

研究论文

计算机辅助抑郁障碍识别研究进展

辜雅婷¹, 张迟¹, 马飞^{2*}, 贾晓健³, 倪士光^{1*}

摘要 抑郁症是一种典型的精神障碍,目前筛查主要以抑郁诊断量表和医生问诊为主。计算机辅助抑郁障碍识别是基于人工智能技术的一种抑郁症筛查的新兴方式。针对传统测量工具的现状和不足,综述了当前计算机辅助抑郁识别方法,论述了抑郁识别数据集和基于面部图片、语音、文本等多模态数据的抑郁识别方法的研究现状,并对计算机辅助抑郁识别的优势与挑战进行了总结与展望。计算机辅助抑郁识别能够提供一种相对简便、规范化的筛查方式,但仍面临模型参数和特征解释不足、中文数据集有待扩充、现有数据集样本量较少的挑战。未来研究人员需要进一步提升抑郁识别数据集的样本量及模型准确率等指标,进行特征提取及模型构建的理论及实验分析,推进计算机辅助抑郁识别的临床应用。

关键词 抑郁障碍;抑郁识别;筛查量表;计算机辅助检测;多模态数据

近年来,各国关于抑郁症(major depressive disorder, MDD)的报道层出不穷,抑郁患者自杀的新闻频出。根据美国精神病学会的《精神障碍诊断与统计手册》(DSM-5)诊断标准,抑郁症是一种典型的精神障碍,临床特征包括显著而持久的心境低落、兴趣丧失和精力缺乏^[1]。据世界卫生组织(WHO)发布的《抑郁症及其他常见精神障碍》报告显示,全球范围内抑郁症患者超过3亿人,预计在未来10年将会成为全球第一大疾病负担源^[2]。《柳叶刀·精神病学》上发表的对中国全国性精神障碍流行病学调查研究显示,抑郁症的终生患病率为6.8%,12个月患病率为3.6%^[3]。抑郁症的发病机制比较复杂,遗传因素、心理-社会因

素均可能导致抑郁症的产生^[4]。社会对抑郁症的认知偏见,患者自身及家庭对心理问题、精神疾病的认识不足,导致抑郁症识别率和就诊率仍处在较低水平^[5-7]。

目前,抑郁症的筛查主要以患者自评量表和医师临床诊断为主,可能受到医生主观判断^[8]和患者自我认知^[9]的影响,筛查工具和诊断方式具有一定局限性。抑郁症心理行为特征相关研究表明,患者在运动行为^[10-11]、语言表达^[12-14]、身体姿态^[15-16]和生理指标等方面表现出不同于健康人群的特征。在抑郁症的筛查和诊断中,对患者的心理和生理症状进行准确的识别至关重要。通过医生的临床经验和大规模的研究分析,能够提取出用于识别抑郁患者的特征标志。基于这些抑郁症状标志特征,计算机辅助抑郁识别技术能从行为、语言等维度提取有效区分抑郁患者与健康人群的特征标志,构建抑郁筛查模型,辅助医生进行

1. 清华大学深圳国际研究生院,深圳 518055

2. 人工智能与数字经济广东省实验室(深圳),深圳 518107

3. 深圳市精神卫生中心/深圳市康宁医院研究所,深圳 518118

收稿日期:2023-03-12;修回日期:2024-12-27

基金项目:深圳市可持续发展科技专项(KCXFZ20230731093600002);深圳市自然科学基金重点项目(JCYJ20241202124000001);深圳市哲学社会科学规划课题(SZ2024B038);广东省数字心理健康与智能生成普通高校哲学社会科学重点实验室项目(2023WSYS010)

作者简介:辜雅婷,硕士研究生,研究方向为人工智能、心理和行为大数据,电子信箱:gyt20@mails.tsinghua.edu.cn;马飞(通信作者),研究员,研究方向为媒体内容理解与生成,电子信箱:mafei@gml.ac.cn;倪士光(共同通信作者),教授,研究方向为积极和健康心理学及人工智能交叉创新,电子信箱:ni.shiguang@sz.tsinghua.edu.cn

引用格式:辜雅婷,张迟,马飞,等.计算机辅助抑郁障碍识别研究进展[J].科技导报,2025,43(14):82-93;doi:10.3981/j.issn.1000-7857.2023.03.00331

抑郁症状的诊断和筛查。

计算机辅助识别算法能够从大规模的数据集中提取特征进行分类,输入机器学习模型中进行参数训练和模型调整,并基于训练的模型对新的数据样本进行预测。该方法的有效性依赖于特征提取的准确性和可靠性,特征选择容易对模型的性能和预测效果产生明显影响。随着深度学习等技术的发展,基于神经网络的分类器能够与深度特征提取器训练得到端到端的分类器,该方法相对手工提取特征的方法更加简便,但这种方法存在模型特征的解释性较弱、计算复杂度较高等问题。目前上述2类计算机辅助筛查算法均有相关学者进行研究,相比于抑郁自评量表等测量工具,计算机辅助方法包含更加丰富的多模态信息,能够进一步提升筛查的客观性,具有较高的应用和研究价值。

1 传统测量工具概述

目前,广泛应用的抑郁症筛查主要通过患者自评量表、医师临床诊断等方式,根据情绪低落、睡眠障碍、能力减退等抑郁典型症状出现的频次、严重程度及持续时长,评估是否有抑郁障碍及其程度。

抑郁障碍自评或他评量表等工具能够较直观地体现抑郁障碍的诊断标准,帮助医生和研究人员快速而准确地评估患者的情况,是常用的评估方法之一。贝克抑郁自评量表(BDI)^[17]、PHQ-9 抑郁症筛查量表(PHQ-9)^[18]、汉密尔顿抑郁量表(HAMD)^[19]、抑郁自评问卷(SDS)^[20]等是目前主流的评估筛查量表,能够用于评估测评者是否有抑郁症状及抑郁程度。临床上对抑郁症的识别诊断主要基于国际疾病分类(ICD-10)^[21]或 DSM-V^[1]中抑郁症诊断标准,结合病人的访谈情况、筛查量表及医生诊疗经验进行。

BDI 是一份自我报告问卷,包括 21 项,一般耗时 5~10 min。BDI 在应用中发现部分抑郁患者完成 21 项测评时存在困难,因此该量表也有修订编制的 13 项版本。评估者需要从认知、情感和躯体症状等方面对自我进行评估,用于诊断自我的抑郁症状程度。

PHQ-9 是临床上对抑郁进行初步筛查的简易量表。PHQ-9 主要用于测评在过去 2 周内是否表现出相应的抑郁症状、严重程度及持续时间,根据测评者

不同症状出现的频次总分进行评估。该筛查量表的测评时间往往在 5 min 内,临床研究证明具有良好的信度和效度。

HAMD 是抑郁症标准的评估工具之一,主要用于诊断和研究目的。HAMD 分为 17 项、21 项和 24 项等多个版本,一般需要 20~30 min 完成。评估者采用交谈和观察的方式,从抑郁情绪、自杀倾向、睡眠情况等角度对测评者进行评定,评定分数能够较好反映是否患有抑郁症及抑郁症的严重程度。

