

DOI:10.16305/j.1007-1334.2025.z20241008001

# 基于中医“司外揣内”法探讨人工智能诊断抑郁症的可行性

李红培<sup>1,2</sup>, 刘 龙<sup>3</sup>, 胡文悦<sup>1</sup>, 马华萍<sup>1,2</sup>, 钟碧莹<sup>1,2</sup>, 王淦宇<sup>1,2</sup>, 韩振蕴<sup>4</sup>

1. 北京中医药大学深圳医院(龙岗)脑病科(广东 深圳 518172); 2. 北京中医药大学第五临床医学院(深圳医院)(广东 深圳 518172); 3. 北京中医药大学深圳医院(龙岗)临床心理科(广东 深圳 518172); 4. 北京中医药大学东方医院脑病科(北京 100078)

**【摘要】** 作为一种具有高患病率、高致残率、高自杀率及高复发率的慢性精神疾病, 抑郁症的发病机制尚不明确, 也缺乏客观诊断指标, 确诊仍依赖精神专科医生一对一判断, 不仅耗时, 还需匹配大量的医疗资源。同时, 由于抑郁症早期症状的不典型以及患者在专科就医可能感到的“病耻感”, 导致患者更倾向于在非专科就诊, 容易造成病情延误。中医“司外揣内”法从整体出发, 通过望神(目)、面、舌和问不适、闻声音及切脉搏等多种方式与患者互动, 来推断其内在心理认知变化。这种诊断方法具备无创易操作、动态多维度、省时又经济的诊断优势, 更易获得患者配合。随着人工智能技术的快速发展, 通过医工结合的跨学科合作方式, 可以对抑郁症体现在外的信息进行标准化采集和客观化分析, 有望减少人工诊断带来的主观影响, 提高抑郁症早期筛查和诊断效率。

**【关键词】** 抑郁症; 人工智能; 司外揣内; 机器学习; 中医诊断

## Exploring feasibility of artificial intelligence in diagnosing depressive disorder based on traditional Chinese medicine principle of "governing exterior to infer interior"

LI Hongpei<sup>1,2</sup>, LIU Long<sup>3</sup>, HU Wenyue<sup>1</sup>, MA Huaping<sup>1,2</sup>, ZHONG Biying<sup>1,2</sup>, WANG Hanyu<sup>1,2</sup>, HAN Zhenyun<sup>4</sup>

1. Department of Encephalopathy, Beijing University of Chinese Medicine Shenzhen Hospital (Longgang), Shenzhen, Guangdong 518172, China; 2. Fifth Clinical Medical College, Beijing University of Chinese Medicine (Shenzhen Hospital), Shenzhen, Guangdong 518172, China; 3. Department of Clinical Psychology, Beijing University of Chinese Medicine Shenzhen Hospital (Longgang), Shenzhen, Guangdong 518172, China; 4. Department of Encephalopathy, Dongfang Hospital, Beijing University of Chinese Medicine, Beijing 100078, China

**Abstract:** Depressive disorder (DD) is a chronic psychiatric disorder characterized with high morbidity, disability, suicide rates, and recurrence, yet its pathogenesis remains unclear, and objective diagnostic indicators are lacking. Diagnosis confirmation still largely relies on one-on-one evaluations by psychiatric specialists, which is time-consuming and requires substantial medical resources. Furthermore, due to the atypical early symptoms of DD, coupled with the stigma associated with seeking psychiatric care, patients tend to consult non-specialists, which may increase the risk of delayed diagnosis and treatment. Traditional Chinese medicine (TCM) diagnostic principle of "governing exterior to infer interior" adopts a holistic approach, interacting with patients through inspecting spirit (eyes), facial expressions and tongue, inquiring about discomfort, listening to the voice, and taking the pulse to infer patients' internal psychological and cognitive changes. This diagnostic approach is non-invasive, easy to perform, dynamic, multi-dimensional, time-efficient, and cost-effective, making it more acceptable to patients. With the rapid advancement of

artificial intelligence, interdisciplinary collaborations between medicine and engineering can standardize the collection and objective analysis of outward manifestations of DD, potentially reducing subjective biases in manual diagnosis and improving the efficiency of early screening and diagnosis.

