

人工智能大模型驱动磁共振多序列融合与精准诊疗研究

曾 畅

常德市第一中医医院 湖南 常德 415000

【摘 要】：磁共振成像（magnetic resonance imaging, MRI）具有多序列、多参数和高软组织分辨率等优势，但设备、场强和扫描协议差异也带来明显的数据异质性，传统人工智能模型因而常局限于特定序列、器官或任务。视觉基础模型、视觉—语言模型和医学大语言模型借助大规模预训练、三维表征及影像—文本联合学习，逐步用于图像重建、多序列融合、病灶识别、分子状态预测、报告生成和疗效评价。本文梳理 MRI 大模型的技术基础及其在精准诊疗中的应用，并讨论序列缺失、合成影像失真、跨中心泛化不足、无依据内容生成和临床证据有限等问题。现阶段，MRI 大模型更适合承担图像质量控制、多序列信息整合和辅助决策，其临床价值仍需多中心前瞻性研究验证。

【关键词】：人工智能大模型；磁共振成像；多序列融合；影像表型；精准诊疗

DOI:10.12417/2811-051X.26.12.009

1 引言

MRI 可通过 T1 加权、T2 加权、液体衰减反转恢复、弥散加权、灌注及增强成像等序列，从组织结构、含水状态、细胞密度、血流动力学和屏障完整性等不同层面呈现病变。多序列信息有助于病灶检出、鉴别诊断和疗效评价，但也伴随配准偏差、信号尺度差异、扫描时间延长和序列缺失等问题。即使采用同名序列，场强、线圈、重复时间、回波时间和重建算法不同，也可能改变图像对比度和数据分布。

传统深度学习模型多围绕特定器官、序列或任务建立，依赖大量精细标注，设备和采集参数变化后性能也容易下降。基础模型则利用未标注或弱标注数据学习可迁移表征，视觉—语言模型进一步连接 MRI 影像、放射学报告和临床资料，使模型由分类、分割逐步延伸至影像描述、风险预测和临床信息整合^[1-4]。对 MRI 而言，大模型的价值不在参数规模本身，而在于面对序列缺失、协议变化和中心差异时，能否稳定识别具有临床意义的影像表型。

2 MRI 大模型与多序列融合的技术基础

MRI 大模型主要包括视觉基础模型、视觉—语言模型、重建与增强模型，以及与医学知识资源连接的大语言模型。此类模型常采用自监督学习、对比学习或影像—文本联合训练，以减少对精细标注的依赖。BrainIAC 基于大规模多参数脑 MRI 预训练，并在序列分类、脑龄估计、肿瘤分割、分子状态预测和生存分析等任务中评估迁移性能^[2]。BrainSegFounder 利用三维多模态脑 MRI 预训练，在有限标注条件下提高了脑肿瘤和卒中病灶分割的数据利用效率^[3]。

多序列融合并非简单叠加 T1、T2、FLAIR、DWI 和增强序列。模型需要处理空间配准、分辨率匹配、信号归一化和序

列权重分配，而不同序列承担的信息也并不相同。DWI 及表观弥散系数主要反映水分子扩散，增强序列更多体现灌注和血管通透性。关键序列缺失时，模型应降低结论确定性，而不能用其他序列无依据地“补全”信息。

Decipher-MR 利用大规模 MRI 及配对报告进行三维视觉—语言预训练，可用于疾病分类、解剖定位和跨模态检索^[4]。MRI 信号同时受脉冲序列、场强和组织弛豫特性影响，仅依据图像外观训练，模型可能把设备或协议差异误作疾病特征。为减轻这一问题，模型设计可引入 k 空间数据、扫描参数、数据一致性约束和解剖先验。跨中心影像协调也不能只追求图像分布一致，还需尽量保留病灶形态、组织信号和定量指标，并通过独立验证判断协调过程是否损伤真实生物学信息。

3 多序列融合在精准诊疗中的应用

病灶分析更能体现多序列融合的价值。模型可联合评估病变位置、体积、边界、信号特征、弥散受限、强化方式以及水肿和坏死范围，把分散在不同序列中的信息整合为较完整的影像表型。相关应用已涉及脑肿瘤、脑卒中、神经退行性疾病，以及前列腺癌、乳腺癌和直肠癌等。这样的任务设置更接近放射科医师综合判读不同序列的过程，也有利于更完整地描述病灶性质、分级依据和局部侵犯情况。类似思路还可延伸至心脏 MRI 的心肌结构与功能评价、肌肉骨骼 MRI 的软骨和韧带分析，以及腹盆部疾病的局部分期。不同临床场景依赖的关键序列不同，对错误结果的容忍度也不同，通用模型仍需结合器官特异性模块并进行任务级适配。

多参数 MRI 还能提取肉眼难以稳定辨认的影像特征，用于预测肿瘤分子亚型、治疗反应和预后。Glio-LLaMA-Vision 联合多参数 MRI 与放射学报告，对成人弥漫性胶质瘤进行分