SDS 是美国教育卫生部推荐用于精神药理研究的量表之一,包含 20 个评分题,涵盖精神性-情感症状、躯体性障碍、精神运动性障碍、抑郁性心理障碍相关问题,上述 4 方面的问题占比为 10%、40%、10%、40%。SDS 的使用和分析较为简单方便,不需要经过专门训练的医师指导评定,评定分数能够直观反映出患者的主观感受,目前已广泛应用于门诊病人的粗筛和情绪状态评定。

上述测评量表在临床实践中得到广泛应用,每个量表都具有其独特的优缺点,研究人员或专业医师可根据具体情况进行选择。在计算机辅助抑郁障碍识别数据集的构建中,抑郁测评量表也具有重要作用。由于有监督学习的机器学习算法依赖于准确的标签数据,目前广泛使用的基于人工智能技术的抑郁识别数据集中,往往以一个或多个上述量表的测评结果或医生诊断结果作为数据的真实标签。

2 计算机辅助抑郁障碍识别

传统测量工具可能受到医生主观判断和患者自我认知的影响,而计算机辅助识别能够通过患者图像、语音和文本等获取更加丰富的多模态信息,进一步提升筛查的客观性。基于人工智能技术的计算机辅助识别作为一种跨学科技术,利用计算机视觉、自然语言处理等人工智能技术,建立机器学习模型,基于输入模型的数据训练调整模型参数,再根据构建模型计算新样本的预测结果和类别概率,辅助疾病的诊断和筛查。根据抑郁患者在面部图像、语音、文本、脑电(EEG)等模态数据上的差异,国内外的研究提取不同模态的特征,可以从多角度建立抑郁诊断评估工具(图 1)。

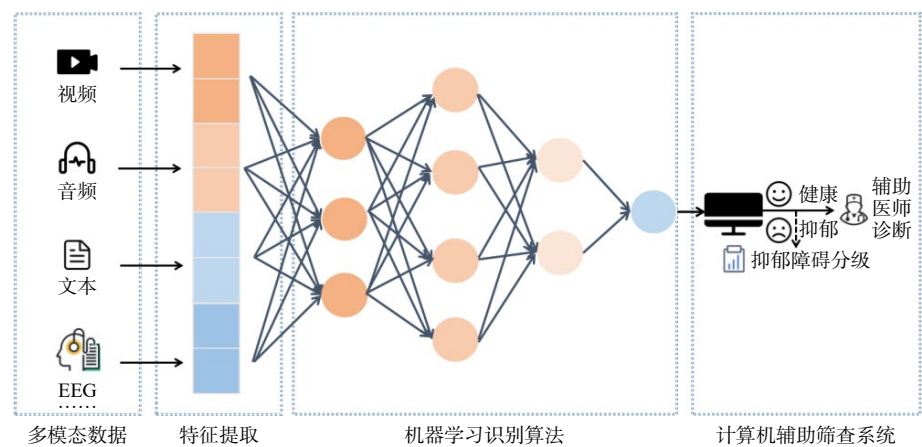


图 1 基于人工智能技术的计算机辅助抑郁诊断示意

2.1 计算机辅助抑郁识别数据集

计算机辅助筛查通过采集抑郁患者和健康对照者的视频、语音等模态数据,提取用于区分抑郁症状的特征,训练机器学习模型对新的潜在抑郁症患者进行预测。机器学习算法模型的训练依赖足够的数据以减少模型过拟合、提高模型精度,标准的抑郁识别

数据集的建立对于筛查算法的训练至关重要。近 10 年,该领域开放的外文数据集包括: AVEC2013^[22]、AVEC2014^[23]、DAIC-WOZ^[24]、Pittsburgh^[25]等,中文数据集包括: MODMA^[26]、EATD^[27]、CMDC^[28]、Wenzhou-Kangning^[29]等,数据集概述如表 1 所示。

表 1 近 10 年抑郁识别领域开源数据集

数据集	时间/语音	数据模态	数据量(抑郁样本:健康样本)
AVEC2013 ^[22]	2013/德文	视频、音频	292
AVEC2014 ^[23]	2014/德文	视频、音频	84
DAIC-WOZ ^[24]	2017/英文	视频、音频、文本	189(56:133)
Pittsburgh ^[25]	2017/英文	视频、音频	49(49:-)
MODMA ^[26]	2020/中文	音频、脑电	128通道脑电: 53(24:29) 3通道脑电: 55(26:29) 音频: 52(23:29)
EATD ^[27]	2022/中文	语音、文本	162(30:132)
CMDC ^[28]	2022/中文	视频、音频	视频+音频: 45(19:26) 音频: 78(26:52)
Wenzhou-Kangning ^[29]	2023/中文	音频、文本	113(61:52)

开源数据集为机器学习模型的训练提供了基础,推动了人工智能辅助筛查领域的发展,对抑郁识别领域的发展有明显的推动作用。数据集往往根据量表筛查或医师诊断出的抑郁人群或健康对照者给定标签,通过患者参与朗读、描述、访谈等实验采集视频、语音、脑电等多种数据模态。

在抑郁识别领域,自行采集和标注的数据集也具有一定的研究和应用价值。虽然这些数据集通常并未公开获取,但它们为研究不同文化背景对抑郁症状表达的影响,以及基于不同数据模态的抑郁识别算法

的构建提供了理论基础和实验支撑。为分析中国抑郁症患者的面部特征, Wang 等^[30]采集了山东省精神卫生中心的 26 名抑郁症住院患者和 26 名健康对照者在观看不同情绪图片时的面部视频数据,以分析抑郁症患者眉毛、眼睛和嘴巴处的特征,进一步扩充了中国患者的抑郁筛查数据。根据抑郁症患者肢体运动等模态上的特点, Wang 等^[31]通过 Kinect 采集了 126 名抑郁症患者和 121 名健康者的步态数据,分析抑郁症患者肢体摆动、步幅变化和头部姿势及运动的特征,进一步扩充了与肢体运动相关的抑郁筛查数据。

2.2 面部视觉模态模型构建

抑郁人群的面部表达能力受损,且更易于表现出眼睛松弛、皱眉等悲伤情绪特征^[32]。根据抑郁症患者的面部表达特点,目前抑郁识别算法通过提取人脸面部关键特征进行筛查(图2),包括人脸特征点坐标、面部动作单元(AU)等。美国心理学家 Ekman 提出了一套面部表情编码系统(FACS)以描绘不同脸部肌肉动作和表情之间的对应关系^[33]。该系统根据人脸学特点,将面部表情划分得到若干相互独立又相互联系的运动单元,以准确地识别面部情绪和情感表达。2009年, Cohn 等^[34]对参与者在回答 HAMD 中涉及情绪低落、内疚和自杀的3个问题时的面部行为进行手动 FACS 编码,结果显示根据所有 AU 进行分类能够得到 79% 的准确率,其中涉及夹肌收缩的 AU14 分类的准确率最高,得到的灵敏度和特异度分别为 87%

和 89%。2013 年, Meng 等^[35]从连续的图像序列中计算运动历史直方图,通过局部二值模式(LBP)描述时序空间的细节,拼接得到的时空描述子通过偏最小二乘回归,在 AVEC2013 数据集上达到平均绝对误差(MAE)=7.08、均方根误差(RMSE)=8.81。2015 年, Pampouchidou 等^[36]提出一种使用鲁棒描述符动态进行面部表情分析的方法,结合 Curvelet 变化和 LBP-TOP 得到人脸特征,建立对抑郁症无、轻度、中度和重度等严重程度的有效的分类系统,混淆矩阵显示各分级的识别准确率分别为 51.0%、63.4%、55.0%、55.5%。2018 年, Wang 等^[30]根据从山东省精神卫生中心采集到的抑郁症住院患者的临床视频样本,从中提取眼睛、眉毛和嘴角的运动变化,手动提取计算了左右瞳孔间距、眨眼频率、双侧眉毛和眼角间距等特征并通过 SVM 算法进行分类,达到了 78.85% 的准确率。