**Keywords:** depressive disorder; artificial intelligence; governing exterior to infer interior; machine learning; traditional Chinese medicine diagnosis

[基金项目] 国家重点研发计划项目(2019YFC1710103); 深圳市“医疗卫生三名工程”项目(SZZYSM202105010); 深圳市龙岗区医疗卫生技术攻关项目(LGKCYLWS2022012); 深圳市龙岗区医学重点学科建设项目(2024年至2027年)

[作者简介] 李红培, 女, 博士研究生, 主要从事中医脑病的临床与基础研究工作

[通信作者] 韩振蕴, 主任医师, 教授, 博士研究生导师;

E-mail: tohazhenyun@sina.com

抑郁症是一种具有高患病率、高致残率、高自杀率及高复发率的慢性精神疾病<sup>[1-3]</sup>。全球流行病学数据<sup>[2]</sup>显示,抑郁症的患病率约为5%。在我国,成人人群的终生患病率为3.4%,其中女性患病率普遍高于男性,且多数患者伴有不同程度的社会功能障碍<sup>[3]</sup>。自2010年以来,抑郁症已跃升为导致全球健康寿命损失的第二大疾病<sup>[4]</sup>。可见,对抑郁症进行精准识别与及时干预具有重要的公共卫生意义。然而,抑郁症早期症状更多表现为不典型的情绪波动或躯体不适,且由于发病机制尚未完全明确,也缺乏客观特异的生物标志物。故目前确诊抑郁症仍高度依赖精神专科医生对患者主观症状的综合判断,存在误诊与漏诊的可能,尤其是在基层或非精神专科场景中这种现象更为突出<sup>[3,5-6]</sup>。

中医“司外揣内”法<sup>[7]</sup>主张从宏观角度审视健康问题,可以通过四诊互动全面捕捉在外疾病信息,合参以精准辨别内在病机证候,从而为抑郁症患者制定个性化的论治方案。其中,望诊(尤其是望神)可以初步判断患者的精神情志状态,在深入问诊挖掘主观不适时,还能根据所闻语音变化直观感知其情绪波动,并结合脉象虚实进一步了解脏腑气血盈亏。相较于西医诊断拘泥于临床访谈和量表,不仅诊断视角单一,还需要匹配大量的专科资源,中医“司外揣内”的辨证方式具有无创易操作、动态多维度、省时又经济的优势,并能在一定程度上削弱患者对精神心理疾病的“病耻感”,拥有更高的就医和治疗配合度。但传统中医同样存在经验为主导而难以实现客观记录与重复验证的问题,导致诊断结果因医生水平差异而一致性不足。

近年来,人工智能,尤其是深度学习技术,在医学图像特征识别、语音与脉搏可视化分析等领域取得了突破性进展,可以对抑郁症患者的望诊(如眼动、面部表情和舌图)、闻诊与脉诊等显露在外的非结构化客观信息进行深度挖掘与提取<sup>[8-13]</sup>。与此同时,机器学习的非线性建模能力与中医“司外揣内”以诊病辨证的整体观念不谋而合,通过多模态融合,有望复现中医四诊合参的推理过程和结果<sup>[8,14-17]</sup>。故本文基于中医“司外揣内”法探讨人工智能在抑郁症诊断中的可行性,以期为后续研究及临床应用提供新的思路 and 依据。

## 1 “司外揣内”与抑郁症

1.1 抑郁症的疾病特点和诊断挑战 严重的抑郁症表现为典型的“三低”(情绪低落、思维减慢、行动减少)、“三无”(无望、无助、无用)及“三自”(自责、自罪、自杀)等。然而,抑郁症早期的临床表现往往并不典型,通常会以睡眠障碍、食欲紊乱或躯体不适等为主。不同年龄层的首发症状也不尽相同,如儿童和青少年可能表现为头痛、胃痛等躯体不适,或易激惹、注意力不集中,

或学习动力不足等<sup>[5]</sup>;高龄患者则以记忆力下降等认知功能改变多见<sup>[6]</sup>。首发症状的不典型性容易造成患者和家属对病情的不重视,再加上精神心理专科就诊可能带来的“病耻感”,使得患者更倾向于选择非专科就诊,这些因素均可导致诊断时机延误,进而加重病情。

目前抑郁症的诊断尚未有明确且特异的客观生物标志物,确诊仍须同时满足症状、病程、严重程度及排除标准<sup>[3]</sup>。然而,全面采集这些信息需要精神专科医生对患者进行“一对一”面诊和深入访谈<sup>[3]</sup>,这不仅耗时,也需大量的专科医疗资源支持。虽然量表测评有助于早期筛查和诊断,但其准确性比较容易受到患者主观感受以及评估者主观判断的影响。尽管已有研究<sup>[18-19]</sup>发现,抑郁症患者体内普遍存在代谢组学和蛋白质组学水平的变化,但这种变化并非在所有个体中都一致。另外,抑郁症患者在免疫炎症指标、神经影像学指标、神经电生理学指标等方面也呈现出了不同程度的异常表现<sup>[20]</sup>,但这些指标在临床中也因特异性不强、操作复杂、成本高昂等情况而应用受限。因此,如何在充分保证患者配合度的前提下,高效、全面获取抑郁症临床资料,从而精准识别和诊断是首要难题。

