336126](https://doi.org/10.1109/ICOT.2017.8336126)]
- [33] Wang F, Zhong SH, Peng JF, Jiang JM, Liu Y. Data augmentation for eeg-based emotion recognition with deep convolutional neural networks. In: *Proc. of the 24th Int'l Conf. on Multimedia Modeling*. Bangkok: Springer, 2018. 82–93. [doi: [10.1007/978-3-319-73600-6_8](https://doi.org/10.1007/978-3-319-73600-6_8)]
- [34] Yang BH, Han X, Tang JZ. Three class emotions recognition based on deep learning using staked autoencoder. In: *Proc. of the 10th Int'l Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. Shanghai: IEEE, 2017. 1–5. [doi: [10.1109/CISP-BMEI.2017.8302098](https://doi.org/10.1109/CISP-BMEI.2017.8302098)]
- [35] Wang XH, Zhang T, Xu XM, Chen L, Xing XF, Chen CLP. EEG emotion recognition using dynamical graph convolutional neural networks and broad learning system. In: *Proc. of the 2018 IEEE Int'l Conf. on Bioinformatics and Biomedicine (BIBM)*. Madrid: IEEE, 2018. 1240–1244. [doi: [10.1109/BIBM.2018.8621147](https://doi.org/10.1109/BIBM.2018.8621147)]
- [36] Luo Y, Zhang SY, Zheng WL, Lu BL. WGAN domain adaptation for EEG-based emotion recognition. In: *Proc. of the 25th Int'l Conf. on Neural Information Processing*. Siem Reap: Springer, 2018. 275–286. [doi: [10.1007/978-3-030-04221-9_25](https://doi.org/10.1007/978-3-030-04221-9_25)]
- [37] Alhagry S, Fahmy AA, El-Khoribi RA. Emotion recognition based on EEG using LSTM recurrent neural network. *International Journal of Advanced Computer Science and Applications*, 2017, 8(10): 355–358. [doi: [10.14569/IJACSA.2017.081046](https://doi.org/10.14569/IJACSA.2017.081046)]
- [38] Wen ZY, Xu RF, Du JC. A novel convolutional neural networks for emotion recognition based on EEG signal. In: *Proc. of the 2017 Int'l Conf. on Security, Pattern Analysis, and Cybernetics (SPAC)*. Shenzhen: IEEE, 2017. 672–677. [doi: [10.1109/SPAC.2017.8304360](https://doi.org/10.1109/SPAC.2017.8304360)]
- [39] Cimtay Y, Ekmekcioglu E. Investigating the use of pretrained convolutional neural network on cross-subject and cross-dataset EEG emotion recognition. *Sensors*, 2020, 20(7): 2034. [doi: [10.3390/s20072034](https://doi.org/10.3390/s20072034)]
- [40] Li YJ, Huang JJ, Zhou HY, Zhong N. Human emotion recognition with electroencephalographic multidimensional features by hybrid deep neural networks. *Applied Sciences*, 2017, 7(10): 1060. [doi: [10.3390/app7101060](https://doi.org/10.3390/app7101060)]
- [41] Li X, Song DW, Zhang P, Yu GL, Hou YX, Hu B. Emotion recognition from multi-channel EEG data through convolutional recurrent neural network. In: *Proc. of the 2016 IEEE Int'l Conf. on Bioinformatics and Biomedicine (BIBM)*. Shenzhen: IEEE, 2016. 352–359. [doi: [10.1109/BIBM.2016.7822545](https://doi.org/10.1109/BIBM.2016.7822545)]