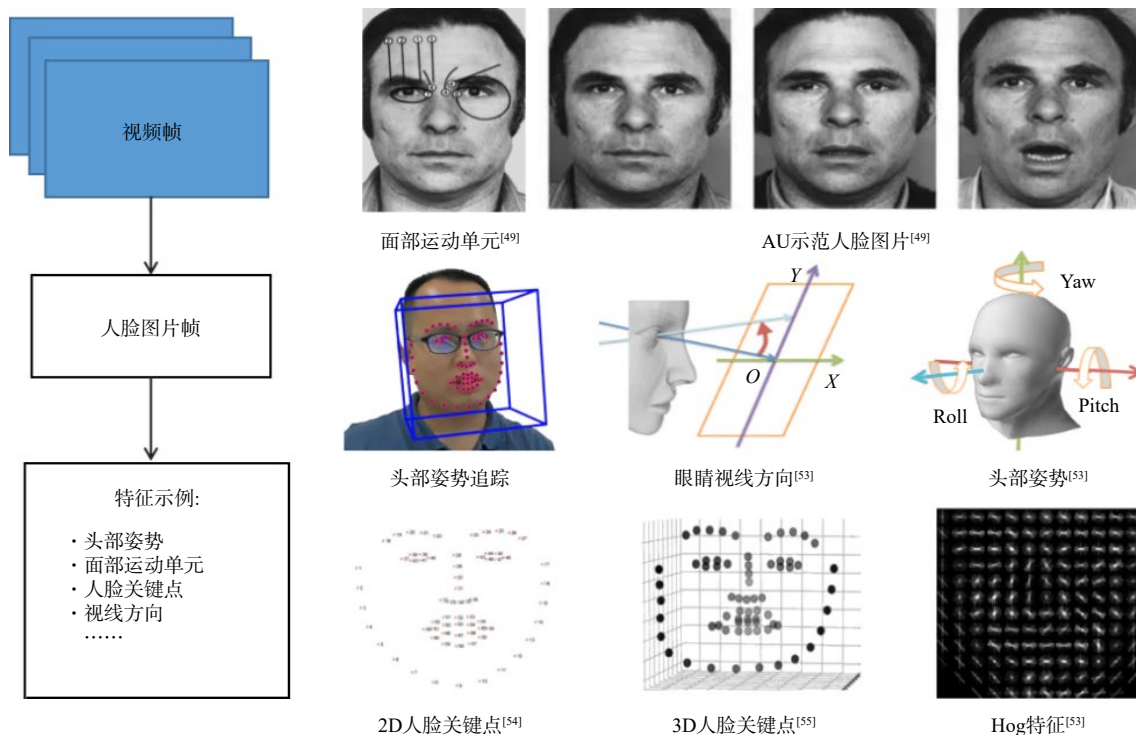


图2 抑郁识别视觉模态特征标志

随着深度学习技术的发展,目前的研究不再局限于手工提取特征——从原始图像数据中提取特征后进行分类、回归,而是通过卷积神经网络(CNN)等端到端学习的方法进行特征提取和模型训练,该模型也表现出了较高的准确性和鲁棒性。CNN 是一种在图像识别领域表现优异的深度学习模型。通过卷积层、

池化层和全连接层等基本结构,能够从原始图像数据中自动提取具有高层次语义的特征表示,以进行分类、回归等任务。在抑郁症分类诊断方面,CNN 模型通过训练较大规模的抑郁样本数据,自动学习面部表情特征与抑郁症之间的联系,并对新的面部图像进行准确的分类预测。2017 年, Zhu 等^[37]通过深度卷积神

神经网络(DNN)分别提取视频中人脸静态外观和跨帧动态特征,构建联合调整层进行最终微调。在 AVEC2014 数据集上跨帧动态特征模型相较静态外观特征模型在 RMSE、MAE 上分别提升 0.56、0.3,联合微调模型达到 RMSE=9.55、MAE=7.47。

2D CNN 已被广泛用于图像分类等领域,但其在处理带有时间信息的数据,如时间序列和视频数据方面存在局限性。为了解决这一问题,研究者提出了 3D CNN,该模型利用卷积操作在时间维度上提取特征,从而能够有效地处理带有时间信息的数据。通过引入时间维度,3D CNN 可以从时空域的角度分析数据,并且能够在时域上学习数据中的时间相关性,从而提高对时间序列和视频数据的建模能力。相比于 2D CNN,3D CNN 在处理时间序列和视频数据方面具有更好的表现,能够有效地应用于抑郁识别模型的构建。2019 年,de Melo 等^[38]从采集对象脸部全局和眼睛局部区域中提取时空特征,通过 3D 卷积网络(C3D)融合来提高抑郁症预测的准确率,在 AVEC2013 和 AVEC2014 上的实验结果表明,结合全局和局部的 C3D 方法 RMSE 达到 8.26、MAE 达到 6.40,在 AVEC2014 数据集上相对于全局 C3D 方法在 RMSE 和 MAE 指标上分别有 0.68 和 0.64 的提升。

3D CNN 能够广泛应用于视频分类、动作识别等任务,并取得了显著的成果。然而,3D CNN 在处理序列数据时存在局限性——它无法充分考虑序列中不同时间步之间的关系。为了解决这个问题,学者们引入了注意力机制(attention),通过为不同时间步赋予不同的权重来强调序列中的关键部分,从而提高 3D CNN 的性能。这种方法能够让 3D CNN 更加关注序列中重要的部分,提高其对于序列数据的建模能力。注意力机制的引入可以使 3D CNN 能够更好地处理时间序列数据,并在视频模态的抑郁识别任务上获得了良好的表现。2021 年,孙浩浩等^[39]提出通道层注意力机制的 DNN,通过多支路卷积网络分别提取眼睛区域、嘴巴区域和面部的特征向量进行融合,在全连接层输出最后分数。该实验在 AVEC2013 和 AVEC2014 训练集上得到加入了注意力机制、融合了多特征的模型,在 AVEC2013 数据集上达到 MAE=6.74、RMSE=8.70,AVEC2014 数据集上达到 MAE=6.56、RMSE=8.56,优于基线模型和其他对比模型。2022 年,Chen

等^[40]通过神经网络结构搜索技术(NAS)设计了一个针对多面部特征的优化模型,通过 CNN-GCN 的端到端网络,在 AVEC2016 数据集上 RMSE 和 MAE 相较于 SOTA 提升了 27% 和 30%。2023 年,Liu 等^[41]提出一种部分和关系注意力网络,通过 2 种注意力机制计算不同局部特征对抑郁识别的贡献,进一步将所有特征聚合成更能提供抑郁症信息的表示并用于抑郁识别,在 AVEC2013 和 AVEC2014 上实现了最优性能,MAE 和 RMSE 分别达到 6.08 和 7.59。