1.2 “司外揣内”的科学内涵和诊断优势 《灵枢·本脏》载“有诸于内,必形于外”,指出内在的病理变化必然在外部有所体现。“司外揣内”正是中医通过观察患者显露在外的各种征象,来推测其内在病变的一种独特诊断手段。《素问·阴阳应象大论》载“善诊者,察色按脉,先别阴阳;审清浊,而知部分;视喘息、听音声,而知所苦;观权衡规矩,而知病所主;按尺寸,观浮沉滑涩,而知病所生”,详细阐明了中医诊病的核心思想,即通过观察面色、脉象、呼吸、音声等外在信息,综合分析患者的内在病理变化,以判断病因与病位。《医宗金鉴》道:“望以目察,闻以耳占,问以言审,切以指参,明斯诊道,识病根源,能合色脉,可以万全。”这进一步强调了中医司外的具体手段为望、闻、问、切,只有通过这4种手段,医生才能够比较全面地获取患者的症状和体征,从而精准揣内。

1.2.1 望诊抑郁症 望诊能从神(表情)、色、形、态等多方面获取疾病信息,为诊断抑郁症提供有力证据。其中,面部表情是表达和传递情感最直接的方式之一,可以反映个体的内心体验和认知变化,是初步评估患者精神心理状态的重要线索。随着面诊的逐步深入,医生可以动态观察到患者神的微妙变化,进而追踪其情绪波动。在情绪失控或者陷入负面思维模式时,可见抑郁症患者出现垂目、皱眉、嘴角下拉等面部表情特征<sup>[21]</sup>。而望神之精要,又尽在望目中。正如《灵枢·大惑》有云:“目者,五脏六腑之精也,营卫魂魄之所常营也,神气之所生也。”心理状态良好的个体通常目光柔

和、眼动灵活、注意力集中、思维敏捷,能与医生进行自然有效的眼神交流。而抑郁症患者往往会表现出眼神空洞和忧郁、目光飘忽、注意力分散、思维反应迟缓,病情严重时,还可能回避眼神接触,眼动也会减慢<sup>[22]</sup>。现代眼动追踪技术观察到,和健康对照组相比,抑郁症患者确实存在多种异常眼动表现,如扫视幅度更大和无意识眼动频繁、眼跳幅度减少和速度减慢、扫视路径缩短、瞳孔大小的动态范围及大小比率降低等<sup>[23]</sup>。此外,情绪波动还会在面色上有所体现。如《灵枢·口问》载:“悲哀愁忧则心动,心动则五脏六腑皆摇。”《难经》指出:“五脏有五色,皆见于面。”当人们感到悲哀或愁忧时,心神会被扰动,心神的不安会导致五脏六腑功能失调,进而影响五志化生七情,同时引起面部色泽和状态的变化。这一现象的发生可能在于情绪起伏触发了身体的应激反应,继而激活自主神经系统,导致血管扩张或收缩,造成面部血流量和温度改变<sup>[24]</sup>。例如,愤怒或紧张可能引起面部泛红,而恐惧或震惊可能导致脸色苍白。《医碥》曰:“郁则不舒,则皆肝木之病矣。”抑郁症的发生归咎于情志内伤,核心病位在肝。长期郁怒、忧思、恐惧等情志刺激可致肝失条达、气机不畅,而成肝郁,面部还会出现“蒙尘脱色”状,即面部总体 Lab-L 值(明度)及 a 值、b 值(色度)均降低<sup>[25]</sup>。当病久累及脾胃、心、肾,还会进一步产生湿、痰、热、瘀等病理产物,这些病理变化则会在舌象上得到具体呈现。如抑郁症患者肝郁日久最易克犯脾土,致脾虚失运,气血生化乏源,水谷精微上输困难,舌体失养,而见舌淡<sup>[26]</sup>,当脾失健运,聚湿成痰,又可见苔腻之象<sup>[27]</sup>。

1.2.2 问诊抑郁症 《黄帝内经》中多次强调“未诊先问,最为有准”。如《素问·疏五过论》中提到:“凡欲诊病者,必问饮食居处,暴乐暴苦,皆伤精气。”《素问·微四失论》言:“诊病不问其始,忧患饮食之失节,起居之过度,或伤于毒,不先言此,卒持寸口(切脉),何病能中?”对抑郁症等情志病而言,更应谨守问诊为“诊病之要领,临证之首务”的原则。《丹溪心法》言:“一有怫郁,诸病生焉。”抑郁症的症状表现十分广泛,往往涉及全身多个系统,如《赤水玄珠》言“心郁者,神气昏昧,心胸微闷,主事健忘”“肝郁者,两胁微膨,暖气连连有声”“脾郁者,中脘微满,生涎少食,四肢无力”“肺郁者,皮毛燥而不润,欲嗽而无痰”“肾郁者,小腹微硬,精髓乏少,或浊或淋,不能久立”。除了核心的情感症状外,抑郁症还多伴见自主神经系统症状,且不同抑郁人群发病特征也存在差异。如《金匱要略》载:“妇人脏躁,喜悲泣欲哭,象如神灵所作,数欠伸……妇人咽中如有炙脔。”只有细致问诊,才能全面了解抑郁症患者既往的发病经过和当下的症状表现。