- [42] Donmez H, Ozkurt N. Emotion classification from EEG signals in convolutional neural networks. In: Proc. of the 2019 Innovations in Intelligent Systems and Applications Conf. (ASYU). Izmir: IEEE, 2019. 1–6. [doi: [10.1109/ASYU48272.2019.8946364](https://doi.org/10.1109/ASYU48272.2019.8946364)]
- [43] Wang F, Wu SC, Zhang WW, Xu ZF, Zhang YH, Wu CD, Coleman S. Emotion recognition with convolutional neural network and EEG-based EFDMs. *Neuropsychologia*, 2020, 146: 107506. [doi: [10.1016/j.neuropsychologia.2020.107506](https://doi.org/10.1016/j.neuropsychologia.2020.107506)]
- [44] Wang KY, Ho YL, Huang YD, Fang WC. Design of intelligent EEG system for human emotion recognition with convolutional neural network. In: Proc. of the 2019 IEEE Int'l Conf. on Artificial Intelligence Circuits and Systems (AICAS). Hsinchu: IEEE, 2019. 142–145. [doi: [10.1109/AICAS.2019.8771581](https://doi.org/10.1109/AICAS.2019.8771581)]
- [45] Jirayucharoensak S, Pan-Ngum S, Israsena P. EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation. *The Scientific World Journal*, 2014, 2014: 627892. [doi: [10.1155/2014/627892](https://doi.org/10.1155/2014/627892)]
- [46] Moon SE, Jang S, Lee JS. Convolutional neural network approach for Eeg-based emotion recognition using brain connectivity and its spatial information. In: Proc. of the 2018 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP). Calgary: IEEE, 2018. 2556–2560. [doi: [10.1109/ICASSP.2018.8461315](https://doi.org/10.1109/ICASSP.2018.8461315)]
- [47] Xu HY, Plataniotis KN. Affective states classification using EEG and semi-supervised deep learning approaches. In: Proc. of the 18th IEEE Int'l Workshop on Multimedia Signal Processing (MMSP). Montreal: IEEE, 2016. 1–6. [doi: [10.1109/MMSP.2016.7813351](https://doi.org/10.1109/MMSP.2016.7813351)]
- [48] Tripathi S, Acharya S, Sharma RD, Mittal S, Bhattacharya S. Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset. In: Proc. of the 31st AAAI Conf. on Artificial Intelligence. San Francisco: ACM, 2017. 4746–4752.
- [49] Zhang T, Zheng WM, Cui Z, Zong Y, Li Y. Spatial-temporal recurrent neural network for emotion recognition. *IEEE Trans. on Cybernetics*, 2018, 49(3): 839–847. [doi: [10.1109/TCYB.2017.2788081](https://doi.org/10.1109/TCYB.2017.2788081)]
- [50] Candra H, Yuwono M, Chai R, Handojoseno A, Elamvazuthi I, Nguyen HT, Su S. Investigation of window size in classification of EEG-emotion signal with wavelet entropy and support vector machine. In: Proc. of the 37th Annual Int'l Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC). Milan: IEEE, 2015. 7250–7253. [doi: [10.1109/EMBC.2015.7320065](https://doi.org/10.1109/EMBC.2015.7320065)]
- [51] Cheng CL, Wei XW, Jian Z. Emotion recognition algorithm based on convolution neural network. In: Proc. of the 12th Int'l Conf. on Intelligent Systems and Knowledge Engineering (ISKE). Nanjing: IEEE, 2017. 1–5. [doi: [10.1109/ISKE.2017.8258786](https://doi.org/10.1109/ISKE.2017.8258786)]
- [52] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Proc. of the 27th Int'l Conf. on Neural Information Processing Systems. Montreal: ACM, 2014. 2672–2680.
- [53] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science*, 2006, 313(5786): 504–507. [doi: [10.1126/science.1127647](https://doi.org/10.1126/science.1127647)]
- [54] Chai X, Wang QS, Zhao YP, Liu X, Bai O, Li YQ. Unsupervised domain adaptation techniques based on auto-encoder for non-stationary EEG-based emotion recognition. *Computers in Biology and Medicine*, 2016, 79: 205–214. [doi: [10.1016/j.combiomed.2016.10.019](https://doi.org/10.1016/j.combiomed.2016.10.019)]
- [55] Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol PA. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *The Journal of Machine Learning Research*, 2010, 11: 3371–3408.
- [56] Li X, Zhao ZG, Song DW, Zhang YZ, Niu CY, Zhang JW, Huo JD, Li J. Variational autoencoder based latent factor decoding of multichannel EEG for emotion recognition. In: Proc. of the 2019 IEEE Int'l Conf. on Bioinformatics and Biomedicine (BIBM). San Diego: IEEE, 2019. 684–687. [doi: [10.1109/BIBM47256.2019.8983341](https://doi.org/10.1109/BIBM47256.2019.8983341)]
- [57] Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, 18(7): 1527–1554. [doi: [10.1162/neco.2006.18.7.1527](https://doi.org/10.1162/neco.2006.18.7.1527)]
- [58] Chao H, Zhi HL, Dong L, Liu YL. Recognition of emotions using multichannel EEG data and DBN-GC-based ensemble deep learning framework. *Computational Intelligence and Neuroscience*, 2018, 2018: 9750904. [doi: [10.1155/2018/9750904](https://doi.org/10.1155/2018/9750904)]
- [59] LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series. In: Arbib MA, ed. *The Handbook of Brain Theory and Neural Networks*. Cambridge: MIT Press, 1995.
- [60] Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. In: Proc. of the 27th Int'l Conf. on Machine Learning. Haifa: ACM, 2010. 807–814.
- [61] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proc. of the IEEE*, 1998, 86(11): 2278–2324. [doi: [10.1109/5.726791](https://doi.org/10.1109/5.726791)]
- [62] He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. In: Proc. of the 2016 IEEE Conf. on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 770–778. [doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90)]
- [63] Bai SJ, Kolter JZ, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv: 1803.01271, 2018.
- [64] Sabour S, Frosst N, Hinton GE. Dynamic routing between capsules. In: Proc. of the 31st Int'l Conf. on Neural Information Processing Systems. Long Beach: ACM, 2017. 3859–3869.

