

◆ 计算科学 ◆

大语言模型与多模态模型在临床医学中的应用与挑战^{*}邹 源, 谈玉平^{**}

(广西医科大学第二附属医院, 广西南宁 530007)

摘要:大语言模型(Large Language Models, LLMs)和多模态模型(Multimodal Models, MMLs)通过整合文本、图像、语音等多模态数据,为临床诊断、个性化治疗及慢性病管理提供了全新的技术支持。本文系统梳理了 LLMs 和 MMLs 的技术基础及其在临床医学中的应用场景,包括临床诊断与决策支持、个性化医疗、慢性病管理等领域,探讨了其在提升诊断准确性、优化治疗方案及改善患者健康管理等方面的潜力与局限性。同时,深入分析了 LLMs 和 MMLs 在医疗领域面临的技术挑战,包括模型泛化能力不足、可解释性与透明性欠缺、隐私与数据安全风险,以及与现有医疗系统的兼容性问题,并阐述了这些挑战对技术落地和推广的影响。最后,本文展望了模型优化、数据融合及隐私保护等方面的发展方向,提出通过技术创新与跨领域协作,推动人工智能(Artificial Intelligence, AI)技术在医学领域的深度应用,为提升医疗服务效率和质量提供参考。

关键词:大语言模型;多模态模型;临床大数据;临床辅助决策;个性化医疗

中图分类号:R-05 文献标识码:A 文章编号:1005-9164(2025)01-0088-08

DOI: 10.13656/j.cnki.gxkx.20250521.010

近年来,人工智能(Artificial Intelligence, AI)技术,尤其是大语言模型(Large Language Models, LLMs)和多模态模型(Multimodal Models, MMLs),在自然语言处理与多模态数据处理领域取得了显著进展。LLMs 凭借强大的语言生成与理解能力,显著提升了文本数据的分析与处理水平;而 MMLs 通过整合文本、图像、语音等多模态数据,进一步提高了信息处理的全面性和准确性^[1]。同时,临床医学正经历

一场由大数据驱动的变革。电子病历、医学影像、基因组数据及可穿戴设备等多源数据的积累,为医学研究与实践提供了前所未有的丰富资源。然而,这些数据的复杂性与多样性也对现有的临床数据处理能力提出了严峻挑战。只有在临床大数据充分积累且能够有效整合的前提下,LLMs 和 MMLs 的优势才能得以充分发挥,通过高效处理海量的结构化与非结构化数据,提取关键信息,并实现多模态数据融合,从而

收稿日期:2025-01-10

修回日期:2025-02-03

^{*} 广西高校中青年教师科研基础能力提升项目(2022KY0093)和广西壮族自治区卫生健康委西医课题(Z-A20230708, Z-A20240684)资助。

【第一作者简介】

邹 源(1973—),女,高级经济师,主要从事卫生事业管理研究, E-mail:175043696@qq.com。

【**通信作者简介】

谈玉平(1987—),女,副主任技师,主要从事卫生事业管理研究, E-mail:402903066@qq.com。

【引用本文】

邹源,谈玉平. 大语言模型与多模态模型在临床医学中的应用与挑战[J]. 广西科学, 2025, 32(1): 88-95.

ZOU Y, TAN Y P. Application and Challenges of Large Language Models and Multimodal Models in Clinical Medicine [J]. Guangxi Sciences, 2025, 32(1): 88-95.

大幅提升临床数据处理的效率和决策支持能力。这些技术的应用不仅帮助临床工作者快速获取有价值的信息,还能通过融合多模态数据实现更精准的疾病诊断、个性化治疗方案制定及健康管理服务,对医疗行业产生深远影响^[2]。

本文拟梳理 LLMs 和 MMLs 在临床医学中的技术基础、应用场景及面临的挑战,重点探讨它们在临床诊断、个性化医疗及慢性病管理中的潜力与局限性,并展望未来在模型优化、数据融合及隐私保护等方面的发展方向,以期为推动 AI 技术在医学领域的深度应用提供参考。