1.2.3 闻诊抑郁症 在自然状态下获取的语音信息对

抑郁症诊断同样有重要意义。中医通过“视喘息,听声音”可以判断患者“所苦”。而声音特性(如声调、音高、强度)和语音特征(如词汇、语法)又能够即时反映说话人的精神心理状态,《金匱要略》中有百合病“意欲食复不能食,常默默”的记载,指出情志抑郁之人往往不喜言谈、不善言辞。现代声学研究<sup>[28]</sup>也发现,抑郁症患者多有音高降低、音域变窄、语速减慢等异常表现。即使在不同的语音情境中,抑郁症患者说话也常以低沉无力为主,且少有抑扬顿挫。某些声学参数,如基频(F0)范围和均值、Mel 频率倒谱系数(MFCCs)、频率微扰和振幅微扰等,已被证实与患者的抑郁程度存在一定的相关性<sup>[29-31]</sup>。此外,抑郁情绪也可能影响发声肌肉的张力和声道的几何形状等,这些声道物理特性的改变会导致说话人的语音音色和元音质量的变化,可表现为共振峰频率的协调性较差<sup>[30]</sup>。

1.2.4 切诊抑郁症 作为一种相对安全的触摸手段,切诊摸脉能够以非语言的形式实施情感安抚,有助于医生与抑郁症患者建立初步信任和情感联结,而且抑郁症患者在脉象上也确有异常表现。古有文献记载,如《景岳全书》言:“凡郁证之脉,在古人皆以结促止节为郁脉……惟情志之郁,则如弦紧、沉涩、迟细、短数之类皆能为之。”又《张氏医通》言:“郁脉多沉伏,或结或促,或沉或涩,郁在肝肾则见于左,郁在心脾则见于右。”现代研究<sup>[32]</sup>证实,不同证型抑郁症患者脉象图表现与传统中医脏腑辨证理论基本一致,如肝郁脾虚型抑郁症患者表现为左手关部脉(候肝胆)波幅增大,甚至一脉独大,且负向波加深,甚至大于正波波幅,右手关部脉(候脾胃)则波幅偏小;心脾两虚型抑郁症患者则表现为左手寸部脉(候心)波幅小,基线不平稳,右手关部脉波幅也小。

由此可见,抑郁症在望、闻、问、切方面均存在不同程度的特征变化。在基于中医“司外揣内”法诊断抑郁症的过程中,望诊首当其冲并贯穿始终,问诊必不可少且占据主导,闻诊、切诊随之自然展开并起到辅助作用。医生通过与患者细致沟通和互动,关心其身心状态,综合挖掘其在“外”信息,合参四诊以窥其“内”在的心理认知变化,在保证诊断全面性的同时,充分展现了中医独特的亲和力和人文性,有助于缓解抑郁症患者的就医顾虑,减少他们对精神心理疾病的“病耻感”,并可以在一定程度上提高就医配合度和治疗依从性。

## 2 人工智能与抑郁症

2.1 四诊信息的标准采集和客观量化 人眼望诊难免受到环境光线和主观判断的影响。时下成熟的舌、面信息采集系统<sup>[25-26]</sup>可以在配置有高清摄像镜头、“暗箱”和辅助光源的统一环境中,进行面图、舌图的标准化的采

集,再通过动态轮廓(Snake)模型、随机森林(RF)、卷积神经网络(CNN)、K-means 聚类等算法,结合小波变换、Open CV 等技术,进行图像预处理(如色彩校正、降噪锐化、图像分割)和特征识别(包括面部特征、舌象特征),进而实现面部信息和舌象信息的数字化提取。现代眼动追踪技术通常应用红外摄像镜头来精准捕捉和记录眼球运动,通过进一步测量眼球运动的速度、持续时间和轨迹等关键参数,可以实现对“眼神”的客观量化,并揭示情绪变化对注意力分配的影响,以帮助研究者间接了解个体在特定情境下的认知过程和情绪状态<sup>[23]</sup>。

口头问诊基于耳闻口述展开,最易受到医患双方语言习惯、个人经验和认知模式等主观影响。虽然联合使用心理学量表、问卷等可以对患者的症状、病情进行系统评估和量化,但自评量表的结果和质量极易受到患者当下情绪状态的干扰,而他评量表又存在需医生引导而导致耗时较长的缺点。随着“互联网+医疗”的发展,集成大数据分析、自然语言处理(NLP)<sup>[33]</sup>、多层次深度强化学习与主动式学习、多维度推理引擎以及多专科医学知识图谱等先进人工智能技术的诊前预问诊系统<sup>[34]</sup>在一定程度上能够规范化、标准化模拟医生的问诊过程。在候诊期间,借助智能手机、门诊问诊机器人等即可与患者进行人机交互和对话<sup>[35]</sup>,记录其一般情况,并围绕其主诉展开疾病发生发展、就医经过等相关信息的结构化问询,最终生成初步电子病历,帮助医生在有限的面诊时间内更有针对性地进行诊疗。