在抑郁识别领域,识别模型的迭代与深度学习的演化发展紧密相关。最初,应用 2D CNN 技术进行面部静态二维图像的分类和识别,虽然简单高效,但存在对时间序列信息处理不足的问题,对采集的视频数据无法很好挖掘帧与帧间的信息。3D CNN 技术的引入扩展了模型对于动态视频数据的建模能力,更好地利用了时间维度的信息。随着注意力机制的广泛应用,有注意力的 3D CNN 框架进一步增强了模型对数据的理解 and 处理能力,为处理复杂数据的抑郁识别问题提供了更有效的工具和方法(图 3)。这种演化从简单到复杂、从二维到三维、从静态到动态,提高了对于数据中关键特征的挖掘,并进一步提升了模型在准确率等方面的表现,对于抑郁识别问题的研究具有重要的意义和实际应用价值。

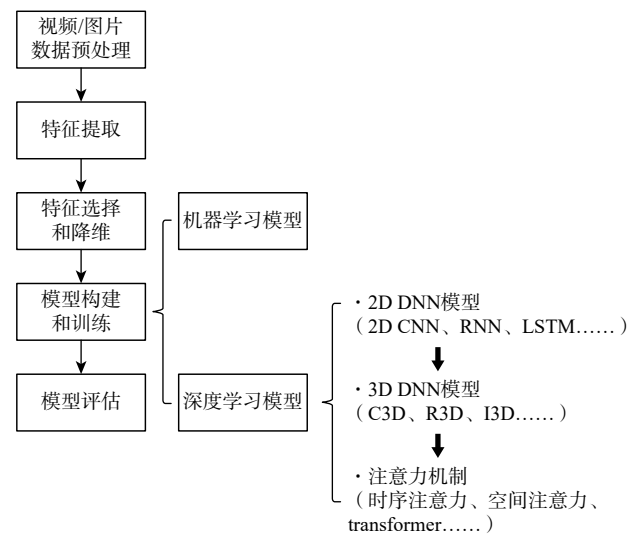


图3 抑郁识别视觉模态模型流程

2.3 语音模态模型构建

音频信号是抑郁症筛查中重要模态之一,能够有效反映人的心理状态和病理学特征。目前针对音频

模态的抑郁识别研究以手动提取特征为主,即针对抑郁患者语速慢、声音低沉等言语表达中的标志特点,通过对采集的抑郁患者的声音信号进行分析处理,提取音频的能量、响度、梅尔频谱、过零率等方面的差异特征,建立算法模型进行训练和预测,从而实现较客观有效的辅助筛查(图4)。由于特征提取的有效性对模型的性能有重要影响,因此特征的设计一直是研究者关注的重点。手工特征的提取需要通过多个预处理步骤,例如根据短时傅里叶变换(STFT)等方法,将原始音频信号转换为频谱图等表示形式。这些特征经过降维、归一化等处理后,输入到分类模型进行训练和预测。该过程需要专家的领域知识和经验,其设计能够反映研究者对标志特征的构建和筛选,加强对特征的全面解释,但往往也比较耗时和复杂。2013年,Joshi等^[42]计算了基频 f_0 、响度、强度和梅尔

频率倒谱系数(MFCC)等音频特征的有效性,融合从视频中提取的视觉特征在30名抑郁患者和30名健康对照者的临床实验数据上进行分析,得到不同特征组合在各分类算法上的最大灵敏度、特异度和准确率,最高准确率达到65.92%。2014年,Ooi等^[43]提出多通道加权语音分类方法,对韵律、声门和频谱特征独立分类后进行加权,判断一个人在未来2.5年内出现抑郁症状的风险,准确率达到74%。2017年,Kiss等^[44]分别对阅读和自发讲话场景下的特征进行提取,分析得到在自发讲话任务场景下的差异体现在与速度(语速、停顿时长等)相关的特征,在朗读任务场景下的差异体现在共振峰轨迹等语音特征,最终在朗读语音样本上达到83%的检测准确率,在自发语音样本上达到86%的检测准确率。

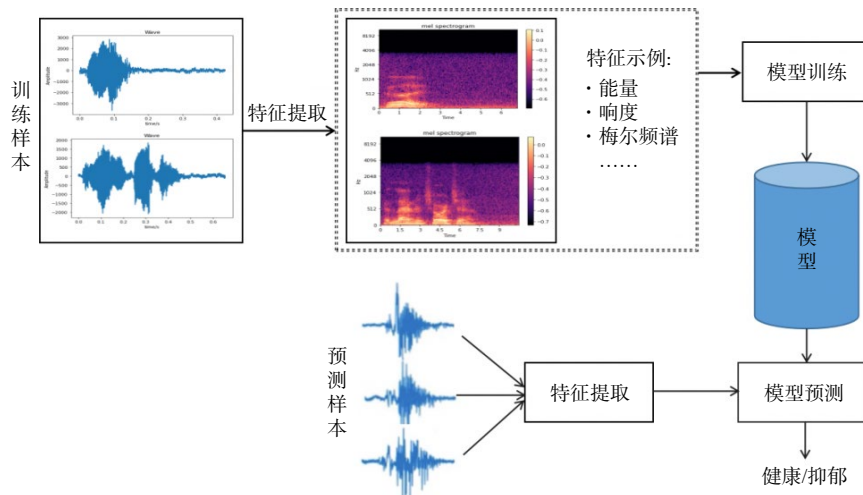


图4 抑郁识别语音模态模型流程

随着深度学习技术的发展,越来越多的研究开始探索在音频信号处理中使用端到端的深度网络。端到端的音频识别网络能够直接从原始音频数据中学习特征,并进行分类任务。这种方法使用CNN、循环神经网络(RNN)等深度学习模型,通过多个卷积层、池化层等对原始音频数据进行处理,提取出音频的高层次特征,然后再将这些特征输入到全连接层进行分类预测。由于不需要额外的特征提取步骤,端到端的音频识别网络能够有效减少特征工程的工作量和误差,且在抑郁识别实验中表现良好。2018年,He等^[45]提出了一种手动特征提取和深度学习特征提取相结合的方式,先通过DNN从频谱图和原始语音波形

图中学习特征,然后手工从频谱图中提取纹理描述符,并通过联合微调层结合手工和网络特征以提高抑郁症识别性能,在AVEC2013和AVEC2014数据集上的RMSE和MAE指标优于对比算法。2021年,Zhang等^[46]提出了抑郁检测的音频嵌入方法DEPA,该方法训练了自监督学习音频嵌入模型,在抑郁症和健康对照者数据集上进行预训练后应用于下游任务的模型分类,F1值达到94%。2022年,Sardari等^[47]通过端到端卷积神经网络的自动编码器CNN AE自动提取相关特征,并采用基于聚类的抽样技术,在DAIC-WOZ数据集上相较于对比方法在F1值上提升了7%。2023年,Du等^[48]提出一种结合线性预测编

码(LPC)和 MFCC 的语音链模型,分别描述语音生成和感知的过程,通过 CNN 和长短时记忆网络(LSTM)依次捕获段内和段间的动态抑郁特征进行分类,在 DAIC-WOZ 和 MODMA 数据集上分别取得 77% 和 86% 的准确率。2024 年, Das 等^[49]结合 MFCC 和通过 CNN 提取的高级频谱特征,在 DAIC-WOZ 数据集和 MODMA 数据集上的检测准确率超过 90%。

