

## • 专论:大语言模型及其医学应用 •

**编者按:** 近期 DeepSeek 在大语言模型领域的突破引发了广泛关注。大语言模型基于深度神经网络架构,能够对复杂医疗数据进行高效处理与模式识别,为智慧医疗提供新型决策支持。其在医学领域的应用正逐步从理论探索走向实践,如智能问诊、影像解读、智慧病案、康复管理等,显著提高诊断准确率、病历处理效率,并为医学教育提供支持。然而医疗领域涉及数据复杂,对错误容忍度低,对可解释性要求高,医学大语言模型面临着与通用人工智能系统完全不同的技术挑战和行业需求,如技术瓶颈突破、临床转化路径构建、伦理及法律监管体系完善等。本期专论着眼于大语言模型及其医学应用,内容包括 DeepSeek 的技术创新、医学大语言模型对医疗服务模式的重构,以及大语言模型在医疗质量控制、出院小结生成“幻觉”抑制等方面的应用实践,以期利用大语言模型提高医疗服务质量和效率提供参考。

# DeepSeek 与医学大语言模型:技术创新与医疗服务模式重构\*

闫海荣 江 瑞 张学工 王存库

(清华大学自动化系生命基础模型实验室 北京 100084)

**[摘要]** **目的/意义** 从 DeepSeek 出发,探讨医学大语言模型的技术创新,以应对医学领域独有的技术难题和挑战。**方法/过程** 重点分析医疗人工智能在多模态数据融合、医生思维方式适配以及医疗决策高风险性的技术要求等方面的突破路径,并展望未来医学大语言模型如何推动数智医院建设与医疗服务模式重构。**结果/结论** 随着人工智能技术的不断发展,医学大语言模型将在未来医疗实践中发挥越来越重要的作用。通过多模态数据融合、因果推理的引入、模型可解释性的提升,大语言模型不仅能提升诊断效率和准确性,还将为数智医院的建设和医疗服务模式重构提供强有力的技术支持。

**[关键词]** 大语言模型; 医学服务模式; DeepSeek

**[中图分类号]** R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2025.02.001

**DeepSeek and Medical Large Language Model: Technological Innovation and Reconstruction of Medical Service Mode**

LYU Hairong, JIANG Rui, ZHANG Xuegong, WANG Cunku

Life Foundation Model Laboratory, Department of Automation, Tsinghua University, Beijing 100084, China

**[Abstract]** **Purpose/Significance** Starting from DeepSeek, the paper explores the technological innovation and challenges of medical large language models, in order to address the unique technical difficulties and challenges in the medical field. **Method/Process** It focuses on analyzing the breakthrough path of medical artificial intelligence (AI) in multimodal data fusion, adaptation of doctors' thinking

**[修回日期]** 2025-02-15

**[作者简介]** 闫海荣, 博士, 副研究员, 发表论文 80 余篇。

**[基金项目]** 国家重点研发计划 (项目编号: 2022YFF1202400)。

patterns, and technical requirements for high - risk medical decision - making, and looks forward the construction of the smart hospitals and the reconstruction of medical service modes using medical large language models in the future. **Result/Conclusion** With the continuous development of AI technology, medical large language models will play an important role in future medical practice. By integrating multimodal data, introducing causal reasoning, and improving model interpretability, large language models can not only enhance diagnostic efficiency and accuracy, but also provide strong technical support for the construction of digital and smart hospitals and the reconstruction of medical service modes.

[**Keywords**] large language model; medical service mode; DeepSeek

## 1 引言

人工智能 (artificial intelligence, AI) 在医疗领域的应用正逐步从理论探索走向实践, 深度学习和大语言模型 (large language model, LLM, 以下简称大模型) 的引入, 为医学数据分析和疾病诊断提供了新方向。随着医疗行业对 AI 技术的关注日益增强, AI 在辅助诊疗、疾病预防、个性化治疗等方面展现出巨大潜力<sup>[1]</sup>。尤其是 DeepSeek 的出现, 引发了医疗体系的大量关注和广泛应用。然而, 医疗 AI 的应用远非简单地将现有大模型算法应用于医学数据。医学领域具有独特的技术难题和挑战, 需要根据其特殊要求进行算法创新、数据融合与应用场景定制。在医疗 AI 的发展过程中, 如何有效整合和分析不同来源、不同形式的多模态数据, 确保 AI 决策的可信度与可解释性, 以及将 AI 应用于实际医疗决策, 依然是亟待解决的核心问题。