语音信息是一组涉及频谱、韵律、频率、振幅、时长等多个物理元素的高维数据,普通人耳无法对其复杂特征做到直接提取、定量识别及重复验证。目前高质量声学样本的获取,除需配备专业的录音设备外,还需满足无噪音、无回声的环境条件。而智能声诊的采样硬件相对简单,故通常被作为组件整合在中医四诊仪中<sup>[36]</sup>。考虑到含有实意的语句研究具有强社会性,客观化分析难度较大,为了准确获取到语音中相对客观且容易量化的物理元素,一般需要将环境音控制在 30 分贝或 45 分贝以下<sup>[36]</sup>,并采取指定元音与字句的方式引导受试者发声<sup>[37]</sup>。常用的声音信号处理和特征提取手段包括降噪、滤波、变换及近似熵(ApEn)、小波包能量熵(WPEE)等。经这些处理后的声音信号会以某时域上的一系列波形图的形式得到可视化呈现,以便科研分析与临床理解。

传统脉诊“在心易了,指下难明”,全赖医生的抽象感知。将传统脉诊与高精度传感器(如压力式传感器、光电式传感器、传声式传感器、超声多普勒式传感器、网格图像法传感器等)、大数据分析、人工智能算法等现代科技融合,同时借助智能设备强大的数据存储能力,可以实现患者脉象数据的客观保存和长期追踪。

在脉象采集设备方面,已从最初的大体积逐渐发展为如今的便携式脉诊仪<sup>[38]</sup>和可穿戴智能手环<sup>[39]</sup>等,其内置的柔性压力传感器可以实时监测人体脉搏波变化,并将压力信号转化为电信号,以便计算机进行识别处理<sup>[40]</sup>。目前的脉诊智能设备还普遍配备了更高精度的压力感受器,其工作原理模拟了医生“举、按、寻”的切脉手法,已经能够做到同步采集双手桡动脉“寸、关、尺”三部在“浮、中、沉”不同深度下的搏动状态和变化;并通过时域分析、频域分析、时频分析和非线性动力学等方法对采集到的脉象关键特征,如位(位置)、数(频率)、形(形态)、势(力度)进行数字化处理,最终以可视化的形式直观展示脉搏变化。<sup>[41-42]</sup>

2.2 抑郁症疾病特征的深度挖掘和模式识别 作为人工智能的核心,无需规则设定,机器学习就能够从庞大繁杂的数据中自主学习,获取知识或经验,随着数据的积累和算法的迭代,模型性能还会不断得到优化。传统机器学习在小样本数据训练和减少过拟合方面表现良好。在疾病诊断中,常用的数据挖掘算法<sup>[8, 14]</sup>包括 Logistic 回归、RF、支持向量机(SVM)、k 近邻(k-NN)、多层感知器(MLP)、BP 神经网络、决策树(DT)、主成分分析(PCA)、独立成分分析(ICA)、朴素贝叶斯(NB)、聚类算法等。这些机器学习算法可以从高维多源的症状信息中挖掘疾病规律和隐性关联,以辅助诊断。特别是基于 MLP 演变而成的深度学习,能够模拟人脑的神经网络结构,无需人工干预,就能对输入数据进行逐层特征提取,其强大的学习能力在处理如视觉、语音及脉搏等复杂数据方面性能突出<sup>[8]</sup>。如具有强鲁棒性和可迁移性的 CNN 尤其擅长自适应学习和提取图片中抑郁相关的面部表情和舌象形态学等深层特征<sup>[9-11]</sup>,还能有效挖掘脉象的时域、频域信息<sup>[8]</sup>,并具有较强的泛化能力;长短期记忆网络(LSTM)或时序卷积网络(TCN)等<sup>[12]</sup>,则可以有效提取抑郁症患者面部特征点的时间序列数据,捕捉面部微表情的动态特征;基于自适应注意力机制的迁移学习框架构建的多特征融合分类网络模型,能够有效融合舌象深层特征(如剥苔、点刺、裂纹)和浅层特征(如边缘、纹理),并对这些特征进行精准定位<sup>[13]</sup>。

