

• 专题研究:神经精神疾病 •

从数据整合到临床转化:机器学习在抑郁症诊断中的应用

高子祥¹, 李翔¹, 张欣¹, 钱昭均¹, 嵇子慧¹, 朱俐璇¹, 徐忆初², 陈亚丽^{1*}, 葛菲菲^{1*}

¹南京中医药大学医学院, ²人工智能与信息技术学院, 江苏 南京 210023

[摘要] 抑郁症起病隐匿、症状异质性明显, 现行诊断仍主要依赖临床访谈与量表评估, 存在主观性较强、早期识别不足以及对分型和预后判断能力有限等问题。随着神经影像、脑电图、语音与数字行为、临床量表及多组学等客观数据的不断积累, 机器学习为抑郁症客观识别提供了新的研究路径。文章围绕抑郁症客观识别这一核心问题, 综述了机器学习在神经影像及其他单模态客观数据中的应用进展, 进一步总结了多模态数据在抑郁症诊断、分型、病程评估和疗效预测中的整合价值, 并对常见融合策略、模型可解释性及临床转化问题进行了归纳分析。

[关键词] 抑郁症; 机器学习; 神经影像学; 多模态数据; 辅助诊断

[中图分类号] R749.4

[文献标志码] A

[文章编号] 1007-4368(2026)07-1011-09

doi: 10.7655/NYDXBNSN260333

From data integration to clinical translation: a review of machine learning applications in depression diagnosis

GAO Zixiang¹, LI Xiang¹, ZHANG Xin¹, QIAN Zhaojun¹, JI Zihui¹, ZHU Lixuan¹, XU Yichu², CHEN Yali^{1*}, GE Feifei^{1*}

¹School of Medicine, ²School of Artificial Intelligence and Information Technology, Nanjing University of Chinese Medicine, Nanjing 210023, China

[Abstract] Depression is characterized by an insidious onset and marked symptom heterogeneity. Current diagnostic practice still mainly relies predominantly on clinical interviews and rating scales, which are constrained by substantial subjectivity, limited capacity for early identification, and insufficient performance in subtype differentiation and prognostic evaluation. With the growing availability of objective data, including neuroimaging, electroencephalography, speech and digital behavioral features, clinical scales, and multi-omics profiles, machine learning has opened new avenues for the objective identification of depression. Focusing on this central issue, the present review summarizes recent advances in the application of machine learning to neuroimaging and other unimodal objective data, further discusses the integrative value of multimodal data in the diagnosis, subtyping, disease-course assessment, and treatment-response prediction of depression, and provides a systematic overview of common fusion strategies, model interpretability, and issues related to clinical translation.

[Key words] depression; machine learning; neuroimaging; multimodal data; auxiliary diagnosis

[J Nanjing Med Univ, 2026, 46(07): 1011-1019]

[基金项目] 国家自然科学基金(82574736); 国家级大学生创新训练计划项目(202510315005)

*通信作者 (Corresponding author), E-mail: ffg@njucm.edu.cn (ORCID: 0000-0003-1336-8587); ylchen@njucm.edu.cn (ORCID: 0009-0004-8738-1666)

1 抑郁症诊断现状及机器学习在单模态数据中的应用

抑郁症是一种常见的慢性复发性精神障碍, 起病隐匿、病程迁延, 复发率与致残率居高不下, 严重影响患者社会功能并加重公共卫生负担^[1]。目前临床诊断主要依据精神障碍诊断操作性标准, 结合汉

密尔顿抑郁量表(Hamilton depression rating scale, HAMD)、患者健康问卷抑郁量表(patient health questionnaire-9, PHQ-9)等开展评估工作。这一诊断模式为抑郁症的识别和治疗规范化提供了基础支撑,但也带来一系列临床问题^[2]:第一,诊断高度依赖医师经验与患者主观报告,难以及时发现早期和隐匿病例;第二,传统量表以总分为主,难以充分反映症状维度和病程特征的异质性,对复发和治疗抵抗等预后事件的个体化预测能力不足^[3];第三,在轻度抑郁、老年抑郁(late-life depression, LLD)以及焦虑、双相情感障碍等共病或谱系性疾病中,仅依靠症状往往难以实现准确鉴别。

1.1 客观指标与机器学习引入抑郁症诊断的必要性

为了弥补抑郁症诊断的缺陷,近年来,越来越多研究尝试将血清神经递质及其代谢产物、炎症因子、应激激素水平以及多导睡眠图等客观指标,引入抑郁症的诊断与疗效评估中^[4]。例如C反应蛋白(C-reactive protein, CRP)、白细胞介素-6(interleukin-6, IL-6)等外周炎症标志物在抑郁患者中可出现轻度升高^[5];代谢组学层面, Yin 等^[6]对青少年重度抑郁障碍(major depressive disorder in adolescents, MDD)、双相情感障碍、精神分裂症及健康对照的血浆代谢组进行系统比较,并提出用于不同疾病-对照比较的潜在诊断生物标志物面板;在睡眠客观评估方面, Enkhbayar 等^[7]基于多导睡眠图数据构建可解释机器学习模型,并通过特征重要性分析提示睡眠效率等指标具有较高贡献度。但在数据层面,这类客观指标还存在检测平台和批次差异、测量下限和偏态分布、采样时间点不一致、缺失值比例较高等问题。总体来看,单一生物标志物的效应量仍然有限,不同研究在检测技术、采样方案和判定阈值等方面差异显著,尚不足以在常规临床实践中独立承担诊断角色,构成了“数据挑战”。上述现象提示,后续研究需扩大样本量,整合多种客观指标,并借助机器学习等方法,从脑-身-心理一体化的视角提炼出更加稳定、可重复的特征模式。