1 技术基础

1.1 LLMs 的技术原理与发展现状

LLMs 是基于深度学习的自然语言处理模型,其核心架构为 Transformer。该模型通过自注意力机制有效捕捉文本中的长距离依赖关系,并采用两阶段训练范式:首先通过海量无监督数据进行预训练,学习通用语言表征,包括语法、语义和上下文信息;随后基于少量标注数据进行微调,以适配文本生成、问答、机器翻译及逻辑推理等下游任务^[3]。当前,全球范围内已涌现多款具有代表性的 LLMs。国际主流模型包括 OpenAI 的 GPT 系列、Google 的 Gemini、Anthropic 的 Claude 以及 Meta 的 LLaMA 系列;国内典型模型则有百度的“文心一言”、清华大学与智谱 AI 联合研发的“智谱清言”、阿里巴巴的“通义千问”等,这些模型在中文场景及垂直领域(如医疗、教育、金融)的表现可圈可点^[4-5]。此外,最新发布的国产大模型 DeepSeek 在多项基准测试中表现优异,其综合性能已达到与 ChatGPT-4.0 比肩的水平,进一步彰显了我国在 AI 领域的技术实力与创新潜力。

1.2 MMLs 的特性与应用

MMLs 通过整合文本、图像、语音等多模态数据,在统一的表示空间中实现跨模态对齐与信息融

合,从而整合不同数据源的互补信息,显著提升任务处理的综合性能和准确性^[6]。目前,该领域的代表性模型包括 OpenAI 的 GPT-4V、Google 的 Gemini 以及阿里巴巴的“通义万相”。这些模型支持多模态交互,适用于图像描述、视觉问答等复杂任务。此外,清华大学与智谱 AI 联合研发的 GLM-4V-Plus 进一步扩展了多模态处理能力,在图像分析、视频理解和时序感知等专业领域展现出突出优势^[6]。

1.3 临床大数据的支持与整合

临床医学中大数据为 LLMs 和 MMLs 的发展提供了丰富的数据支撑,包括电子病历、医学影(图)像、基因组数据和可穿戴设备数据等^[7]。LLMs 通过分析海量医学文本数据,能够掌握专业知识并应用于临床辅助决策、智能问诊和患者教育;而 MMLs 则通过融合文本、医学影(图)像和基因组数据等多模态数据,提供更精准的诊断支持^[7]。此外,大数据技术在数据清洗、标注、存储和分析流程中的优化,进一步提升了这些模型在复杂临床环境中的适应能力。

1.4 临床 AI 技术框架整合

基于上述技术基础,本文构建了一个 LLMs 与 MMLs 在临床 workflows 中的整合框架(图 1)。该框架由 4 个核心层次构成:临床数据输入层、AI 模型层、临床应用层和反馈优化层。临床数据输入层中,系统整合了电子健康记录(Electronic Health Record, EHR)、医学影像、基因组数据、可穿戴设备数据以及医学文献知识库等多源异构数据,为 AI 模型提供全面支撑。AI 模型层中,LLMs 主要负责医学文本语义理解、临床决策支持和知识问答,而 MMLs 则专注于医学影像分析、多模态数据融合和跨模态推理。临床应用层将模型输出运用到诊断支持、个性化治疗方案、慢性病管理和健康风险预测等实际场景。反馈优化层通过临床验证、模型优化和性能评估形成闭环,不断提升系统效能,确保 AI 技术能够精准、高效地服务于临床实践。