本文从 DeepSeek 出发, 探讨医学大模型的技术创新与挑战, 重点分析其在多模态数据融合、医生思维方式适配以及医疗决策高风险性的技术要求等方面的突破路径, 并展望未来医学大模型如何推动数智医院建设与医疗服务模式重构。本文中所提的 DeepSeek, 在某些情景也泛指一般的医学大模型, 包括多模态模型。

## 2 DeepSeek 核心技术优势

DeepSeek 在模型架构设计、训练效率优化、推理能力强化及知识迁移等方面展现出显著优势, 通过不同版本 (DeepSeek - V3、DeepSeek - R1 等)

的迭代, 构建了行业领先的技术壁垒。

### 2.1 高效架构设计: 平衡性能与成本

2.1.1 多头潜注意力 DeepSeek 的多头潜注意力 (multi - head latent attention, MLA) 技术通过智能压缩与动态计算优化, 在保持模型能力的同时大幅节省资源<sup>[2]</sup>, 其核心优势可概括如下。(1) 高效记忆存储。MLA 采用优化的注意力计算方式, 在减少计算资源的同时, 提升长文本处理能力。通过高效的存储和计算优化, 使 DeepSeek 在长文本任务中的处理能力远超传统 Transformer 结构。(2) 智能信息筛选。根据任务难度动态调整信息保留比例, 例如数学推理时保留更多细节, 简单对话时减少冗余计算, 降低整体计算量。(3) 硬件深度适配。针对不同显卡 (如 A100/H100) 优化计算流程, 在精度几乎不损失的前提下, 速度相对传统方法提升数倍。该技术使大模型在普通显卡上也能流畅运行, 特别适合处理长文档分析、代码生成等场景, 为 AI 应用普及提供技术基础。

2.1.2 DeepSeekMoE 架构 DeepSeek 的混合专家 (mixture - of - experts, MoE) 架构通过智能任务分配与动态资源调度, 显著提升模型效率, 其核心设计如下。(1) 精细化专家分工。DeepSeekMoE 采用动态专家激活策略, 每个 token 仅激活部分专家, 减少计算冗余, 提高推理效率。(2) 混合并行策略。结合专家并行与数据并行技术, 配合动态路由算法减少跨节点通信, 训练速度提升数倍。(3) 资源优化机制。实时监测硬件负载, 动态调整冗余专家分布, 显著提升资源利用率并减少显存占用。该架构在代码生成、长文本理解等场景中表现突出, 推理速度达传统方案的 2.3 倍, 已广泛应用

于金融、工业等高精度领域。

2.1.3 多令牌预测 DeepSeek 的多令牌预测 (multi-token prediction, MTP) 技术通过分层解码架构与动态路由优化, 突破传统自回归模型的单步预测限制, 其核心机制如下<sup>[2]</sup>。(1) 分层预测。MTP 采用分层预测结构, 主模型负责基础预测, 多个 MTP 模块预测未来 token, 通过共享嵌入层和输出头实现多 token 预测的协同优化。(2) 动态深度调整。MTP 模块通过序列级负载均衡调整 token 分配。(3) 推测解码加速。MTP 通过预计算未来 token 概率分布加速推理, 为长逻辑链任务提供新的效率-精度平衡范式。

## 2.2 训练效率突破: FP8 与分布式优化

一是 FP8 混合精度框架。DeepSeek 是首个在超大规模模型上成功应用 FP8 训练的框架。通过精确的数值缩放策略, 有效降低计算负担, 提高训练效率, 同时保持模型的高精度表现<sup>[2]</sup>。二是 DualPipe 并行算法。该算法是一种面向大规模 MoE 模型的高效分布式训练框架, 其核心目标是通过双向流水线调度与计算-通信深度重叠, 显著降低分布式训练中的通信与显存开销<sup>[2]</sup>。三是极简内存优化。通过 RMSNorm 重计算 (反向传播时动态重算梯度, 节省 33% 激活存储)、EMA 参数 CPU 卸载 (节省 15% 模型状态内存)、嵌入层共享 (消除冗余参数) 及参数分片缓存 (Zero-1 并行下单卡显存降至 41GB), 使 671B 模型训练成本降至 278.8 万 GPU 小时, 显存效率远高于同规模模型<sup>[2]</sup>。