机器学习是一类以数据驱动为核心的建模方法,可通过在样本中学习特征与标签之间的映射关系,实现对未知个体的自动分类或预测^[8]。从算法选型的角度看,不同临床任务和数据模态对应的建模需求并不相同:在线性特征维度相对有限、样本量中等的场景,逻辑回归(logistic regression, LR)等线性模型结构简单、易于解释,有利于量化各变量对结果的贡献^[9];对于特征维度较高、存在明显非线性

性关系的数据,如脑影像或高密度脑电图,支持向量机(support vector machine, SVM)、随机森林(random forest, RF)和梯度提升决策树(gradient boosting decision tree, GBDT)等非线性模型更适合捕捉复杂特征交互。近年来也有研究尝试采用卷积神经网络(convolutional neural network, CNN)、循环神经网络(recurrent neural network, RNN)等深度学习模型^[9]。在抑郁症相关研究中,一般流程为:在严格质控基础上采集临床资料、神经影像、脑电图、语音与行为、多组学等多源数据,经预处理和标准化后构建症状维度、脑结构与功能指标、频谱特征、语音参数及行为指标等特征,随后通过特征选择或降维方法减少冗余信息,最终训练分类或回归模型,并采用交叉验证或独立测试集评估性能。现有工作聚焦于三个方面:抑郁症的诊断及与其他精神障碍的鉴别,自然病程和复发风险预测,以及抗抑郁治疗反应和治疗抵抗的预测^[10]。

1.2 神经影像学与机器学习

神经影像学作为观察抑郁症患者脑结构与功能异常的重要手段,目前已成为抑郁症机器学习研究领域中的应用最为广泛的数据模态之一。结构磁共振成像(structural MRI, sMRI)可量化全脑及特定脑区的灰质体积、皮层厚度和白质高信号负荷,功能磁共振成像(functional MRI, fMRI)可反映静息态或任务态下脑活动时间序列及脑区间功能连接,扩散张量成像(diffusion tensor imaging, DTI)则用于刻画白质纤维完整性和脑网络结构^[11]。大量研究表明,抑郁症患者在前额叶-边缘系统、默认模式网络(default mode network, DMN)、认知控制网络和奖赏网络等关键环路中,普遍存在不同程度的结构和功能异常^[12]。在此基础上,神经影像学可以对这些异常模式进行客观量化,为提取与抑郁症诊断、分型及预后相关的影像特征提供重要依据。

在sMRI领域,早期研究多采用sMRI结合传统机器学习算法,对MDD与健康对照进行二分类。常见做法为提取海马、杏仁核、前额叶皮层和扣带回等情绪调节相关脑区的体积及皮层厚度,或利用全脑体素形态学及基于区域的形态测量结果,输入SVM或正则化LR模型。部分研究将对象拓展至LLD,通过整合额叶皮质萎缩、海马萎缩和白质高信号负荷等特征,区分LLD与正常老年、轻度认知功能障碍等人群,并尝试预测抗抑郁治疗反应及随访期间功能结局^[13]。面向LLD的机器学习模型在建模目标和特征选择上具有一定特殊性,通常需要在

同龄健康老年人、老年期抑郁、轻度认知障碍或痴呆以及血管性认知障碍之间进行鉴别,因此更强调白质高信号负荷、脑萎缩模式以及脑小血管病相关改变等结构损害特征,并常将高血压、糖尿病、脑卒中史等血管危险因素以及简易精神状态检查量表(mini-mental state examination, MMSE)、蒙特利尔认知评估(Montreal cognitive assessment, MoCA)等认知量表纳入特征空间。

静息态fMRI研究常基于全脑或特定网络的血氧水平依赖(blood oxygenation level dependent, BOLD)信号构建功能连接矩阵,计算DMN、认知控制网络、边缘系统网络和显著性网络的连接强度及图论指标,再输入分类模型区分MDD与健康对照或识别不同临床亚型,并对症状维度或病程特征进行预测。部分多中心研究还使用静息态连接模式预测重复经颅磁刺激(repetitive transcranial magnetic stimulation, rTMS)等干预的疗效,为个体化神经调控方案的制定提供依据。

任务态fMRI则多围绕情绪加工、奖赏和认知控制范式下的激活模式,用于预测抗抑郁药物、rTMS、电休克治疗等干预的疗效^[14]。现有结果表明,不同机器学习模型中常出现前额叶-边缘系统功能连接异常、DMN内部连通性增高或与认知控制网络解耦、奖赏相关纹状体反应减弱等特征,且上述特征与抑郁症的临床诊断及治疗反应密切相关^[15]。

DTI等白质结构影像为抑郁症脑网络异常提供了补充信息。通过提取前额叶-边缘系统、额叶-顶叶网络和胼胝体等区域的各向异性分数和平均弥散率,以及全脑白质纤维追踪和图论分析得到的网络效率和节点度等指标,可构建用于区分抑郁与健康、不同发病亚型及预后结局的机器学习模型。部分研究还将代谢或受体成像指标与结构影像联合分析,从神经递质功能改变与结构损害的对应关系出发,探索更具生物学意义的影像特征。