2.4 文本模态模型构建

文本数据也是抑郁识别的重要模态之一,研究人员通常会收集社交媒体平台如 Twitter、Facebook 等用户发布的文本数据,通过自然语言处理和文本挖掘等技术挖掘数据中有关个体的健康状况、情绪状态等方面的信息,用于辅助评估抑郁的严重程度及潜在风险。然而,在许多国家和地区存在法律限制和隐私保密问题,诸多研究采用自行采集的文本数据集或从音频数据中转录得到文本数据。

目前的研究根据抑郁症语言使用模式特点,通过词频计数、情感极性分析、词嵌入等方法提取文本中与抑郁症状标志关联的特征,从而进行抑郁识别算法分类(图 5)。抑郁文本分析的相关研究主要包括基于规则的方法和基于机器学习的方法。基于规则的文本分析利用预定义的规则和模式来解析和提取文本中的有用信息。这种方法通常使用正则表达式和语法解析器来识别和提取文本中的关键信息。基于机器学习的方法使用大规模文本数据进行训练,可以在分类和识别等下游任务中应用,具有高准确、可扩展等优点,通过改变模型参数和算法,适应各种不同的文本数据,还能够通过增加训练数据和特征选择等方法,进一步提高模型的性能。但这种方法需要大量的训练数据和计算资源,对于文本特征的解释性较弱,使得该方法在深入了解抑郁症的机理和相关因素方面存在局限性。

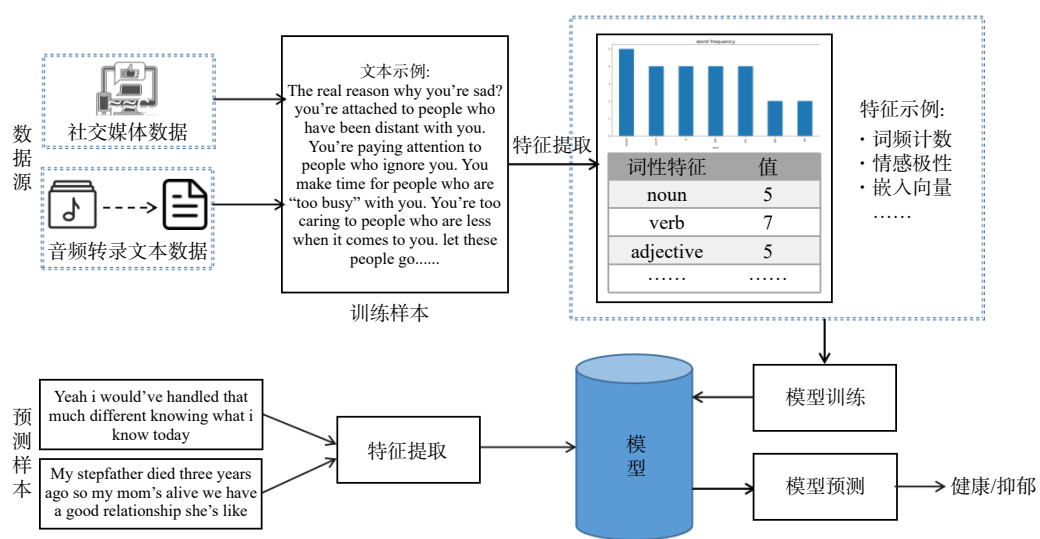


图 5 抑郁识别语音模态模型流程图

随着深度学习等技术的发展,通过词嵌入等方法将文本中的单词或短语进行向量表示的研究愈来愈普遍。目前应用于文本分类模型的典型特征生成方法包括 OneHot 编码、关键词权重(TF-IDF、Text-Rank 等)、主题模型(LSA、LDA、LDA2Vec 等)、词嵌入(BERT Embedding)等。2016 年, Williamson 等^[50]从 DAIC-WOZ 数据集的文本中分别提取了语义内容特征和语义上下文特征。语义内容特征由 GloVe 词嵌入模型生成词向量的平均向量表示对话语句,用主成分分析(PCA)和零相位成分分析(ZCA)进行稀疏

编码后,通过支持向量回归(SVR)算法对表示向量和 PHQ-9 分数进行回归。语义上下文特征从文本中提取抑郁、感受、治疗和自杀 4 个指标,计算 4 个指标的总和。结果显示, DAIC-WOZ 数据集本文对话中提问部分的最大 F1 值达到 75%, 回答部分的最大 F1 值达到 62%。2022 年, Amanat 等^[51]通过独热编码和主成分分析从 Twitter 抑郁文本数据集中提取特征,进一步使用 RNN-LSTM 模型进行预测, 10 折交叉验证的结果显示模型达到 99.44% 的灵敏度和 99.7% 的特异度。2022 年, Uddin 等^[52]从医学和心理

学专家预定义的抑郁症状作为特征,采用 LSTM 算法进行分类,在挪威公共在线频道年轻用户的文本数据集拆分的 2 个数据集上分别达到 98%、99% 的准确率,并通过可解释人工智能(XAI)对机器学习模型的结果进行了解释。2023 年, Yadav 等^[53]提出一种基于患者访谈文本转录的抑郁识别方法,结合了用于处理语言信息的双向门控循环单元(BGRU)和用于集成模型输出的全耦合网络,该方法可以提高识别的准确性和效率, F1 值达到 92%。

2.5 多模态融合及模型构建

计算机辅助抑郁识别依赖视觉、语言和社会特征等多维指标的综合判断,计算机视觉、自然语言处理、语音处理等领域的专家倾向于应用自己对应领域的算法模型,但单一模态数据不足以提供足够的信息,需要综合多个领域、多种模态的专家知识,以实现抑郁相关多维指标特征之间的有效信息融合^[54]。多模态技术运用多种不同的信息以更好地捕捉和理解事物的复杂性,广泛应用于情感识别等领域。在这些领域中,多模态数据可提供关于事物的多方面信息,对于理解事物的复杂性和捕捉不同维度的特征提供有力支持,进一步提高抑郁检测识别模型的准确性。

多模态特征融合抽取来自不同模态数据的信息,整合成更丰富的多模态特征表征,以提高模型的鲁棒性。但是,特征的有效融合仍是目前的研究难点之一。目前比较常见的特征融合方法有特征拼接^[55]、跨模态注意力机制^[56]和条件批归一化(CBN)^[57]等。其中,特征拼接是一种简单直观的特征拼接方式,将来自不同模态的特征向量进行拼接或加权,拼接后的融合特征向量输入神经网络。跨模态注意力机制则通过计算模态间的注意力权重,对模态特征进行加权融合。CBN 则是通过在批归一化中引入额外的条件信息,来增强模型对不同模态数据的适应能力。

随着人工智能和深度学习技术的进步,尤其是跨模态的特征融合技术的不断发展,近几年利用多模态数据进行抑郁识别的研究受到了越来越多研究者的关注。通过结合多模态的数据,能够更全面地了解患者的情况,提高抑郁识别模型的准确性和有效性。2019 年, Wang 等^[58]融合了眼球运动、记忆特征、认知风格特征和网络行为特征用于诊断抑郁倾向,通过模型堆叠的方式提取非线性特征,达到模型泛化效果。