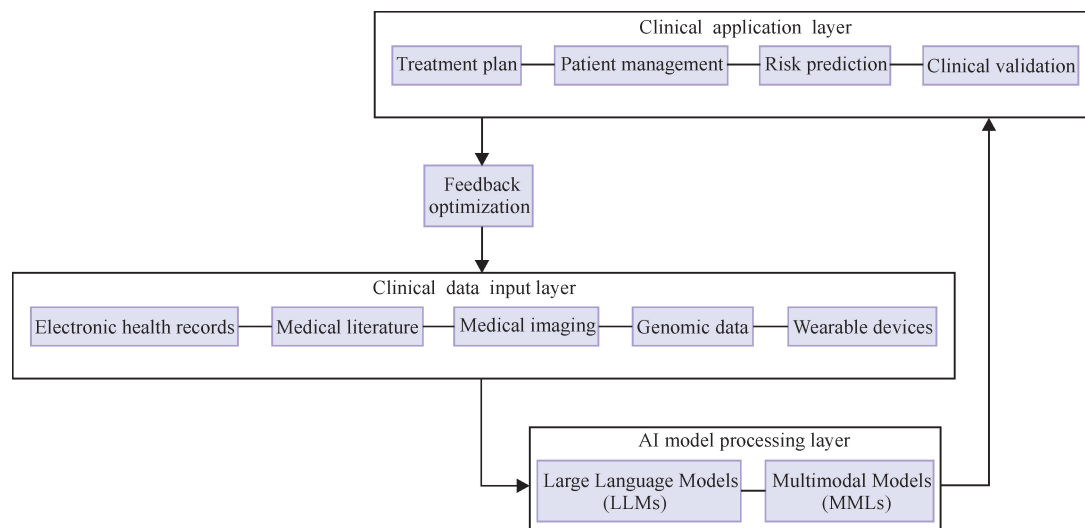


图 1 LLMs 与 MMLs 在临床 workflow 中的整合框架

Fig. 1 Integration framework of LLMs and MMLs in clinical workflow

2 应用场景

2.1 临床诊断与决策支持

2.1.1 LLMs 在临床诊断与决策支持中的应用潜力与挑战

LLMs 在临床诊断与决策支持中展现出巨大的潜力,能通过分析海量医学文本数据(如电子健康记录、指南和医学文献等),为医生提供高效、精准的辅助工具。在病例分析中,LLMs 可快速提取患者主诉、病史、检查结果等关键信息,并生成结构化摘要,帮助医生梳理诊疗思路。在决策支持方面,LLMs 能够基于医学知识库和实时患者数据回答复杂医学问题,为医生提供诊断建议或治疗方案参考。例如,输入患者症状和体征后,模型能结合临床指南生成可能的诊断列表及后续检查项目建议。然而,LLMs 在临床实践中的应用仍存在局限性。一方面,受训练数据质量影响,LLMs 在真实临床场景中的诊断准确率仍有待提高。研究显示,ChatGPT 在临床工作流程中的平均诊断准确率为 71.8%,而初始鉴别诊断准确率仅为 60.3%^[5,8-9]。另一方面,LLMs 对罕见病和新发疾病的识别能力有限,且缺乏动态更新医学知识的机制^[10]。此外,由于医疗隐私保护和数据共享限制,LLMs 难以获得充足的临床数据进行专项训练或微调,进一步制约了其应用效果。

2.1.2 MMLs 在医学图像诊断中的应用与挑战

近年来,MMLs 在医学图像诊断领域取得了显著进展。这些模型能够通过学习大量标注数据,自动

识别和分析医学图像中的病灶特征,为临床医生提供精准的辅助诊断。例如,CHIEF 模型由哈佛大学开发,专门用于组织病理学成像评估,在 19 种癌症的诊断中准确率高达 94%^[11];PathChat 系统是哈佛医学院开发的视觉语言通用 AI 助手,其诊断准确率超过 80%,甚至超越 GPT-4V^[12],该系统不仅能分析病理图像,还能提供诊断推理、辅助检测以及疾病风险因素、预后和治疗建议;SkinGPT-4 基于 MMLs,可通过图像分析提供交互式皮肤病诊断和治疗建议,其诊断结果与皮肤科医生的一致性达到 78.76%^[13];MMLs 在 B 细胞淋巴瘤(B-NHLs)亚型分类和治疗预测任务中的准确率为 92.4%,相比传统的 COO 分类器(准确率为 70%—80%),提升幅度达 12%—22%^[14](表 1)。可见,MMLs 显著提升了疾病检测的效率和准确性,尤其在病理切片、X 光、磁共振成像(MRI)和计算机断层扫描(CT)等影像分析中展现出巨大潜力。

然而,MMLs 在医学图像诊断中的应用仍面临诸多挑战。首先,多模态数据的融合过程较为复杂,不同成像设备的成像原理和参数差异可能导致数据在时间和空间上出现不一致性,进而影响模型的诊断准确性^[15]。其次,现有算法的泛化能力有限,在不同任务中的表现差异较大,如某些算法在微血管浸润识别任务中表现优异,但在脑肿瘤病理分级任务中的性能却相对较差^[16]。此外,尽管多模态医学图像技术已在部分领域(如肾小球疾病分类与乳腺癌的诊断)得到应用,但其在临床实践中的普及程度仍然较低,