总体而言,基于神经影像学数据的机器学习在抑郁症识别、亚型划分及疗效预测等方面均展现出一定应用潜力,但现有研究多局限于单中心、小样本范畴,且在特征构建与验证策略上存在明显差异。随着多中心协作项目的推进及大样本队列研究的开展,已有研究指出,小样本研究中常用的留一法或单次划分交叉验证策略往往会高估模型分类性能,而在样本量达数百例以上且采用重复交叉验证或外部验证的数据集上,MDD与健康对照的分类准确率多维持在中等水平^[9]。这提示未来研究需

在扩大样本量、规范数据预处理流程及特征构建标准的基础上,进一步重视模型泛化能力与可重复性的提升。

1.3 其他单模态数据

除神经影像学外,脑电图(electroencephalography, EEG)、临床与心理量表、语音与行为以及多组学等其他数据模态也已应用于抑郁症机器学习研究^[16]。现有EEG研究多基于静息态或任务态信号,提取多类特征并采用SVM、RF等方法区分MDD、健康对照及其他精神障碍^[17]。部分研究借助治疗前以及早期随访的EEG数据预测抗抑郁治疗的临床反应,显示其在抑郁症疗效预测中的独特优势^[18]。临床与心理量表可通过系统整合人口学特征、病程信息、共病情况及多维量表得分,为机器学习模型识别复发风险、高残疾风险及治疗抵抗的抑郁症高危人群提供关键支撑^[19]。语音与数字行为数据则可提供抑郁症患者日常生活中的客观表型,通过提取语音/副语言特征即可实现MDD评估^[20-21]。现有研究多基于语速、停顿时长、基频及其波动、强度、抖动、微扰以及梅尔频率倒谱系数等语言/副语言特征,构建抑郁识别模型。总体来看,这类模型在区分MDD与健康对照或不同抑郁严重程度方面,内部验证的准确率多在70%~85%,曲线下面积(area under the curve, AUC)为0.75~0.90,部分研究的F1值也达到中高水平。然而,方言口音、录音环境噪声、说话任务差异以及共病焦虑、药物使用状态等因素均可能影响语音特征分布,加上外部验证样本有限,导致模型泛化能力和可重复性仍有待进一步评估。未来需在样本量更大、多中心协同及更贴近真实门诊场景的语音队列中,规范特征提取与建立标准化流程,并系统比较不同算法在抑郁症语音识别中的性能与适用范围。此外,多组学研究通过整合基因组、转录组、代谢组等高维分子特征与临床资料,利用特征筛选和集成学习方法识别与抑郁风险、临床亚型和治疗反应相关的分子标志物及通路,为理解抑郁症生物学基础提供了新的线索^[22-23]。

2 多模态数据在抑郁症诊断中的作用

抑郁症的临床核心问题在于:如何在症状高度异质、共病复杂的前提下,实现早期识别、精细分型以及复发和治疗抵抗等预后事件的个体化预测。前文已经指出,仅依赖访谈和量表等主观信息以及单一客观指标,往往难以实现抑郁症早期诊疗和个体化诊疗。诊断依赖经验、亚型界限模糊、预后预

测能力有限等,这构成了现阶段的“数据挑战”^[2]。随着神经影像、电生理、生物标志物及行为监测等技术的发展,在同一患者身上获取多种客观指标成为可能,即形成多模态数据。在本综述中,多模态主要指以神经影像学数据为核心,同时包括EEG、语音与行为、临床与心理量表以及多组学等多源信息。相比单一数据模态,多模态数据为抑郁症的临床诊断、精准分型及病程动态评估提供了更立体全面的研究视角,这也使得从脑结构特征、脑功能活动、身心状态及日常行为表现等多个维度系统且综合地认识抑郁症成为可能(表1)。

2.1 多模态神经影像的结构-功能-连接一体化表征
在神经影像领域,多种成像技术的联合应用是

理解抑郁症脑网络异常的重要基础。常见的组合方式包括:sMRI联合静息态fMRI,用于从灰质形态和静息态网络同步性双层面表征脑网络失衡;sMRI联合DTI,用于同时评估皮层萎缩与白质纤维束完整性;sMRI/DTI联合任务态fMRI或磁共振波谱,从结构-功能-代谢耦合的角度刻画情绪和认知环路的异常。例如在一项研究中,部分患者表现为前额叶皮层、海马等关键区域萎缩且功能连接减弱,呈现结构损害伴功能降低的特征;另一些患者结构指标相对正常,但DMN内部或DMN与认知控制网络间存在过度同步或解耦现象,以脑功能失衡为主要表现^[24]。这类多模态影像观察有助于将抑郁症从单一维度的有无异常转变为多维度的结构-功能表型,

表1 抑郁症多模态数据类型及其在机器学习诊断中的应用概述

Table 1 Overview of multimodal data types for depression and their application in machine learning diagnosis