2021 年, SchulteBraucks 等^[59]从非结构化临床访谈中提取面部、语音、语言和运动特征作为输入对创伤后应激障碍(PTSD)和抑郁症状的分类,解释了患者和健康人群在运动协调、处理速度、情绪偏差、持续注意力、控制注意力、认知灵活性、认知抑制和执行能力等认知功能上的差异,结果显示基于视频和音频的标记能准确区分 PTSD 状态, F1 值分别为 0.83、0.8。2022 年, Liu 等^[60]通过采集参与者在自然行走、情景访谈、阅读情感文本和观看情感视频等任务时的视频和音频,分别提取视频帧和音频特征,利用 VGG-Face、ResNet50、VGG16 和 DenseNet 网络得到不同模态的特征向量,通过门控循环单元(GRU)分别决策后进行结果融合,在测试集上的负性样本、中性样本和正性样本的 RMSE 分别为 10.59、10.64 和 10.13。2023 年, Fang 等^[61]提出了一种具有多层次注意力机制的多模态融合模型,第一阶段通过 2 个 LSTM 和一个双向 LSTM(Bi-LSTM)分别提取视觉、音频和文本特征,第二阶段通过注意力融合网络利用模态之间的多样性和互补性进行抑郁识别,在 DAIC-WOZ 数据集上实现了 RMSE 最优。

2.6 其他模态数据及模型构建

肢体运动、脑部信号等生物标志也能够用于抑郁识别。根据抑郁患者步态和运动等方面表现出的症状, Wang 等^[31]通过参与者的步态数据提取左右臂摆动、头部垂直运动、左右步幅和步行速度等时空特征进行抑郁识别。特征包括 Kinect 采集的 25 个关键点在 x 、 y 、 z 轴上的均值、标准差、偏度和峰度等时域特征,以及离散傅里叶变换得到的频域特征。模型融合所有特征进行 10 折交叉验证,分类灵敏度、特异度和 AUC 分别达到 0.94、0.91 和 0.93。针对抑郁患者的脑电信号差异,深度学习方法,尤其是 CNN 和 LSTM,被广泛应用于特征提取、抑郁识别分类和评估任务^[62]。2020 年, Zhang 等^[63]提出了一种基于静息态脑电的抑郁症分类脑功能网络架构,以考虑电极之间的相关性和大脑异常拓扑结构的变化。对不同脑电频段和不同脑区进行统计分析,结果表明大脑的改变主要在左脑的额叶、颞叶、顶枕区和右脑的颞叶区,且 θ 频带左中央区域的平均最短路径长度、聚类系数以及右顶枕区域的节点中介中心度与抑郁症的 PHQ-9 评分显著相关,最高分类准确率达到 93.31%。

2023 年, Xu 等^[64]针对静息态脑电, 提出一种多分辨率 CNN 结合 LSTM 以及残差挤压和激励(RSE)的深度学习模型, 结果表明脑电频段越高抑郁识别效果越好, 在高频 8~30 Hz 实现了 $(98.48 \pm 0.22)\%$ 的最高分类准确率。

3 计算机辅助抑郁识别的优势与挑战

现有研究根据抑郁患者在面部表情、语言表达和肢体动作等特征上的差异, 从图像、音频和文本多模态数据中提取特征, 通过机器学习算法对抑郁症状进行识别和分类, 在准确率、召回率、F1、RMSE 和 MAE 等指标上达到了良好的效果, 说明心理学、临床医学等领域提取的标识特征在抑郁识别领域的有效性。

计算机辅助识别方法能够提供一种相对简便、规范化的筛查方式, 具备与目前广泛使用的量表筛查、医师诊断等方式协同增效的潜能, 尤其针对传统抑郁筛查量表中存在社会赞许性误差、对主观判断依赖高等问题, 计算机辅助筛查能够提供更加客观的筛查结果, 具体表现为以下 3 方面。

1) 现有诊断方式依赖于患者对自我的合理评估, 但由于心理测量中普遍存在的社会赞许性等问题, 患者对自己的评估往往倾向于偏离难以被社会接受的消极行为, 偏向能够被社会认可的积极行为, 导致心理测评的结果出现偏差。而计算机辅助筛查能够根据患者的面部行为、语言表达等数据, 规范提取筛查指标、构建识别模型, 结果相对客观准确, 能够辅助医师进行抑郁症筛查。

2) 抑郁的诊断筛查方法较为繁琐, 在社会医疗资源紧张的情况下难以广泛大规模推广使用。目前, 抑郁诊断很大程度依赖行为症状学指标, 在临床实践上具有可操作性, 但依赖医师对患者心理状况的问询和对临床表现分析, 对医师的主观判断依赖程度较高, 容易受到医师临床经验、沟通技巧等因素影响, 且诊疗的成本较高, 容易给患者及其家庭带来经济负担。而计算机辅助筛查方式较容易广泛地筛查, 尤其是针对抑郁高风险的人群进行便捷、快速的检测。此外, 人工智能模型的训练和评估过程可以在大规模样本上进行, 从而构建更具代表性和准确性的模型。

3) 由于对抑郁症等心理精神疾病的病耻感与文

化易感性, 中国抑郁患者相较于国外患者更容易表达躯体症状的不适, 包括失眠、体重减轻、食欲不振、头痛、肠胃道或呼吸系统症状等, 在问诊过程中症状表述容易和其他综合性疾病混淆, 出现误诊和漏诊。因此, 采集适用于中国患者的抑郁识别数据、构建相应的模型算法和筛查系统对于筛查的准确性至关重要。目前抑郁识别算法达到了良好的识别效果, 但仍面临诸多挑战, 包括中文数据集有待扩充、现有数据集样本量较少、模型参数和特征解释不足等现象, 有待进一步研究。

首先, 相关研究的开放数据集以英文居多, 中文数据集有待扩充。由于国内外抑郁患者在面部特点、语言使用特点等方面存在一定差异, 国外数据集的研究结果在中国人群中的适用性不明确、相关研究理论不足, 仍需要更多的理论及实验支撑。中国本土化的抑郁筛查工具的发展是近年来备受关注的研究方向之一。本土化的抑郁症筛查工具的发展, 在充分考虑中国人文化背景和心理特点、构建更加精准的指标和方法的基础上, 对于提高应用于中国患者的抑郁筛查工具的有效性和准确性上具有重要的意义。

其次, 目前抑郁识别数据集样本量较少, 且部分数据集有较明显的样本分布不平衡现象, 即健康对照组的数据样本量明显多于抑郁患者的数量。非均衡样本对于抑郁识别的分类算法的准确性和有效性存在考验。

最后, 目前研究对提取特征和模型参数的解释性不足, 模型的输入输出和内部的工作原理和运行机制不透明, 多模态抑郁识别算法在临床上难以得到充分应用。相关研究实验结果及分析中, 不仅仅需要提升模型准确率等指标, 特征提取及模型构建的理论及实验分析也尤为重要。

4 结论

抑郁量表作为目前抑郁症状主要筛查方式, 主要根据典型症状出现的频次、严重程度及持续时长, 评估测评者是否有抑郁症状及抑郁程度, 具有良好的信效度。但传统测量工具可能受到医生主观判断和患者自我认知的影响, 客观性和便捷性不足。随着机器学习等技术的发展, 计算机辅助筛查方式为抑郁症状

筛查提供了新的解决方案。基于视觉、语音、文本等多模态特征,模型能够有效对抑郁症状进行分类分级,提供了一种相对快速、准确的识别方法。目前抑郁识别检测算法在准确率、召回率、F1 值等指标上已达到了较好的指标效果,但开放数据集样本量较少,且部分数据集的样本分布不均衡,为进一步提升算法精度带来了挑战。在抑郁识别工具的应用上,目前抑郁识别研究中模型特征和参数的可解释性不足,相关理论及实验研究对抑郁识别算法未来在临床上的落地应用尤为重要。