表 1 LLMs、MMLs 与传统方法/模型在医学诊断任务中的准确率对比

Table 1 Accuracy of LLMs and MMLs compared with conventional methods/models in medical diagnostic tasks

模型名称 Model name	诊断任务 Diagnosis mission	模型准确率/% Model accuracy/%	对比方法/模型 Compared method/model	对比方法/ 模型准确率/% Compared method/model accuracy/%	提升幅度/% Improved accuracy/%
CHIEF	Certain cancer diagnosis and survival prediction (specific cancer types are not listed)	>80.0	Basic AI model based on pathological images	<70	10—20 ^[11]
PathChat	Pathological diagnosis (based on tissue section images and multimodal input)	81.2	LLaVA 1.5, LLaVA-Med, GPT-4V	50—70	11—30 ^[12]
SkinGPT-4	Diagnosis of skin diseases (based on image analysis)	78.76	MiniGPT-4, ChatGPT, ChatCAD	<50	30 ^[13]
Multimodal models for B-NHLs	Subtype classification and treatment prediction of B-NHLs	92.4	COO classifier	70—80	12—22 ^[14]

这与医疗机构对新技术的接受度、设备投入成本以及医务人员的培训水平密切相关^[17-18]。

2.2 个性化医疗

2.2.1 疾病预测与诊断

个性化医疗的核心是根据患者个体差异(如基因组特征、病史、生活方式等)制定精准的治疗方案。LLMs 和 MMLs 通过处理和分析复杂的异构数据(包括基因组数据、医学影像和电子健康记录),为疾病预测和诊断提供了强有力的支持。这些模型能够整合多种数据来源,帮助医疗专业人员全面评估患者的健康状况,从而制定更精准的治疗方案^[19-20]。例如,Isavand 等^[14]不仅系统梳理了 B 细胞淋巴瘤肿瘤微环境(TME)在肿瘤发生、发展和治疗中的关键作用,还通过 AI 系统整合基因组、转录组、临床及病理成像等多模态数据,构建了 B 细胞淋巴瘤肿瘤微环境的多尺度图谱;Liu 等^[21]采用 Kassandra 算法解析 RNA 测序数据,利用对抗正则化变分图自动编码器分析细胞间空间关系,并结合卷积神经网络(CNN)进行细胞分割和类型识别,从而深入探讨了 TME 对治疗响应的影响。此外,帕金森病预测研究表明,基于多模态数据(包括临床数据、遗传学信息和转录组学数据)构建的模型在预测性能上显著优于单一数据类型的模型。Makariou 等^[22]的研究表明,嗅觉功能测试(University of Pennsylvania Smell Identification Test, UPSIT)和遗传学多态性评分(Polygenic Risk Score, PRS)是帕金森病预测的关键指标,该研究通过无偏网络分析识别出 13 个与帕金森病相关的基因簇,这些基因可为该疾病的治疗提供新的靶点。

2.2.2 药物选择与剂量优化

LLMs 和 MMLs 在个性化医疗的药物选择与剂量优化方面展现出巨大潜力。通过整合基因组数据、电子健康记录和患者症状等多模态数据,这些模型能够解析基因变异与药物反应之间的复杂关联,从而为患者提供个性化的用药方案。以部署在上海东方医院的 MedGo 模型为例,该模型是针对中国医疗环境优化的专业 LLMs,具有强大的多模态数据整合与药物剂量优化功能。MedGo 模型通过自然语言处理(NLP)技术解析文本数据(如电子健康记录、医生笔记),并利用计算机视觉技术处理医学影像数据(如 CT、MRI),实现对患者健康状况的全面评估。具体而言,该模型能够综合患者的个体特征(包括年龄、体重、基因型、病史等)和实时临床数据推荐最佳药物剂量方案,同时通过分析药物相互作用、不良反应史和过敏信息,确保治疗方案的安全性和有效性。研究表明,LLMs 可根据患者的基因型数据调整药物剂量,从而提升疗效并降低不良反应风险^[23]。在临床实践中,MedGo 模型通过结合基因检测结果和治疗反应数据优化抗癌药物剂量,显著提高了治疗效果,同时减少了副作用发生率^[24]。