Data modality	Key information reflected	Typical machine learning tasks	Main advantages and limitations	Representative literature
Neuroimaging	Gray matter volume, cortical thickness, white matter hyper-intensity burden; fractional anisotropy, mean diffusivity, white matter tractography, and network efficiency	Classification: diagnosis(MDD vs. HC or others) & geriatric differential diagnosis; Precision medicine: subtyping & treatment response prediction	Pros: high spatial resolution; maps network topology; potential for objective biomarkers; Cons: high cost; significant batch effects; overfitting risk in small samples	5, 9-12, 14, 26-27
EEG & electrophysiological signals	Temporal dynamics of brain activity, oscillatory rhythms, and neural synchronization	Assessment: automated screening & severity grading; Prediction: forecasting outcomes of drugs or stimulation	Pros: high temporal resolution; cost-effective; portable; ideal for longitudinal monitoring; Cons: low spatial resolution; sensitive to artifacts; lacks standardization	15-19, 28, 31, 37
Speech & digital behavioral data	Paralinguistic acoustic features; ecological behaviors such as physical activity levels and sleep rhythms	Detection: objective recognition of depressive states; Monitoring: continuous tracking of symptom fluctuation & relapse prediction	Pros: non-invasive; high ecological validity; enables passive, continuous sensing; Cons: Environmental confounders; privacy risks; ground-truth labeling difficulties	20-21
Clinical or psychometric Scales & broad clinical data	Symptom dimensions, course characteristics, comorbidity profiles, and treatment history.	Risk stratification: predicting onset and relapse; Trajectory modeling: identifying treatment-resistant or high-risk subgroups	Pros: low cost; scalable to multicenter cohorts; broad clinical coverage; Cons: subjective reporting bias; total scores mask symptom heterogeneity; label noise	2, 3, 8, 23-25
Multi-omics & fluid biomarkers	Genetic susceptibility, transcriptional and metabolic states; inflammatory and endocrine levels	Profiling: biological risk identification & subtyping; Biotyping: linking molecular signatures to clinical phenotypes	Pros: mechanistic insight; facilitates novel target discovery; Cons: high cost; high dimensionality; reproducibility challenges; requires large samples & rigorous QC	22-24

为后续亚型划分提供基础。

相比单模态数据难以全面刻画抑郁症相关脑网络多维异常模式的局限性,多模态神经影像在临床分型和鉴别诊断中具有更重要的价值。以老年人群为例,LLD与轻度认知功能减退、阿尔茨海默病等在情绪和认知症状上存在大量重叠,仅凭临床表现和传统量表往往难以区分。通过联合分析老年患者的皮层萎缩模式、海马体积、白质高信号负荷以及静息态功能连接和网络效率等指标,可在以情绪障碍为主的LLD表型与以神经退行性过程为主的认知障碍表型之间建立相对清晰的界限,并提示两者在脑网络层面的差异^[25]。类似地,在广泛性焦虑障碍、双相情感障碍等情绪障碍之间,通过比较杏仁核-前额叶回路、前扣带和岛叶等关键区域的结构与功能特征,有助于识别以焦虑成分为主的抑郁与以抑郁成分为主的焦虑,提高临床诊断的精细化水平。

2.2 影像联合临床指标

在影像与非影像信息的联合方面,多模态思路逐渐由单纯的多种扫描技术叠加,拓展为脑-身-心理多层信息的综合表征。大量研究通过同时收集结构和功能影像、抑郁与焦虑量表、病程信息及认知功能指标,探讨脑网络异常与具体症状维度之间的对应关系^[26]。例如,快感缺失突出的患者更常表现为前额叶-纹状体奖赏环路异常,而以注意和执行功能障碍为主的患者,其相关改变则更多涉及背侧前额叶及顶叶网络^[27]。这提示影像指标与临床症状并非彼此独立,而是能够在症状维度层面形成相互印证。

相比单纯依赖量表总分或单一影像异常,影像联合临床指标更有助于揭示抑郁症内部异质性。一方面,影像数据可提供脑结构和功能网络异常的客观证据;另一方面,临床症状、病程及认知表现则有助于界定这些脑异常在个体层面的临床含义。二者结合有助于突破传统诊断中同病异质和异病同象的局限,推动构建基于症状维度和脑网络机制的抑郁症识别框架。此外,影像与体液、生理应激等指标的联合,也为识别以炎症或应激反应为主要特征的抑郁亚型提供了可能,从而为临床分层和风险评估提供更丰富的客观依据。

2.3 非影像多模态数据

在更广义的多模态框架下,电生理、语音与行为以及多组学等非影像数据之间的联合,也为抑郁症诊断提供了重要补充。EEG以毫秒级时间分辨率刻画脑活动动态过程,与MRI所表征的空间结构和功能网络信息形成互补^[19];语音和数字行为数据

则能够反映患者在真实生活情境中的日常状态,如语速、停顿、基频波动、活动量及睡眠节律等,为识别精神运动迟滞、社会退缩及病情波动提供了较高生态效度的客观线索。与此同时,多组学研究通过整合基因组、转录组和代谢组等分子层信息,可进一步解释不同临床表型和行为特征背后的生物学差异^[25]。

总体而言,多模态数据的核心价值并不在于简单叠加信息,而在于从脑结构与功能、症状维度、行为表型及分子特征等不同层面提供相互补充的证据,从而为抑郁症的早期识别、亚型划分和风险分层提供更完整的表征基础。本部分重点阐明多模态数据的互补价值及其临床意义;至于如何借助机器学习实现跨模态整合、不同融合策略的适用条件及其具体应用场景,将在下一部分进一步展开讨论。

3 机器学习结合多模态数据在抑郁症诊断中的应用

3.1 多模态融合策略与建模流程

多模态数据为理解抑郁症复杂的脑-身-心理改变提供了更为全面的基础,但在实际应用中,如何将来源、尺度和噪声特性各不相同的多种指标有机整合,仍是一项重要挑战。机器学习在这一过程中发挥关键作用:一方面,可通过特征选择、降维和正则化等方法处理高维、多源数据,缓解维度灾难和多重共线性问题^[28];另一方面,可在统一模型框架下实现多模态信息的加权融合,从而提高诊断、分型和预后预测的准确性与稳定性^[28]。常见的多模态融合策略包括特征层面的早期融合,即在输入阶段将不同模态特征进行拼接,以及决策层面的晚期融合,即分别建模后在输出阶段进行加权整合。不同策略在样本量、特征维度和任务目标不同时各有适用场景(图1)。