参考文献 (References)

- [1] 美国精神医学学会. 精神障碍诊断与统计手册(第五版)[M]. 张道龙, 刘春宇, 张小梅, 等译. 北京: 北京大学出版社, 2015: 345–350.
- [2] World Health Organization. Depression and other common mental disorders: Global health estimates[R]. World Health Organization, 2017.
- [3] Huang Y Q, Wang Y, Wang H, et al. Prevalence of mental disorders in China: A cross-sectional epidemiological study[J]. The Lancet Psychiatry, 2019, 6(3): 211–224.
- [4] Zhao D M, Wu Z G, Zhang H F, et al. Somatic symptoms vary in major depressive disorder in China[J]. Comprehensive Psychiatry, 2018, 87: 32–37.
- [5] 祁荣, 陈军, 余邵民. 关于抑郁症的研究综述[J]. 心理月刊, 2020, 15(17): 238–240.
- [6] Katon W, Kleinman A, Rosen G. Depression and somatization: A review part I[J]. The American Journal of Medicine, 1982, 72(1): 127–135.
- [7] 李聪慧, 杨世平, 阎凤芝. 首次住院抑郁症患者就诊途径调查分析[J]. 中国全科医学, 2006, 9(11): 919.
- [8] Craft L L, Landers D M. The effect of exercise on clinical depression and depression resulting from mental illness: A meta-analysis[J]. Journal of Sport and Exercise Psychology, 1998, 20(4): 339–357.
- [9] Liang Y J, Zheng X L, Zeng D D. A survey on big data-driven digital phenotyping of mental health[J]. Information Fusion, 2019, 52: 290–307.
- [10] Joshi J, Goecke R, Parker G, et al. Can body expressions contribute to automatic depression analysis? [C]//Proceedings of 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG). Piscataway, NJ: IEEE, 2013: 1–7.
- [11] 杜铭浩, 刘爽, 刘潇雅, 等. 面部视觉特征在抑郁症诊断中的研究进展[J]. 小型微型计算机系统, 2022, 43(3): 483–489.
- [12] Silva W J, Lopes L, Galdino M K C, et al. Voice acoustic parameters as predictors of depression[J]. Journal of Voice, 2024, 38(1): 77–85.
- [13] 黄观澜, 周晓璐. 抑郁症患者的语言使用模式[J]. 心理科学进展, 2021, 29(5): 838–848.
- [14] Edwards T, Holtzman N S. A meta-analysis of correlations between depression and first person singular pronoun use[J]. Journal of Research in Personality, 2017, 68: 63–68.
- [15] Robertson R, Robertson A, Jepson R, et al. Walking for depression or depressive symptoms: A systematic review and meta-analysis[J]. Mental Health and Physical Activity, 2012, 5(1): 66–75.
- [16] Li W T, Wang Q X, Liu X, et al. Simple action for depression detection: Using kinect-recorded human kinematic skeletal data[J]. BMC Psychiatry, 2021, 21(1): 205.
- [17] Beck A T, Steer R A, Brown G. Beck depression inventory II[J/OL]. Psychological Assessment, 1996, <https://www.sciencedirect.com/topics/medicine-and-dentistry/beck-depression-inventory>.
- [18] Kroenke K, Spitzer R L. The PHQ-9: A new depression diagnostic and severity measure[J]. Psychiatric Annals, 2002, 32(9): 509–515.
- [19] Hamilton M. The Hamilton rating scale for depression [M]//Assessment of Depression. Berlin, Heidelberg: Springer Berlin Heidelberg, 1986: 143–152.
- [20] Zung W W. A self-rating depression scale[J]. Archives of General Psychiatry, 1965, 12: 63–70.
- [21] Sundararajan V, Henderson T, Perry C, et al. New ICD-10 version of the Charlson comorbidity index predicted in-hospital mortality[J]. Journal of Clinical Epidemiology, 2004, 57(12): 1288–1294.
- [22] Valstar M, Schuller B, Smith K, et al. AVEC 2013: The continuous audio/visual emotion and depression recognition challenge[C]//Proceedings of the 3rd ACM international workshop on Audio/visual emotion challenge. New York: ACM, 2013: 3–10.
- [23] Valstar M, Schuller B, Smith K, et al. AVEC 2014: 3D dimensional affect and depression recognition challenge[C]//Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge. New York: ACM, 2014: 3–10.
- [24] Gratch J, Artstein R, Lucas G, et al. The distress analysis interview corpus of human and computer interviews[R]. California: University of Southern California Los Angeles, 2014.
- [25] Dibeklioglu H, Hammal Z, Cohn J F. Dynamic multimodal measurement of depression severity using deep autoencoding[J]. IEEE Journal of Biomedical and Health Informatics, 2018, 22(2): 525–536.
- [26] Cai H S, Yuan Z Q, Gao Y W, et al. A multi-modal open dataset for mental-disorder analysis[J]. Scientific Data, 2022, 9(1): 178.
- [27] Shen Y, Yang H Y, Lin L. Automatic depression detection: An emotional audio-textual corpus and A gru/bilstm-based