## 2.3 推理能力强化: 强化学习驱动的自我进化

2.3.1 纯强化学习 (DeepSeek-R1-Zero) 传统大模型的推理能力通常需要大量人工标注的监督数据, 但 DeepSeek-R1-Zero 首次验证了无需任何监督微调 (supervised fine-tuning, SFT) 数据, 仅通过纯强化学习 (reinforcement learning, RL) 即可实现推理能力的自主进化<sup>[3]</sup>, 其核心创新如下。(1) 算法框架。采用群组相对策略优化 (group relative policy optimization, GRPO), 通过组内奖励对比优化策略避免传统 RL 中对复杂价值模型的依赖。

(2) 自我进化现象。模型在训练中自发涌现出“反思”“多步验证”等复杂推理行为。例如, 在解决数学方程时模型会主动纠正早期错误步骤。(3) 性能飞跃。在 2024 年美国数学邀请赛任务中, 模型 Pass@1 准确率从初始的 15.6% 提升至 71.0%, 多数投票后更达 86.7%, 与 OpenAI o1 模型持平。然而纯 RL 训练的代价是可读性差与多语言混杂, 限制了实际应用。

2.3.2 多阶段对齐 (DeepSeek-R1) 为解决上述问题, DeepSeek-R1 提出多阶段对齐策略<sup>[3]</sup>。(1) 冷启动阶段。引入数千条人工标注的高质量长推理链数据对基础模型微调, 强制规范输出格式, 提升可读性。(2) 两阶段强化学习。一是推理导向 RL, 结合规则奖励 (答案准确性、语言一致性), 优化数学、编程等结构化任务表现。二是通用对齐 RL, 融入人类偏好奖励模型, 确保模型在开放域任务中的安全性与实用性。(3) 性能对标。DeepSeek-R1 在 MATH-500、Codeforces 等任务上达到与 OpenAI o1 相当的水平, 同时在 MMLU、GPQA Diamond 等知识密集型任务中显著超越前代模型。

## 2.4 知识迁移与生态构建

2.4.1 蒸馏技术 通过将 DeepSeek-R1 生成的 80 万条数据用于微调开源模型 (Qwen、Llama 系列), 实现推理能力的高效迁移<sup>[3]</sup>。一是小模型性能飞跃, DeepSeek 通过高效的知识蒸馏策略, 使小规模模型继承大模型推理能力, 在多个基准测试中展现出显著的性能提升。二是开源贡献, 发布 1.5B 至 70B 的蒸馏模型, 提供低成本、高性能推理解决方案。

2.4.2 数据与工具链 预训练使用 14.8T 高质量多语言数据, 涵盖代码、数学、科学文献等垂直领域。配套开源 DeepSeek 训练工具链支持 FP8 量化、动态负载监控与自动容错, 实现千卡级别训练稳定性<sup>[2]</sup>。

DeepSeek 通过 MLA + MoE 架构实现性能与成本平衡, FP8 与 DualPipe 突破训练效率瓶颈, RL 驱动的自我进化机制提升复杂任务推理能力, 而蒸馏技术则构建了大小模型协同的生态体系。这些技术优

势使其在开源模型中树立标杆,并为通用人工智能时代的模型研发提供可复用的技术范式。

### 3 医疗 AI 的特殊性

医疗 AI 并不是普通的 AI 应用。医疗领域涉及的数据不仅复杂、异质化,而且在应用中对准确性和可靠性的要求极高。医学大模型在此领域面临着与通用 AI 系统完全不同的技术挑战和行业需求。

#### 3.1 医学数据的复杂性

医学数据的复杂性是医疗 AI 面临的最大挑战之一。医学领域的数据具有高度异质性,涉及影像、基因、电子病历、病理报告、生命体征等多个维度。不同类型的数据,不仅结构、存储格式有所不同,其背后所承载的医学信息也具有不同的意义和应用场景。

**3.1.1 多模态数据融合问题** 医疗 AI 的核心挑战之一是如何高效融合多模态数据。这些数据来源各异,格式不一,且信息内容具有一定关联性。以肿瘤的早期诊断为例,影像数据(如 CT、MRI 图像)可以揭示肿瘤的物理表现,基因数据(如基因组学分析)则能够提供肿瘤的分子信息,而电子病历数据(如病史、诊疗过程)则涉及患者的临床表现。这些数据需要通过 AI 模型进行深度融合,才能为医生提供准确、全面的诊断信息。目前的医疗 AI 技术虽然在某些领域取得突破,但多模态数据的有效融合仍然是一个巨大挑战。尤其是在影像和基因数据结合,以及电子病历与影像数据融合方面,依然存在诸多技术难题。