近年来,针对抑郁症患者在眼动、面部表情、语言文本及语音等方面的异常外在表现,许多研究致力于开发抑郁症早期识别和预测模型。Zhang 等<sup>[23]</sup>基于抑郁症患者在注视稳定性范式和自由观看范式等任务中的多种异常眼动表现,分别采用 SVM、NB 和二次判别分析(QDA)算法建立抑郁症预测模型,准确率均超过 80%。Gavrilescu 等<sup>[43]</sup>基于面部动作编码系统(FACS)采集的面部动作单元(AUs)数据,构建了一个 3 层架构的非侵入式预测系统。该系统第 1 层通过活动外观

模型提取面部特征点,并使用多类SVM进行AUs分类,第2层建立AUs强度矩阵,第3层应用前馈神经网络进行模式识别,以预测患者的抑郁、焦虑和压力水平,准确率分别为87.2%、77.9%和90.2%。刘豪等<sup>[44]</sup>从NLP技术开展研究,基于抑郁症患者的文本语料,构建并完善了融合字词特征的中文抑郁症领域情感词典,并由此开发了可以有效处理中文抑郁症文本分类任务的BERT-W模型,该模型对抑郁倾向文本的识别精确率高达97.6%。Chiong等<sup>[45]</sup>基于Twitter上公开标注的抑郁文本,构建了包括SVM、NB、k-NN、DT、梯度提升决策树(GBDT)等多种分类模型,并分别在不同社交媒体平台上对模型进行验证;结果表明,这些模型均能有效检测和识别抑郁相关文本。Shin等<sup>[46]</sup>基于不同程度抑郁人群的F0、共振峰、频率微扰和振幅微扰等21项语音特征值,应用SVM、RF、MLP构建分类模型,其中,MLP表现最佳(灵敏度为65.6%,特异性为66.2%)。

因为单模态数据承载的抑郁症疾病信息相对有限,也有学者从多模态着手进行研究,丰富了“司外揣内”智能识别抑郁症的应用范围,提高了模型识别精度。Zhao等<sup>[15]</sup>应用正交匹配追踪算法及多模态融合算法构建了集面部表情和语音特征为一体的抑郁症识别模型,可以达到81.1%的识别率,相较于语音单模态,识别率提高了6.8%。考虑到访谈语音中的主观文本内容与客观音频信息之间的强互补效应,赵小明等<sup>[16]</sup>提出了一种基于跨模态特征重构与解耦网络(CFRDN)的多模态抑郁症检测方法,该方法通过“文本引导、重构、解耦与融合”策略,引入双向交叉注意力和Transformer融合模块,实现了文本与音频的深度交互和融合,并在DAIC-WoZ与E-DAIC两个公开的抑郁症多模态数据集上进行性能验证;结果表明,该方法在分类任务中的F1值高达0.90、召回率为0.93、精确率为0.87,在回归任务中的一致性相关系数为0.665,均方根误差为4.41,平均绝对误差为3.92,表现出较优的准确性与鲁棒性。王姝予等<sup>[17]</sup>分别应用LSTM结合注意力机制提取视觉脉诊信号,轻量级网络MobileNet V3 + CA(通道注意力)机制提取舌诊特征,再基于GBDT、Logistic回归、SVM、RF、k-NN五种机器学习算法,构建融合患者症状信息的抑郁症多模态辨证模型;结果表明,GBDT模型在“无填充、无采样、无筛选”条件下准确率最高(90.32%),该研究进一步为“司外揣内”理论指导下的抑郁症智能辨证和辅助诊断提供了可靠的证据支持。

2.3 人工智能诊断抑郁症的优势与不足 人工智能的发展,为抑郁症客观化、标准化诊断提供了契机。在中医“司外揣内”思想指导下,人工智能能够将传统的主观诊病逻辑转变为可量化、可验证的数字化推理过程。借助高清摄像头<sup>[43]</sup>、眼动仪<sup>[23]</sup>、录音笔<sup>[46]</sup>等现代专业设

备,以及舌、面、脉、闻等智能诊断仪器<sup>[25-26, 38-39]</sup>,可以对抑郁症患者的中医多维临床数据(如眼动、面部表情、舌象、脉象、语音等)进行标准化采集与可视化处理,进而实现从抽象的中医主观描述到具象的现代客观数据之间的转换。此外,人工智能的无规则自主学习能力正好与“司外揣内”诊病辨证的全局推理思路高度契合。基于患者在望<sup>[22-23, 25-27]</sup>、问<sup>[44-45]</sup>、闻<sup>[29-31]</sup>、切<sup>[32]</sup>诊方面的特征表现,应用机器学习与深度学习算法可以综合挖掘疾病的潜在规律和模式,并辅助医生发现容易被忽视的细微症状和前驱表现,从而提高抑郁症的早期检出率,以便及时实施干预<sup>[15, 23, 46]</sup>。而通过多模态策略全面整合患者外在的疾病信号,不仅能够有效降低单纯症状描述的主观影响,还能进一步提高对患者内在情志变化的识别精度<sup>[15, 17]</sup>。