- model[C]//Proceedings of ICASSP 2022–2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ: IEEE, 2022: 6247–6251.
- [28] Zou B C, Han J L, Wang Y X, et al. Semi-structural interview-based Chinese multimodal depression corpus towards automatic preliminary screening of depressive disorders[J]. IEEE Transactions on Affective Computing, 2023, 14(4): 2823–2838.
- [29] Mao K N, Wang D B, Zheng T S, et al. Analysis of automated clinical depression diagnosis in a Chinese corpus[J]. IEEE Transactions on Biomedical Circuits and Systems, 2023, 17(5): 1135–1152.
- [30] Wang Q X, Yang H X, Yu Y H. Facial expression video analysis for depression detection in Chinese patients[J]. Journal of Visual Communication and Image Representation, 2018, 57: 228–233.
- [31] Wang Y M, Wang J Y, Liu X Q, et al. Detecting depression through gait data: Examining the contribution of gait features in recognizing depression[J]. Frontiers in Psychiatry, 2021, 12: 661213.
- [32] Haque A, Guo M, Miner A S, et al. Measuring depression symptom severity from spoken language and 3D facial expressions[EB/OL]. [2023–05–11]. <https://arxiv.org/abs/1811.08592v2>.
- [33] Friesen E, Ekman P. Facial action coding system: a technique for the measurement of facial movement[J]. Palo Alto, 1978, 3(2): 5.
- [34] Cohn J F, Krueze T S, Matthews I, et al. Detecting depression from facial actions and vocal prosody[C]//Proceedings of 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops. Piscataway, NJ: IEEE, 2009: 1–7.
- [35] Meng H Y, Huang D, Wang H, et al. Depression recognition based on dynamic facial and vocal expression features using partial least square regression[C]//Proceedings of the 3rd ACM international workshop on Audio/visual emotion challenge. New York: ACM, 2013: 21–30.
- [36] Pampouchidou A, Marias K, Tsiknakis M, et al. Designing a framework for assisting depression severity assessment from facial image analysis[C]//Proceedings of IEEE International Conference on Signal and Image Processing Applications (ICSIPA). Piscataway, NJ: IEEE, 2015: 578–583.
- [37] Zhu Y, Shang Y Y, Shao Z H, et al. Automated depression diagnosis based on deep networks to encode facial appearance and dynamics[J]. IEEE Transactions on Affective Computing, 2018, 9(4): 578–584.
- [38] de Melo W C, Granger E, Hadid A. Combining global and local convolutional 3D networks for detecting depression from facial expressions[C]//Proceedings of 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019). Piscataway, NJ: IEEE, 2019: 1–8.
- [39] 孙浩浩, 邵珠宏, 尚媛园, 等. 融合通道层注意力机制的多支路卷积网络抑郁症识别[J]. 中国图象图形学报, 2022, 27(11): 3292–3302.
- [40] Chen M Z, Xiao X, Zhang B, et al. Neural architecture searching for facial attributes-based depression recognition [C]//Proceedings of 26th International Conference on Pattern Recognition (ICPR). Piscataway, NJ: IEEE, 2022: 877–884.
- [41] Liu Z Y, Yuan X Y, Li Y T, et al. PRA-Net: Part-and-Relation Attention Network for depression recognition from facial expression[J]. Computers in Biology and Medicine, 2023, 157: 106589.
- [42] Joshi J, Goecke R, Alghowinem S, et al. Multimodal assistive technologies for depression diagnosis and monitoring[J]. Journal on Multimodal User Interfaces, 2013, 7(3): 217–228.
- [43] Ooi K E B, Lech M, Allen N B. Prediction of major depression in adolescents using an optimized multi-channel weighted speech classification system[J]. Biomedical Signal Processing and Control, 2014, 14: 228–239.
- [44] Kiss G, Vicsi K. Comparison of read and spontaneous speech in case of automatic detection of depression[C]//Proceedings of 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom). Piscataway, NJ: IEEE, 2017: 000213–000218.
- [45] He L, Cao C. Automated depression analysis using convolutional neural networks from speech[J]. Journal of Biomedical Informatics, 2018, 83: 103–111.
- [46] Zhang P Y, Wu M Y, Dinkel H, et al. DEPA: Self-supervised audio embedding for depression detection[EB/OL]. [2023–05–11]. <https://arxiv.org/abs/1910.13028v3>.
- [47] Sardari S, Nakisa B, Rastgoo M N, et al. Audio based depression detection using Convolutional Autoencoder[J]. Expert Systems with Applications, 2022, 189: 116076.
- [48] Du M H, Liu S, Wang T, et al. Depression recognition using a proposed speech chain model fusing speech production and perception features[J]. Journal of Affective Disorders, 2023, 323: 299–308.
- [49] Das A K, Naskar R. A deep learning model for depression detection based on MFCC and CNN generated spectrogram features[J]. Biomedical Signal Processing and Control, 2024, 90: 105898.
- [50] Williamson J R, Godoy E, Cha M, et al. Detecting depression using vocal, facial and semantic communication cues[C]//Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge. New York: ACM, 2016: 11–18.
- [51] Amanat A, Rizwan M, Javed A R, et al. Deep learning for depression detection from textual data[J]. Electronics, 2022, 11(5): 676.
- [52] Uddin M Z, Dysthe K K, Følstad A, et al. Deep learning for prediction of depressive symptoms in a large textual dataset[J]. Neural Computing and Applications, 2022, 34(1): 721–744.
- [53] Yadav U, Sharma A K. A novel automated depression detec-

- tion technique using text transcript[J]. International Journal of Imaging Systems and Technology, 2023, 33(1): 108–122.
- [54] Morales M, Scherer S, Levitan R. A cross-modal review of indicators for depression detection systems[C]//Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology—From Linguistic Signal to Clinical Reality. Stroudsburg, PA, USA: Association for Computational Linguistics, 2017: 1–12.
- [55] Ngiam J, Khosla A, Kim M, et al. Multimodal deep learning[C]//Proceedings of the 28th International Conference on Machine Learning. 2011: 689–696.
- [56] Xu K, Ba J, Kiros R, et al. Show, attend and tell: Neural image caption generation with visual attention[C]//Proceedings of the International Conference on Machine Learning. PMLR, 2015: 2048–2057.
- [57] de Vries H, Strub F, Mary J, et al. Modulating early visual processing by language[J]. Advances in Neural Information Processing Systems, 2017, 30: 1–14.
- [58] Wang H, Zhou Y, Yu F P, et al. Fusional recognition for depressive tendency with multi-modal feature[J]. IEEE Access, 2019, 7: 38702–38713.
- [59] Schultebrasucks K, Galatzer-Levy I. Computer vision and voice analysis for diagnostic assessment of PTSD, depression, and neurocognitive functioning[J]. Biological Psychiatry, 2021, 89(9): S19.
- [60] Liu J Q, Huang Y, Chai S R, et al. Computer-aided detection of depressive severity using multimodal behavioral data[M]//Intelligent Systems Reference Library. Cham: Springer International Publishing, 2021: 353–371.
- [61] Fang M, Peng S Y, Liang Y J, et al. A multimodal fusion model with multi-level attention mechanism for depression detection[J]. Biomedical Signal Processing and Control, 2023, 82: 104561.
- [62] Yasin S, Othmani A, Raza I, et al. Machine learning based approaches for clinical and non-clinical depression recognition and depression relapse prediction using audiovisual and EEG modalities: A comprehensive review[J]. Computers in Biology and Medicine, 2023, 159: 106741.
- [63] Zhang B T, Yan G H, Yang Z F, et al. Brain functional networks based on resting-state EEG data for major depressive disorder analysis and classification[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2021, 29: 215–229.
- [64] Xu Y T, Zhong H Y, Ying S Y, et al. Depressive disorder recognition based on frontal EEG signals and deep learning[J]. Sensors, 2023, 23(20): 8639.

Advanced research progress in computer-aided depression detection

GU Yating¹, ZHANG Chi¹, MA Fei^{2*}, JIA Xiaojian³, NI Shiguang^{1*}

1. Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China

2. Guangdong Laboratory of Artificial Intelligence and Digital Economy(SZ), Shenzhen 518107, China

3. Shenzhen Mental Health Center, Shenzhen Kangning Hospital, Shenzhen 518118, China

Abstract Major depressive disorder is a prevalent psychological disorder, and screening is currently based on depression diagnostic scales and physician interviews. Based on artificial intelligence technology, computer-aided depression detection is an emerging approach to depression screening. Aiming at the current status and shortcomings of traditional measurement tools, this paper reviews current computer-aided depression detection methods, discusses the current research status of depression detection datasets and depression detection methods based on multimodal data such as facial pictures, speech, and text, and summarizes and outlooks the advantages and challenges of computer-aided depression detection. Computer-aided depression detection can provide a relatively simple and standardized screening method with the potential to synergize with the widely used scale screening and physician diagnosis, but still faces the challenges of the insufficient interpretation of model parameters and features, the Chinese dataset to be expanded, and the small sample size of the existing dataset. In the future, researchers need to further improve the sample size and model accuracy of depression detection datasets, conduct theoretical and experimental analyses of feature extraction and model construction, and promote the clinical application of computer-aided depression detection.

Keywords major depressive disorder; depression recognition; screening scale; computer-aided detection; multimodal data ●



(责任编辑 徐丽娇)