2.3 慢性病管理

2.3.1 多模态数据整合与疾病监测

慢性病(如糖尿病、高血压、心血管疾病)的管理需要持续监测患者的多模态数据,包括电子健康记录、可穿戴设备数据采集的生理参数(如心率、血压、血糖水平)、医学影像(如超声、CT)数据以及生活方式(如饮食、运动、睡眠)数据。MMLs 能够整合这些

异构数据源,通过构建统一的表示空间对不同模态数据进行处理和分析,从而实现对患者健康状况的全面监测。例如,基于可穿戴设备数据和电子健康记录构建的 MMLs 可以实时监测患者的血糖波动情况,并结合历史数据预测血糖变化趋势,为糖尿病患者提供动态监测和早期预警支持^[25]。在心血管疾病管理中, MMLs 通过整合心电图 (Electrocardiogram, ECG) 数据、可穿戴设备记录的心率波动数据和医学影像数据,能够实时识别心律失常等异常情况,并评估潜在的心血管风险,为临床干预提供决策支持^[26]。

2.3.2 个性化健康干预与生活方式指导

LLMs 在慢性病管理中的另一重要应用是提供个性化健康干预和生活方式指导。通过综合分析患者的病史、体检报告和生活方式数据, LLMs 能够生成个性化的健康建议^[27]。以高血压患者为例, LLMs 可根据患者的饮食日志和动态血压监测数据,提供具体的饮食调整建议(如减少钠盐摄入量、增加富钾食物)以及个性化的运动处方(包括有氧运动类型、强度和时长)^[26]。

2.3.3 风险预测与长期管理

LLMs 和 MMLs 通过分析患者的长期健康数据,能够识别潜在的疾病风险和健康恶化趋势。具体而言, MMLs 可以整合患者的基因组数据、生活方式参数和环境因素,预测糖尿病患者发生并发症的风险,并生成相应的预警提示^[28]。同样地, LLMs 通过对电子健康记录进行深度语义分析,能够发现患者可能忽略的健康隐患,如未充分记录的症状或潜在的药物不良反应,并及时向医患双方发出提醒。此外, LLMs 还能协助医生制定个性化的长期管理方案,通过持续的行为干预帮助患者改善生活习惯,从而有效延缓疾病进展并提高生活质量^[29]。

3 技术挑战

除上述应用场景中提到的数据质量和多模态融合的具体挑战外, LLMs 和 MMLs 在医疗领域中的应用还需突破以下系统性技术障碍。

3.1 模型泛化与跨模态协同

尽管现有研究已证明 LLMs 和 MMLs 在特定场景下的有效性^[11-13],但跨模态的深度协同仍存在技术瓶颈。不同模态间的信息互补机制尚未完全明确,例如影像特征与自然语言描述的语义对齐可能因疾病类型差异而失效^[30]。此外,模型对医疗数据分布的敏感性导致其在跨机构、跨地域应用时性能显著下

降,因此需要建立更鲁棒的特征提取框架以提升泛化能力。

3.2 模型可解释性与透明性

医疗领域对模型的可解释性要求极高,而 LLMs 和 MMLs 的“黑箱”特性使得其决策过程难以被完全理解。这种透明性不足可能导致医生对模型结果缺乏信任,从而阻碍其临床应用。因此,如何提升模型的可解释性,使其能够清晰展示决策依据,是当前亟待解决的关键问题。

3.3 隐私保护与数据安全

医疗数据涉及高度敏感的个人隐私信息,因此在模型训练和应用过程中必须严格遵守隐私保护法规,如《通用数据保护条例》(GDPR)和《健康保险可携带性和问责法案》(HIPAA)。然而,大规模的模型训练依赖于海量高质量数据,如何在满足隐私保护要求的同时获取足够的数据,仍是一项亟待应对的挑战^[31]。此外,模型在实际应用中还需防范潜在的数据泄露和安全风险,确保患者信息的安全性。