作者简介：曾畅，男（1983—），大学本科，副主任技师，研究方向：磁共振技术。

项目基金：2024 年度常德市科技创新指导性项目，2024ZD80。

子状态预测和报告生成,使MRI分析由形态描述进一步延伸至分子表型预测^[5]。这类结果可为术前风险分层、穿刺区域选择和讨论提供补充信息,但不能替代病理检查和基因检测。影像分子预测仍需在独立外部队列中验证稳定性、校准度和临床增益。

视觉—语言模型还可汇总不同序列中的病灶表现,辅助生成影像所见、诊断印象和纵向比较内容。它的主要价值在于提高信息组织效率并减少重复书写,而不是代替医师完成诊断。虚构病灶、定位错误、否定关系混淆和关键序列遗漏,都是目前需要警惕的问题。MRI报告应尽可能注明关键判断来自哪些序列;当序列缺失、图像质量不足或不同序列证据不一致时,模型应保留不确定性。治疗前后的纵向比较还需区分真实病理变化与设备更换、扫描参数差异或配准误差造成的表现变化。

从实际流程看,大模型更适合承担图像质量控制、病灶定量、多序列信息整合、报告草拟和风险提示。遇到关键序列缺失、结果可靠性较低,或影像与临床资料明显不一致的病例,系统应提示补充检查或转入人工复核,而不宜自动输出确定性结论。最终诊断和治疗决策仍需医师结合病理、实验室结果和患者整体情况作出。

4 临床转化限制与展望

MRI大模型临床转化绕不开数据异质性。MRI信号缺少统一绝对尺度,设备厂商、场强、接收线圈和扫描协议的差异都会改变图像对比度及数据分布。单中心训练得到的模型即使内部表现良好,换到外部机构后也可能出现明显性能下降。扩大样本量能够缓解部分问题,却不能自动消除病例来源和采集条件形成的偏倚。模型仍需在设备、协议和患者群体中进行独立验证,相关论文也应交代训练数据构成、序列缺失情况和主要错误类型。

参考文献:

- [1] Sun Y,Wang L,Li G,et al.A foundation model for enhancing magnetic resonance images and downstream segmentation,registration and diagnostic tasks.Nat Biomed Eng,2025,9:521-538.
- [2] Tak D,Garomsa BA,Zapaishchykova A,et al.A generalizable foundation model for analysis of human brain MRI.Nat Neurosci,2026,29:945-956.
- [3] Cox J,Liu P,Stolte SE,et al.BrainSegFounder:Towards 3D foundation models for neuroimage segmentation.Med Image Anal,2024,97:103301.
- [4] Yang Z,DSouza N,Megyeri I,et al.Decipher-MR:a vision-language foundation model for 3D MRI representations.NPJ Digit Med,2026,9:419.
- [5] Park YW,Kang M,Ryu H,et al.A robust vision language model for molecular status prediction and radiology report generation in adult-type diffuse gliomas.NPJ Digit Med,2026,9:406.
- [6] Tejani AS,Klontzas ME,Gatti AA,et al.Checklist for Artificial Intelligence in Medical Imaging:2024 update.Radiol Artif Intell,2024,6:e240300.
- [7] Lekadir K,Feragen A,Fofanah AJ,et al.FUTURE-AI:international consensus guideline for trustworthy and deployable artificial intelligence in healthcare.BMJ,2025,388:e081554.

序列缺失、合成影像失真和无依据内容生成,则直接影响诊断可靠性。较理想的模型输出不仅要给出结果,还应说明主要判断依赖哪些序列、病灶位于何处、有哪些支持性征象,并对结论可靠性作出估计。证据不足时,模型不宜使用过度确定的诊断语言。评价也不能停留在曲线下面积、Dice系数、峰值信噪比和结构相似度等技术指标。临床应用更应关注微小病灶能否保留、重大诊断错误和重复扫描是否减少,以及模型是否真正缩短判读时间、改善治疗决策或患者结局。筛查、分割、报告生成和预后预测承担的风险不同,性能门槛和人工复核条件也应分别设定。

CLAIM 2024 强调医学影像人工智能中的数据透明性、外部验证和错误分析;FUTURE-AI 则从公平性、通用性、可追溯性、可用性、鲁棒性和可解释性等方面提出可信医疗人工智能要求^[6-7]。后续模型设计需要更充分地纳入MRI成像物理、序列关系和临床知识,并在多中心前瞻性研究及真实工作流程中验证净获益。临床评价还应比较模型单独运行、医师独立判读和人机协同三种模式,明确收益究竟来自诊断准确性、工作效率还是信息完整性的改善,同时记录误诊、漏诊和额外检查等潜在损害。

模型进入临床后,版本信息、输入输出记录和医师修改痕迹都应保留,并按照临床风险设置复核流程。开发者、部署机构和临床使用者的责任边界也需事先明确。MRI大模型的价值不能用合成序列数量或报告篇幅衡量,更关键的是,当设备、协议和序列构成发生变化时,模型能否继续稳定识别与诊断、分型和疗效评价相关的影像表型。临床应用最终应建立在数据标准化、独立外部验证、证据可追溯和医师监督的基础上,评价重点也应回到诊断质量、工作流程和患者结局,而不是单纯追求算法指标提高。