在抑郁症领域,早期融合常用于提升诊断和分型精度,晚期融合则更适合处理模态不全、任务目标多样的情境。现有抑郁症多模态机器学习研究在算法层面大致可分为3类:一类是以正则化LR、线性SVM等为代表的特征级早期融合模型,适用于样本量有限、特征维度极高的场景;一类是以RF、梯度提升树和集成学习为代表的模型级或决策级晚期融合方法,更侧重不同模态子模型的加权组合;还有一类是多分支CNN、图神经网络等深度学习框架,可在大样本、多中心多模态数据上实现端到端表征学习^[29]。

3.2 多模态神经影像模型

多模态机器学习的进一步发展,主要体现在不同信息源的任务导向整合及其临床场景拓展。在

神经影像领域,多模态机器学习模型已被广泛用于提升抑郁症的诊断与分型性能。基于单一结构或功能模态构建的模型往往只能捕捉抑郁症脑网络异常的某一侧面,而将sMRI提取的灰质体积、皮层厚度、白质病灶负荷与fMRI的功能连接和图论指标、DTI的白质完整性和网络效率等特征联合输入模型,可在结构-功能-连接的综合框架下刻画个体的脑网络状态^[2]。部分研究采用早期融合策略,将多模态影像特征统一标准化后直接拼接,再通过LASSO和最小冗余最大相关(minimum redundancy maximum relevance, mRMR)等方法筛选具有代表性的联合特征,进而使用SVM或RF进行分类。结果显示,在区分MDD与健康对照、LLD与认知障碍患者方面,多模态模型的准确率普遍优于任何单一模态。也有研究采用晚期融合方法,对结构、功能和扩散影像分别建立子模型,再根据各模态在验证集中的表现确定权重,构建综合评分用于诊断或疗效预测,在一定程度上缓解了多模态数据量纲差异和缺失不一致的问题。

3.3 影像联合临床与生物学指标的多模态模型

影像联合临床与生物学指标是当前抑郁症多模态机器学习研究中较具临床转化意义的方向。相比单纯依赖影像或量表评分,此类研究更强调在统一模型框架下整合脑网络特征、症状维度、病程信息、认知功能及体液指标等多源信息,以提升鉴别诊断、亚型识别及风险预测的稳定性^[30]。其中,影像特征可提供相对客观的脑结构和功能异常证据,临床与生物学指标则有助于增强模型对当前症状状态和个体差异的识别能力。从具体应用看,这类模型尤其适用于区分临床表现相近但潜在机制不同的人群。例如,在广泛性焦虑障碍与MDD的鉴别研究中,将临床症状结构、应激激素水平与结构MRI特征共同纳入模型后,其性能较仅采用临床信息者明显改善,其中海马和前额叶等脑区指标提供了额外判别信息^[22]。类似地,在LLD及认知障碍相关研究中,影像联合病程和认知指标的模型也更有助于建立面向老年人群的精细化识别框架。

3.4 影像联合EEG、语音-行为及多组学的多模态模型

除影像联合临床资料外,影像与EEG、语音-行为以及多组学的联合建模,进一步拓展了抑郁症机器学习的应用范围。此类研究通常需要处理不同模态在时间尺度、数据结构及缺失模式上的差异,因此更强调根据具体任务选择合适的融合方式。

影像与EEG的联合,主要用于同时整合脑结构和功能网络的稳定特征,以及电生理信号对短期状态变化的敏感性,可用于预测治疗反应^[18];影像与语音、活动量、睡眠节律等行为特征的联合,更适用于病情监测和复发预警^[21];影像与多组学的联合分析,则有助于将脑表型与分子机制联系起来,为精准分型和潜在靶点识别提供依据。总体来看,这类研究的重点已不再是简单叠加不同模态,而是围绕具体临床场景优化融合对象与建模路径,从而提高模型在诊断、预测和分层中的实际应用价值。

3.5 模型可解释性与临床转化

由于多模态特征数量众多且相互关联,若仅追求预测性能而忽视解释,将难以为临床医生所接受^[2]。因此,越来越多研究在建模时引入具有明确特征权重的算法,或在模型训练完成后采用特征重要性排序、部分依赖图和Shapley加性解释等方法,识别对模型输出贡献最大的关键特征^[30]。相关结果显示,多模态影像模型和影像联合临床、生物学特征的模型大体指向相似的一组核心预测因子:前额叶-边缘系统、DMN和认知控制网络的异常,以及残余症状、病程指标和部分炎症、代谢标志物^[31]。这既为抑郁症脑网络及身心相互作用的机制研究提供线索,也为临床报告形式的设计提供依据,如将模型输出转换为高危和低危提示,并附上主要贡献特征及其可能的临床解释,便于医生理解和使用^[32]。

从临床应用角度看,机器学习结合多模态数据的突出价值主要体现在3个方面:提高筛查与辅助诊断效能,支持亚型划分与风险分层,以及为治疗反应预测和随访监测提供量化依据。相关Meta分析显示,多模态抑郁筛查模型的合并AUC可达0.95,提示其在辅助诊断中的应用潜力突出^[33];而在治疗结局预测方面,现有高质量研究的总体效能仍多处于中等水平。总之,前两方面证据相对更成熟,而治疗结局预测虽显示出应用前景,仍需进一步加强外部验证。