3.4 临床部署与实际适用性

将 LLMs 和 MMLs 应用于临床实践面临多重障碍,包括高计算资源需求、实时性要求、系统兼容性问题^[32],具体表现如下:(1)高计算资源需求。模型运行依赖高性能硬件(如 GPU 或 TPU 集群),同时伴随高昂的维护成本和能源消耗,这对资源有限的基层医院和发展中国家医疗机构构成显著负担。(2)实时性要求。在急诊或手术室等高压场景中,系统延迟可能直接影响诊疗效率和患者安全,因此对模型的响应速度提出极高要求。(3)系统兼容性。许多医疗机构仍在使用定制化或老旧的电子健康记录系统,这些系统通常缺乏统一的数据结构和标准化接口,导致与现代 AI 系统的集成困难。为实现无缝对接,需开发复杂接口以弥合技术差异,并对现有数据进行标准化处理,以确保数据格式、语义和结构的一致性。这一过程不仅增加技术开发工作量,还可能涉及系统改造或升级,进一步提高实施成本和技术门槛。(4)隐私法规约束。涉及患者敏感数据的模型部署必须严格遵循 GDPR、HIPAA 等法规,这在一定程度上增加了部署的复杂性。(5)模型更新与扩展成本。随着临床数据的积累和医学知识的快速更新,模型需频繁更新以保持准确性。然而,更新过程通常需要重新训练或微调,不仅耗时耗力,还可能影响系统稳定性^[33]。

上述技术挑战与第 2 节的应用场景限制共同构成双重阻碍:应用层挑战(如诊断准确率问题、设备兼

容性障碍)直接制约技术落地,而系统性技术障碍(如模型泛化能力、可解释性等)则影响整个医学 AI 生态的发展。这种分层结构表明,医学 AI 的应用突破不仅需要解决具体场景的实践问题,还需要攻克基础性技术瓶颈。

4 展望

LLMs 和 MMLs 在临床医学领域展现出巨大潜力。通过整合多模态医疗数据和智能分析技术,这些模型显著推动了精准诊断、个性化治疗及慢性病管理等重要领域的发展。尽管目前仍面临数据质量参差、模型可解释性不足和患者隐私保护等挑战,但随着算法技术的持续优化和相关政策的逐步完善,这些模型必将在未来医学实践中发挥更加重要的作用,为提升医疗服务的效率和质量提供强有力的技术支撑。为充分发挥这些模型的临床价值,未来研究应重点关注以下 4 个方面:(1)算法性能提升方面,需要着力提高数据质量并增强多模态融合能力,开发具有更强泛化能力的预处理和多模态学习框架,同时利用自监督学习等技术缓解医疗数据标签稀缺的问题。(2)可解释性方面,应重点探索能够透明化模型决策过程的创新算法和可视化工具,以增强临床医生对模型输出的信任度^[34]。(3)隐私保护方面,可采用联邦学习和差分隐私等前沿技术,在确保数据安全的前提下实现医疗信息的高效共享与利用^[35]。(4)临床转化应用方面,需优化模型的轻量化设计,加强真实医疗场景下的系统性验证,确保模型在复杂临床环境中的稳定性和适用性^[36]。实现这些目标的关键在于建立高效的跨学科协作机制,促进医学专家、数据科学家和工程师之间的深度合作。只有通过协同创新,才能充分发挥 AI 技术在医疗领域的变革性潜力。