尽管人工智能在抑郁症识别和诊断领域表现出了巨大的应用潜力,但完全实现临床推广仍面临不少挑战,具体如下。①部分数据隐私等级高:抑郁症属于精神心理疾病,采集数据时要严格遵守伦理规范,保护患者隐私,尤其是面部信息和语音私密性极强,需与患者积极沟通,充分尊重患者,避免二次伤害影响病情。②多维数据获取难度大:抑郁症症状多样,涉及情绪、行为和生理多个层面,不同群体间异质性较大,单一诊法特异性不强,不同诊法之间具备互补性,故需要进行多维数据的完整采集;其中又以问诊内容为诊断核心,相较于望、闻、切诊,其交互性更强,耗时较长,开展智能问诊可尝试将NLP技术和心理问卷、症状条目深度融合。③算法参差且模态单一:目前大多数研究还处于小样本、小范围、单模态探索阶段,基于同源数据进行不同模型性能比较的研究也较少,导致模型的泛化能力较低;下一步还需扩大样本人群,在外部验证模型,促进算法优化,或在小样本研究中尝试开展多模态研究及纵向跟踪采集。④单独应用患者配合度低:人工智能并不具备亲和力和人文性,无法完全取代人类医生展开心理层面的深度交互;若单独使用人工智能进行诊断,患者可能会对结果存疑,故应用过程中还需专业人员引导,诊断结果也需专业人类医生把关。

### 3 结语

中医“司外揣内”法从整体出发,通过望神(目)、望面、望舌和问不适(口头问诊)、闻声音及切脉搏等与患者多方面互动,四诊合参,综合挖掘抑郁症在“外”信息,进而探知其“内”在的心理认知变化,在获取完整症状资料的同时,拥有更高的患者配合度,有助于抑郁症的动态跟踪和个性化干预。人工智能的介入,不仅可以帮助实现对抑郁症四诊信息的标准化采集与客观化提取<sup>[23, 25-26, 35-39, 41-42]</sup>,尽可能降低由人工诊断带来的主

观影响,也有助于前驱症状的精准识别与隐性关联的深度挖掘,从而提高抑郁症的早期筛查和诊断效率。相对来说,传统机器学习技术在处理结构化数据方面表现稳健<sup>[8, 14]</sup>,而深度学习算法更擅长处理图像、语音、脉搏及自然语言文本等非结构化数据<sup>[8-12, 33, 39]</sup>。通过多模态数据融合方法<sup>[15-17]</sup>,智能识别和辅助诊断模型的稳定性与泛化性能又将得到进一步增强。

综上,将“司外揣内”法与人工智能、大数据分析等现代信息技术相结合,为构建具有中医特色的抑郁症智能识别和诊断体系提供了切实可行的理论框架与技术支持,并有望推动情志类疾病诊断范式由“经验辨证”迈向“数据驱动”。然而,尽管人工智能在抑郁症早期识别和精准诊断中展现了应用价值,但其真正临床推广仍面临如数据隐私、算法优化、多维数据全面采集与深度融合等挑战。更重要的是,人工智能不能完全替代人类医生在心理层面的互动。因此,在追求精准识别和高效诊断的同时,仍需把控好人工智能在实际应用中的伦理边界,即在医工有机结合之下还是应以临床医生为主、人工智能为辅,以确保抑郁症智能化诊断的科学严谨与人文温度。