从临床部署看,近年来通用大模型在电子病历整理、临床文本理解及多源信息归纳方面显示出较强优势,能够在一定程度上提升临床资源整合效率,并为初步筛查、信息提炼和流程衔接提供支持;而面向特定任务构建的专用机器学习模型则更适用于诊断分类、疗效预测和复发预警等目标明确、标签清晰的具体任务,其结果也更便于量化评估和独立验证。二者并非相互替代,而更可能形成互补:前者侧重整合异构信息、提高临床场景适应性,后者侧重在关键诊疗节点输出稳定、可解释的预测

结果,从而共同服务于临床决策^[34]。

综上,现有多模态机器学习研究已经在多个临床任务上完成了初步的算法分型和实证对比:在样本量有限、特征维度极高且需要较强可解释性的影像和多组学场景中,经特征筛选与正则化的线性模型和基于树的集成方法往往在外部验证中表现更

稳健;在多中心大样本或高频时间序列数据上,深度学习框架在AUC和灵敏度等指标上可获得一定增益,但对数据质量和建模规范更为敏感。此外,多数工作证实在同等任务和样本下,多模态模型相对于最佳单模态模型通常可带来中度但一致的性能提升,尤其在临床亚型划分和治疗反应预测方面更为明显。

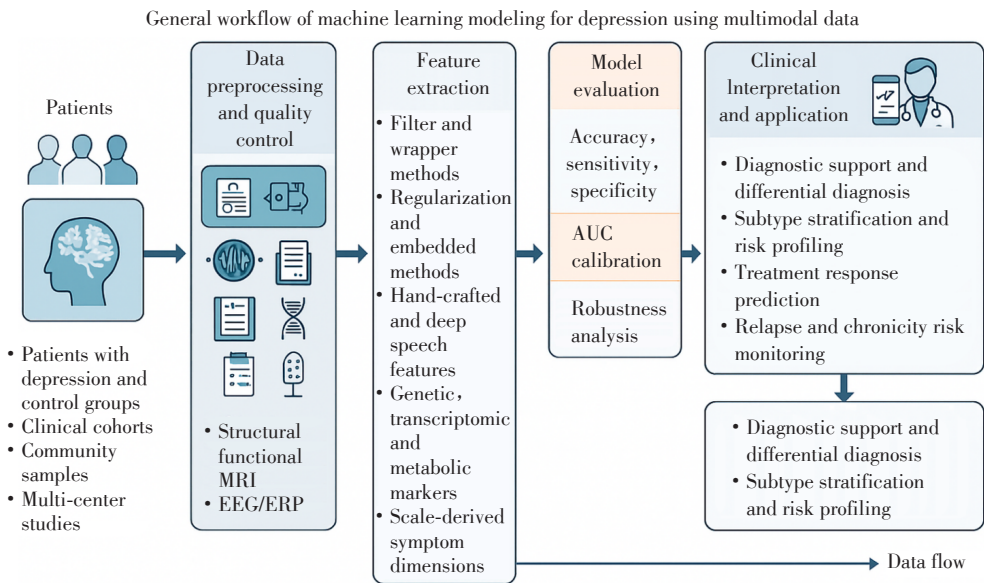


图1 基于多模态数据的抑郁症机器学习建模通用工作流程图

Figure 1 General workflow of machine learning modeling for depression using multimodal data

4 小结与展望

抑郁症具有患病率高、复发率高和致残率高等特点,传统以访谈和量表为核心的诊断模式在早期识别、精细分型以及预后和疗效预测方面存在明显不足^[35]。围绕抑郁症早期识别和精准分型这一核心临床问题,文章从单模态客观指标、多模态数据特性以及机器学习建模策略3个层面梳理了现有证据。纵观现有研究,可得到以下几点认识。第一,机器学习已在临床与心理量表、神经影像、EEG、语音与行为以及多组学等单一模态中得到广泛应用,在抑郁症诊断、病程预测和疗效评估方面显示出一定优势,其中神经影像学数据是当前证据最为集中的领域,距离潜在客观生物标志物最近^[13]。第二,多模态数据可从结构、功能、连接、症状和行为等多个层面刻画抑郁症,为亚型划分和风险分层提供了比单一模态更丰富的信息。第三,机器学习为多模态数据整合提供了关键技术手段,多模态模型在诊断、分型和疗效预测中的表现普遍优于单模态模型,有望成为实现抑郁症精准诊疗的重要工具^[36]。

同时应看到,现有研究仍存在若干共性问题。多数工作基于单中心、小样本,预处理流程、特征构建和验证策略差异较大,模型性能常在内部验证环境中被高估,外部泛化能力和结果可重复性有限。多模态数据采集成本高、流程复杂,真正具备影像、EEG、体液指标和详细临床资料的样本数量有限,不同中心在扫描协议、设备参数和量表组合方面的差异,又进一步加剧了批次效应和系统偏倚问题^[37]。多模态机器学习模型结构复杂,若缺乏充分的可解释性,将难以融入临床决策流程;即便模型具有统计学意义,如何将其输出转化为简明、可操作的临床信息,如风险分层、关键特征提示等,仍需在真实世界环境中反复打磨和验证。

面向未来,抑郁症领域的机器学习与多模态研究可从以下几个方面推进。其一,建设大样本、多中心、多模态协作队列,在保证隐私和数据安全前提下,逐步推动采集方案和预处理流程的标准化与数据共享,为模型训练和外部验证提供更为坚实的基础。其二,在方法学层面重视预注册设计和规范化分析流程,采用更稳健的交叉验证和独立验证