- [65] Liu Y, Ding YF, Li C, Cheng J, Song RC, Wan F, Chen X. Multi-channel EEG-based emotion recognition via a multi-level features guided capsule network. *Computers in Biology and Medicine*, 2020, 123: 103927. [doi: [10.1016/j.combiomed.2020.103927](https://doi.org/10.1016/j.combiomed.2020.103927)]
- [66] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735–1780. [doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735)]
- [67] Du XB, Ma CX, Zhang GH, Li JY, Lai YK, Zhao GZ, Deng XM, Liu YJ, Wang HA. An efficient LSTM network for emotion recognition from multichannel EEG signals. *IEEE Trans. on Affective Computing*, 2020. [doi: [10.1109/TAFFC.2020.3013711](https://doi.org/10.1109/TAFFC.2020.3013711)]
- [68] Torrey L, Shavlik J. Transfer learning. In: Soria E, Martin J, Magdalena R, Martinez M, Serrano A, eds. *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*. Cartagena: IGI Global, 2010. 242–264.
- [69] Ganin Y, Ustinova E, Ajakan H, Germain P, Larochelle H, Laviolette F, Marchand M, Lempitsky V. Domain-adversarial training of neural networks. *The Journal of Machine Learning Research*, 2016, 17(1): 2096–2030.
- [70] Long MS, Cao Y, Wang JM, Jordan MI. Learning transferable features with deep adaptation networks. In: *Proc. of the 32nd Int'l Conf. on Machine Learning*. Lille: JMLR, 2015. 97–105.
- [71] Philippot P. Inducing and assessing differentiated emotion-feeling states in the laboratory. *Cognition and Emotion*, 1993, 7(2): 171–193. [doi: [10.1080/02699939308409183](https://doi.org/10.1080/02699939308409183)]
- [72] Gross JJ, Levenson RW. Emotion elicitation using films. *Cognition and Emotion*, 1995, 9(1): 87–108. [doi: [10.1080/02699939508408966](https://doi.org/10.1080/02699939508408966)]
- [73] Koelstra S, Muhl C, Soleymani M, Lee JS, Yazdani A, Ebrahimi T, Pun T, Nijholt A, Patras I. DEAP: A database for emotion analysis; using physiological signals. *IEEE Trans. on Affective Computing*, 2012, 3(1): 18–31. [doi: [10.1109/T-AFFC.2011.15](https://doi.org/10.1109/T-AFFC.2011.15)]
- [74] Soleymani M, Lichtenauer J, Pun T, Pantic M. A multimodal database for affect recognition and implicit tagging. *IEEE Trans. on Affective Computing*, 2012, 3(1): 42–55. [doi: [10.1109/T-AFFC.2011.25](https://doi.org/10.1109/T-AFFC.2011.25)]
- [75] Katsigiannis S, Ramzan N. DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. *IEEE Journal of Biomedical and Health Informatics*, 2018, 22(1): 98–107. [doi: [10.1109/JBHI.2017.2688239](https://doi.org/10.1109/JBHI.2017.2688239)]
- [76] Gao ZK, Wang XM, Yang YX, Li YL, Ma K, Chen GR. A channel-fused dense convolutional network for EEG-based emotion recognition. *IEEE Trans. on Cognitive and Developmental Systems*, 2021, 13(4): 945–954. [doi: [10.1109/TCDS.2020.2976112](https://doi.org/10.1109/TCDS.2020.2976112)]
- [77] Li Y, Wang L, Zheng WM, Zong Y, Qi L, Cui Z, Zhang T, Song TF. A novel Bi-hemispheric discrepancy model for EEG emotion recognition. *IEEE Trans. on Cognitive and Developmental Systems*, 2021, 13(2): 354–367. [doi: [10.1109/TCDS.2020.2999337](https://doi.org/10.1109/TCDS.2020.2999337)]

附中文参考文献:

- [26] 张冠华, 余旻婧, 陈果, 韩义恒, 张丹, 赵国联, 刘永进. 面向情绪识别的脑电特征研究综述. *中国科学: 信息科学*, 2019, 49(9): 1097–1118. [doi: [10.1360/N112018-00337](https://doi.org/10.1360/N112018-00337)]



李锦瑶(1997—), 女, 硕士, 主要研究领域为情感计算, 情绪可视化与交互.



邓小明(1980—), 男, 博士, 副研究员, CCF 高级会员, 主要研究领域为计算机视觉, 人机交互.



杜肖兵(1992—), 女, 博士, 主要研究领域为人机交互, 多模态情感计算.



马翠霞(1975—), 女, 博士, 研究员, CCF 高级会员, 主要研究领域为人机交互, 媒体大数据可视分析.



朱志亮(1988—), 男, 博士, CCF 专业会员, 主要研究领域为图像智能感知与增强, 人机交互.



王宏安(1963—), 男, 博士, 研究员, CCF 高级会员, 主要研究领域为自然人机交互, 实时智能计算.