#### 参考文献:

- [1] RIDLEY M, RAO G, SCHILBACH F, et al. Poverty, depression, and anxiety: causal evidence and mechanisms [J]. *Science*, 2020, 370(6522): eaay0214.
- [2] HERRMAN H, PATEL V, KIELING C, et al. Time for united action on depression: a lancet-world psychiatric association commission [J]. *Lancet*, 2022, 399(10328): 957-1022.
- [3] 中华医学会行为医学分会, 中华医学会行为医学分会认知应对治疗学组. 抑郁症治疗与管理的专家推荐意见(2022年) [J]. *中华行为医学与脑科学杂志*, 2023, 32(3): 193-202.
- [4] LU J, XU X F, HUANG Y Q, et al. Prevalence of depressive disorders and treatment in China: a cross-sectional epidemiological study [J]. *Lancet Psychiatry*, 2021, 8(11): 981-990.
- [5] RAO W W, XU D D, CAO X L, et al. Prevalence of depressive symptoms in children and adolescents in China: a meta-analysis of observational studies [J]. *Psychiatry Res*, 2019, 272: 790-796.
- [6] BABA H, KITO S, NUKARIYA K, et al. Guidelines for diagnosis and treatment of depression in older adults: a report from the Japanese society of mood disorders [J]. *Psychiatry Clin Neurosci*, 2022, 76(6): 222-234.
- [7] 乔世杰, 王章林, 付超, 等. 中医诊断基本原理中的象思维运用 [J]. *中医学报*, 2022, 37(11): 2261-2264.
- [8] 罗思言, 王心舟, 饶向荣. 人工智能在中医诊断中的应用进展 [J]. *中国医学物理学杂志*, 2022, 39(5): 647-654.
- [9] PAMPOUCHIDOU A, PEDIADITIS M, KAZANTZAKI E, et al. Automated facial video-based recognition of depression and anxiety symptom severity: cross-corpus validation [J]. *Mach Vis Appl*, 2020, 31(4): 1-19.
- [10] HE L, CHAN J C W, WANG Z M. Automatic depression recognition using CNN with attention mechanism from videos [J]. *Neurocomputing*, 2021, 422: 165-175.
- [11] WANG X, LIU J W, WU C Y, et al. Artificial intelligence in tongue diagnosis: using deep convolutional neural network for recognizing unhealthy tongue with tooth-mark [J]. *Comput Struct Biotechnol J*, 2020, 18: 973-980.
- [12] YUAN X Y, LIU Z Y, CHEN Q Q, et al. Combining informative regions and clips for detecting depression from facial expressions [J]. *Cognit Comput*, 2023, 15(6): 1961-1972.
- [13] 徐雍钦, 杨晶东, 姜泉, 等. 基于多特征融合的中医症候舌象分类方法研究 [J]. *智能计算机与应用*, 2022, 12(7): 25-34.
- [14] 崔骥, 许家伦. 人工智能信息技术在中医四诊现代化研究中的应用现状与展望 [J]. *上海中医药杂志*, 2025, 59(1): 7-12.
- [15] ZHAO J, SU W, JIA J, et al. Research on depression detection algorithm combine acoustic rhythm with sparse face recognition [J]. *Cluster Comput*, 2019, 22: 7873-7884.
- [16] 赵小明, 谌自强, 张石清. 基于跨模态特征重构与解耦网络的多模态抑郁症检测方法 [J]. *计算机应用研究*, 2025, 42(1): 236-241.
- [17] 王姝予. 基于机器视觉的舌脉客观化技术构建抑郁症辨证模型 [D]. 成都: 成都中医药大学, 2024.
- [18] SILVA-COSTA L C, CARLSON P T, GUEST P C, et al. Proteomic markers for depression [J]. *Adv Exp Med Biol*, 2019, 1118: 191-206.
- [19] 刘振, 周宁, 刘通, 等. 基于代谢组学的抑郁症研究进展 [J]. *中国病理生理杂志*, 2020, 36(12): 2264-2275.
- [20] 祁娜, 冯媛, 王刚. 抑郁症客观评估方法的研究进展 [J]. *神经疾病与精神卫生*, 2020, 20(5): 341-346.
- [21] DEMCHENKO I, DESAI N, IWASA S N, et al. Manipulating facial musculature with functional electrical stimulation as an intervention for major depressive disorder: a focused search of literature for a proposal [J]. *J Neuroeng Rehabil*, 2023, 20(1): 64.
- [22] ZHOU X Z, JIN K, SHANG Y Y, et al. Visually interpretable representation learning for depression recognition from facial images [J]. *IEEE Trans Affect Comput*, 2020, 11(3): 542-552.
- [23] ZHANG D, LIU X, XU L H, et al. Effective differentiation between depressed patients and controls using discriminative eye movement features [J]. *J Affect Disord*, 2022, 307: 237-243.
- [24] GIOIA F, NARDELLI M, SCILINGO E P, et al. Autonomic regulation of facial temperature during stress: a cross-mapping analysis [J]. *Sensors (Basel)*, 2023, 23(14): 6403.
- [25] 杨娜婷, 周旋, 马庆宇, 等. 抑郁症肝郁脾虚证患者面诊客观参数分析 [J]. *中华中医药杂志*, 2023, 38(8): 3836-3839.
- [26] 杨娜婷, 周旋, 马庆宇, 等. 抑郁症肝郁脾虚证舌诊客观参数分析及应用探索 [J]. *中华中医药杂志*, 2023, 38(3): 1269-1273.
- [27] 刘素彤, 张丽慧, 顾亚娇, 等. 赵文霞教授基于“肝胃不和-舌象变化”诊治功能性消化不良伴抑郁状态的经验 [J]. *中国中西医结合消化杂志*, 2023, 31(1): 13-15.
- [28] FU H L, ZHUANG Z H, WANG Y, et al. Cross-corpus speech emotion recognition based on multi-task learning and subdomain adaptation [J]. *Entropy (Basel)*, 2023, 25(1): 124.
- [29] WANG J Y, ZHANG L, LIU T L, et al. Acoustic differences between healthy and depressed people: a cross-situation study [J]. *BMC Psychiatry*, 2019, 19(1): 300.

(下转第74页)