集,结合迁移学习、领域自适应等技术,提升模型在不同中心和人群中的泛化能力^[38]。其三,加强模型可解释性和生物学合理性,结合可解释机器学习、网络神经科学和多组学分析,探索前额叶-边缘系统、DMN与炎症、代谢等通路之间的潜在联系,从相关性逐步走向机制假设的提出。其四,在临床转化层面,开展前瞻性及嵌入式研究,评估多模态机器学习辅助系统对诊断一致性、治疗决策和复发预警的实际影响,同时关注患者体验、伦理规范和监管路径^[38]。

随着多模态数据资源的持续积累和建模方法的不断完善,有望逐步形成以神经影像为核心、整合电生理、生物标志物、临床信息与数字行为的综合评估框架,为抑郁症的早期识别、精准分层和个体化干预提供更加坚实的证据支持^[39-40]。

利益冲突声明:

所有作者声明无利益冲突。

Conflict of Interests:

All authors declare no conflict of interests.

作者贡献声明:

高子祥、李翔、张欣负责论文撰写;钱昭均、嵇子慧负责材料整合;朱俐璇、徐忆初负责图表制作;陈亚丽、葛菲菲负责方向指导与修改。

Author's Contributions:

GAO Zixiang, LI Xiang, and ZHANG Xin drafted the manuscript; QIAN Zhaojun and JI Zihui were responsible for material integration; ZHU Lixuan and XU Yichu were responsible for chart preparation; CHEN Yali and GE Feifei were responsible for direction guidance and revisions.

[参考文献]

- [1] 王 颖,侯宇威,张 哲. 中国人群抑郁症疾病负担的年龄-时期-队列分析[J]. 四川精神卫生, 2025, 38(3): 254-260
WANG Y, HOU Y W, ZHANG Z. Age-period-cohort analysis of the disease burden of depression in the Chinese population[J]. Sichuan Journal of Mental Health, 2025, 38(3): 254-260
- [2] LIN S K, KUO P H, HSU C Y, et al. The effects of *Lactobacillus plantarum* PS128 in patients with major depressive disorder: an eight-week double-blind, placebo-controlled study[J]. Asian J Psychiatr, 2024, 101: 104210
- [3] XU R F, LIU Z J, OUYANG S, et al. Machine learning-driven development of a stratified CES-D screening system: optimizing depression assessment through adaptive item selection[J]. BMC Psychiatry, 2025, 25(1): 286
- [4] GAN X, LI X, CAI Y, et al. Metabolic features of adolescent major depressive disorder: a comparative study between treatment-resistant depression and first-episode drug-naïve depression[J]. Psychoneuroendocrinology, 2024, 167: 107086
- [5] BATAIL J M, COROUGE I, BLANCHARD T, et al. Inflammatory and MRI perfusion biomarkers in predicting persistence of depression: a 6-month longitudinal study[J]. Transl Psychiatry, 2025, 15(1): 370
- [6] YIN B, CAI Y, TENG T, et al. Identifying plasma metabolic characteristics of major depressive disorder, bipolar disorder, and schizophrenia in adolescents[J]. Transl Psychiatry, 2024, 14(1): 163
- [7] ENKHBAYAR D, KO J, OH S, et al. Explainable artificial intelligence models for predicting depression based on polysomnographic phenotypes[J]. Bioengineering, 2025, 12(2): 186
- [8] HU Y X, SHI J Y, XIA G Y, et al. Analysis of functional connectivity changes in attention networks and default mode networks in patients with depression and insomnia[J]. Sleep Breath, 2024, 28(4): 1731-1742
- [9] GRIMM B, YILMAM P, TALBOT B, et al. Multimodal machine learning for video based single question mental health assessment[J]. NPJ Digit Med, 2025, 9(1): 65
- [10] CLAAS F, MICAH C, NILS O, et al. Systematic misestimation of machine learning performance in neuroimaging studies of depression[J]. Neuropsychopharmacology, 2021, 46(8): 1510-1517
- [11] GUO Y, CHU T, LI Q, et al. Diagnosis of major depressive disorder based on individualized brain functional and structural connectivity[J]. J Magn Reson Imaging, 2025, 61(4): 1712-1725
- [12] DAI P, HUANG K, HU T, et al. Altered effective connectivity in patients with drug-naïve first-episode, recurrent, and medicated major depressive disorder: a multi-site fMRI study[J]. Behav Brain Res, 2025, 495: 115756
- [13] GUO P, LIU X, SUN Y, et al. Effects of sMRI-guided rTMS on brain function and structure in major depressive disorder[J]. J Affect Disord, 2025, 390: 119772
- [14] VU T, DAWADI R, YAMAMOTO M, et al. Prediction of depressive disorder using machine learning approaches: findings from the NHANES[J]. BMC medical informatics and decision making, 2025, 25(1): 83
- [15] LUO Y, SHEN Y, FAN X. EEG microstates in adolescent depression: effects of depression severity and overall symptoms[J]. J Affect Disord, 2025, 390: 119819
- [16] MULC D, VUKOJEVIC J, KALAFATIC E, et al. Opportunities and challenges for clinical practice in detecting depression using EEG and machine learning[J]. Sensors, 2025, 25(2): 409-409
- [17] YANG CY, CHEN Y Z. Support vector machine classification of patients with depression based on resting-state