参考文献

- [1] CHEEMA B, PANDIT J. AI and heart failure: present state and future with multimodal large language models [J]. *Journal of the American College of Cardiology: Advances*, 2024, 3(9):101029.
- [2] OMIYE J A, GUI H, REZAEI S J, et al. Large language models in medicine: the potentials and pitfalls: a narrative review [J]. *Annals of Internal Medicine*, 2024, 177(2):210-220.
- [3] 陈润生. 医疗大数据结合大语言模型的应用展望[J]. *四川大学学报(医学版)*, 2023, 54(5):855-856.
- [4] KOWALEWSKI K F, RODLER S. Large language models in science [J]. *Urologie*, 2024, 63(9):860-866.
- [5] SINGHAL K, AZIZI S, TU T, et al. Large language models encode clinical knowledge [J]. *Nature*, 2023, 620(7972):172-180.
- [6] YIN S K, FU C Y, ZHAO S R, et al. A survey on multimodal large language models [J]. *National Science Review*, 2024, 11(12):nwae403.
- [7] QIU P C, WU C Y, ZHANG X M, et al. Towards building multilingual language model for medicine [J]. *Nature Communications*, 2024, 15(1):8384.
- [8] VAN VEEN D, VAN UDEN C, BLANKEMEIER L, et al. Adapted large language models can outperform medical experts in clinical text summarization [J]. *Nature Medicine*, 2024, 30(4):1134-1142.
- [9] RAO A, PANG M, KIM J, et al. Assessing the utility of ChatGPT throughout the entire clinical workflow: development and usability study [J]. *Journal of Medical Internet Research*, 2023, 25:e48659.
- [10] CAN E, ULLER W, VOGT K, et al. Large language models for simplified interventional radiology reports: a comparative analysis [J]. *Academic Radiology*, 2025, 32(2):888-898.
- [11] WANG X Y, ZHAO J H, MAROSTICA E, et al. A pathology foundation model for cancer diagnosis and prognosis prediction [J]. *Nature*, 2024, 634(8035):970-978.
- [12] LU M Y, CHEN B, WILLIAMSON D F K, et al. A multimodal generative AI copilot for human pathology [J]. *Nature*, 2024, 634(8033):466-473.
- [13] ZHOU J X, HE X N, SUN L Y, et al. Pre-trained multimodal large language model enhances dermatological diagnosis using SkinGPT-4 [J]. *Nature Communications*, 2024, 15(1):5649.
- [14] ISAVAND P, AGHAMIRI S S, AMIN R. Applications of multimodal artificial intelligence in non-Hodgkin lymphoma B cells [J]. *Biomedicine*, 2024, 12(8):1753.
- [15] 文含, 赵莹, 蔡秀定, 等. 一种具有域自适应反标准化的多模态医学图像对比学习算法[J]. *生物医学工程学杂志*, 2023, 40(3):482-491.
- [16] 邢素霞, 方俊泽, 鞠子涵, 等. 基于记忆驱动的多模态医学影像报告自动生成研究[J]. *生物医学工程学杂志*, 2024, 41(1):60-69.
- [17] 龙楷兴, 翁丹仪, 耿舰, 等. 基于多模态多示例学习的免疫介导性肾小球疾病自动分类方法[J]. *南方医科大学学报*, 2024, 44(3):585-593.
- [18] 姜良, 张程, 曹慧, 等. 基于深度学习的乳腺病理图像诊