- electroencephalography [J]. Asian Biomed (Res Rev News), 2024, 18(5): 212-223
- [18] DENIER N, WALTHER S, BREIT S, et al. Electroconvulsive therapy induces remodeling of hippocampal co-activation with the default mode network in patients with depression[J]. Neuroimage Clin, 2023, 38: 103404
- [19] ALOTAIBI N M, BAKHEET D M. Predicting depression severity using effective and functional brain connectivity of the electroencephalography signals [J]. Comput Biol Med, 2025, 190: 110045
- [20] RAVAN M, NOROOZI A, GEDIYA H, et al. Using deep learning and pretreatment EEG to predict response to sertraline, bupropion, and placebo [J]. Clin Neurophysiol, 2024, 167: 198-208
- [21] BAUER J F, GERCZUK M, SCHINDLER-GMELCH L, et al. Validation of machine learning-based assessment of major depressive disorder from paralinguistic speech characteristics in routine care [J]. Depress Anxiety, 2024, 2024(1): 9667377
- [22] IKÄHEIMONEN A, LUONG N, BARYSHNIKOV I, et al. Predicting and monitoring symptoms in patients diagnosed with depression using smartphone data: observational study[J]. J Med Internet Res, 2024, 26: e56874
- [23] EDER J, PFEIFFER L, WICHERT P S, et al. Deconstructing depression by machine learning: the POKAL - PSY study[J]. Eur Arch Psychiatry Clin Neurosci, 2023, 274(5): 1153-1165
- [24] ALOTAIBI N M, ALHOTHALI A M, ALI M S. Multi-atlas ensemble graph neural network model for major depressive disorder detection using functional MRI data [J]. Front Comput Neurosci, 2025, 19: 1537284
- [25] SAMPRIT B, YIYUAN W, S K B, et al. Trajectories of remitted psychotic depression: identification of predictors of worsening by machine learning[J]. Psychol med, 2024, 54(6): 1142-1151
- [26] MARDINI M T, KHALIL G E, BAI C, et al. Identifying adolescent depression and anxiety through real-world data and social determinants of health: machine learning model development and validation[J]. JMIR Ment Health, 2025, 12: e66665
- [27] MA S, NIE Z, ZHANG M, et al. Towards precision psychiatry: metabolomics identifies three biological subtypes of depression [J]. PLoS Digit Health, 2025, 4(12): e0001125
- [28] TOZZI L, ZHANG X, PINES A, et al. Personalized brain circuit scores identify clinically distinct biotypes in depression and anxiety[J]. Nat Med, 2024, 30(7): 2076-2087
- [29] JIAO Y, ZHAO K, WEI X, et al. Deep graph learning of multimodal brain networks defines treatment - predictive signatures in major depression[J]. Mol Psychiatry, 2025, 30(9): 3963-3974
- [30] HAN W, WU X, WANG L, et al. Altered brain function in treatment - resistant depression patients: a resting - state functional magnetic resonance imaging study [J]. Neurosci Lett, 2024, 842: 138004
- [31] LEE D Y, KIM N, PARK C, et al. Explainable multimodal prediction of treatment-resistance in patients with depression leveraging brain morphometry and natural language processing[J]. Psychiatry Res, 2024, 334: 115817
- [32] QAYYUM A, RAZZAK I, TANVEER M, et al. High-density electroencephalography and speech signal based deep framework for clinical depression diagnosis [J]. IEEE/ACM Trans Comput Biol Bioinform, 2023, 20(4): 2587-2597
- [33] WEBER J, WEBER M, LOPEZ A. Depression diagnosis from patient interviews using multimodal machine learning[J]. Front Psychiatry, 2025, 16: 1694762
- [34] KIM J, MA S P, CHEN M L, et al. Optimizing large language models for detecting symptoms of depression/anxiety in chronic diseases patient communications [J]. NPJ Digit Med, 2025, 8(1): 580
- [35] YANG W, WANG X, KANG C, et al. Establishment of a risk prediction model for suicide attempts in first-episode and drug naïve patients with major depressive disorder[J]. Asian J Psychiatr, 2023, 88: 103732
- [36] SCHNELLBÄCHER G J, RAJKUMAR R, VESELINOVIĆ T, et al. Structural alterations as a predictor of depression-a7-Tesla MRI-based multidimensional approach[J]. Mol Psychiatry, 2025, 30(6): 2517-2524
- [37] COLOMBO F, CALESELLA F, BRAVI B, et al. Multimodal brain-derived subtypes of major depressive disorder differentiate patients for anergic symptoms, immune-inflammatory markers, history of childhood trauma and treatment-resistance[J]. Eur Neuropsychopharmacol, 2024, 85: 45-57
- [38] LIU R, HOU X, LIU S, et al. Predicting antidepressant response *via* local-global graph neural network and neuroimaging biomarkers[J]. NPJ Digit Med, 2025, 8(1): 515
- [39] MAZURKA R, CUNNINGHAM S, HASSEL S, et al. Relation of hippocampal volume and SGK1 gene expression to treatment remission in major depression is moderated by childhood maltreatment: a CAN-BIND-1 report[J]. Eur Neuropsychopharmacol, 2024, 78: 71-80
- [40] WANG M, WANG H, FENG Z, et al. Predicting depression among chinese patients with narcolepsy type 1: a machine - learning approach [J]. Nat Sci Sleep, 2024, 16: 1419-1429
- (收稿: 2026-03-19; 修回: 2026-04-28; 录用: 2026-04-30)
- (本文编辑: 唐震)