- 断研究进展[J]. 生物医学工程学杂志, 2024, 41(5): 1072-1077, 1084.
- [19] ALSAAD R, ABD-ALRAZAQ A, BOUGHORBEL S, et al. Multimodal large language models in health care: applications, challenges, and future outlook [J]. *Journal of Medical Internet Research*, 2024, 26: e59505.
- [20] KHALIGHI S, REDDY K, MIDYA A, et al. Artificial intelligence in neuro-oncology: advances and challenges in brain tumor diagnosis, prognosis, and precision treatment [J]. *NPJ Precision Oncology*, 2024, 8(1): 80.
- [21] LIU Y Y, ZHOU X X, WANG X. Targeting the tumor microenvironment in B-cell lymphoma: challenges and opportunities [J]. *Journal of Hematology & Oncology*, 2021, 14(1): 125.
- [22] MAKARIOUS M B, LEONARD H L, VITALE D, et al. Multi-modality machine learning predicting Parkinson's disease [J]. *NPJ Parkinson's Disease*, 2022, 8(1): 35.
- [23] JOHNSON K B, WEI W Q, WEERARATNE D, et al. Precision medicine, AI, and the future of personalized health care [J]. *Clinical and Translational Science*, 2021, 14(1): 86-93.
- [24] ZHANG H T, AN B. MedGo: a Chinese medical large language model [Z/OL]. (2024-10-27) [2025-01-08]. <https://arxiv.org/abs/2410.20428>.
- [25] THAO P N M, DAO C T, WU C W, et al. MEDFuse: multimodal EHR data fusion with masked lab-test modeling and large language models [C]//Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. New York: ACM, 2024: 3974-3978.
- [26] PAPASTRATIS I, STERGIOULAS A, KONSTANTINIDIS D, et al. Can ChatGPT provide appropriate meal plans for NCD patients? [J]. *Nutrition*, 2024, 121: 112291.
- [27] MONDAL H, DASH I, MONDAL S, et al. ChatGPT in answering queries related to lifestyle-related diseases and disorders [J]. *Cureus*, 2023, 15(11): e48296.
- [28] IQBAL J, CORTÉS JAIMES D C, MAKINENI P, et al. Reimagining healthcare: unleashing the power of artificial intelligence in medicine [J]. *Cureus*, 2023, 15(9): e44658.
- [29] HALEEM A, JAVAID M, SINGH R P. Exploring the competence of ChatGPT for customer and patient service management [J]. *Intelligent Pharmacy*, 2024, 2(3): 392-414.
- [30] BELLANDA V C F, SANTOS M L D, FERRAZ D A, et al. Applications of ChatGPT in the diagnosis, management, education, and research of retinal diseases: a scoping review [J]. *International Journal of Retina and Vitreous*, 2024, 10(1): 79.
- [31] 刘喜恩, 刘少辉, 周开银, 等. 大语言模型基本医学能力及其在脑血管病等临床应用上的研究进展[J]. *中国卒中杂志*, 2024, 19(6): 614-620.
- [32] MENG X B, YAN X Y, ZHANG K, et al. The application of large language models in medicine: a scoping review [J]. *iScience*, 2024, 27(5): 109713.
- [33] ZHANG D Z, YU Y H, DONG J H, et al. MM-LLMs: recent advances in MultiModal large language models [Z/OL]. (2024-01-24) [2025-01-08]. <https://arxiv.org/abs/2401.13601>.
- [34] HOLZINGER A, SARANTI A, MOLNAR C, et al. Explainable AI methods: a brief overview [M]//HOLZINGER A, GOEBEL R, FONG R, et al. *xxAI: beyond explainable AI*. Cham: Springer International Publishing, 2022: 13-38.
- [35] 陈晶, 彭长根, 谭伟杰, 等. 基于差分隐私和秘密共享的多服务器联邦学习方案[J]. *信息安全*, 2023, 23(7): 98-110.
- [36] JIN Y Z, LI J, LIU Y X, et al. Efficient multimodal large language models: a survey [Z/OL]. (2024-03-17) [2025-01-08]. <https://arxiv.org/abs/2405.10739>.

Application and Challenges of Large Language Models and Multimodal Models in Clinical Medicine

ZOU Yuan, TAN Yuping^{* *}

(The Second Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, 530007, China)

Abstract: Large Language Models (LLMs) and Multimodal Models (MMLs) integrate text, images, and audio data to provide innovative technical support for clinical diagnosis, personalized treatment, and chronic disease management. The technical foundations of LLMs and MMLs and their applications in clinical medicine, including clinical diagnosis and decision support, personalized treatment, and chronic disease management, are systematically reviewed. The potential and limitations of LLMs and MMLs in enhancing diagnostic accuracy, optimizing treatment plans, and improving patient health management are explored. Furthermore, the technical challenges faced by LLMs and MMLs in the medical domain, such as limited generalization capability, issues with interpretability and transparency, risks related to privacy protection and data security, and compatibility challenges with existing medical systems, are examined. These challenges are highlighted as key barriers to the implementation and widespread adoption of these technologies. Finally, the future directions in model optimization, data integration, privacy protection are prospected, and it is proposed that technological innovation and multidisciplinary collaboration are needed in advancing the application of Artificial Intelligence (AI) in medicine. The paper provides a reference for improving healthcare service efficiency and quality.

Key words: Large Language Models (LLMs); Multimodal Models (MMLs); clinical big data; clinical decision support; personalized medicine

责任编辑:陆媛峰



微信公众号投稿更便捷

联系电话:0771-2503923

邮箱:gxkx@gxas.cn

投稿系统网址: <http://gxkx.ijournal.cn/gxkx/ch>