**3.1.2 数据结构化问题** 医学数据往往具有较高的非结构化特征<sup>[4]</sup>。例如,病历文本往往包含大量自由文本内容,医生的描述具有个性化差异和语言模糊性。如何将这些自由文本转换为结构化数据,便于 AI 系统处理,是当前医疗 AI 面临的又一难题。传统自然语言处理技术在该领域的应用仍然处于初级阶段。DeepSeek 等大模型取得了巨大进展,但现阶段其准确性和泛化能力仍有提升空间。为了提升 AI 的准确性,DeepSeek 等大模型正在努力加强对医

学文本的语义理解,尤其是对医学术语和患者健康信息的深度挖掘。这需要结合医学领域的专有知识库,建立医学知识图谱,以帮助 AI 更好地理解 and 推理患者的健康状态。

#### 3.2 医疗决策的高风险性

**3.2.1 错误容忍度低** 医疗 AI 的另一大特殊性在于其决策的高风险性。与其他领域的 AI 应用不同,医疗 AI 的决策往往涉及患者的生命健康。一旦 AI 作出错误决策,可能导致患者死亡或病情加重,造成不可挽回的后果。医疗 AI 的容错率远低于其他应用领域,因此医疗 AI 不仅要具备高效的推理和预测能力,还需要确保极高的准确性、可解释性和可靠性,具备强大的容错机制和可靠的验证体系。对于 DeepSeek 等大模型而言,如何从复杂的医学数据中提取最具代表性的信息,并生成可行的决策方案,成为当前研究的关键。

**3.2.2 可解释性要求高** 医疗 AI 的可解释性是影响其广泛应用的一个重要因素。医生和患者在接受 AI 诊疗建议时,往往需要理解 AI 给出的诊断依据和推理过程。如果 AI 无法提供透明的决策过程,医生和患者就无法对其结果进行信任和验证。因此,医疗 AI 需要具备强大的可解释性,能够清晰地展示其推理路径和决策依据。DeepSeek 等在这方面做了大量尝试。通过引入符号推理和统计学习的融合方法,DeepSeek 不仅能够提供预测结果,还能够解释其推理过程。例如在肿瘤筛查中,DeepSeek 能够说明为什么某一患者更可能患有某种类型的肿瘤,并给出具体证据和生物标志物支持。这种可解释性极大地提高了医生对 AI 的信任和接受度。

#### 3.3 AI 需要适应医生的思维方式

医生的诊疗过程通常依赖循证医学和因果推理,而现有医疗 AI 模型大多基于模式匹配和统计学习。两者之间的差异,导致现有 AI 模型在临床应用中难以完全适应医生思维方式。医生在诊断过程中,往往不仅依赖于数据和模型的预测结果,还结合临床经验、病情进展和患者个体差异进行综合

判断。因此, 医疗 AI 不仅要能够处理大规模数据, 还要能够理解和模拟医生思维过程。目前, DeepSeek 等大模型更多依赖于大数据和模式匹配的方式, 进行疾病预测和风险评估。然而, 医生决策更多依赖于因果推理, 即通过疾病的发病机制、临床表现和患者个体差异做出判断。为了使医疗 AI 能够更好地与医生思维方式对接, 未来的医学大模型需要向基于因果推理的方向发展。如此则 AI 不仅能够预测疾病发生概率, 还能够分析疾病发生的潜在机制, 并为治疗方案提供依据。

## 4 医疗 AI 落地的挑战与技术突破

尽管医疗 AI 的潜力巨大, 但要将其从理论阶段推向临床实践, 仍然面临诸多挑战。与其他领域 AI 应用相比, 医疗 AI 不仅需要解决算法本身技术问题, 还需要与医院实际工作流程紧密对接, 确保其能够在日常的医疗服务中产生实际价值。

### 4.1 从“模型好”到“流程适配好”

医疗 AI 系统的成功不仅仅依赖于模型精度和预测能力, 还需要深度融合到医院实际工作流程中。医疗 AI 不应当仅仅作为一个独立工具存在, 而是应该成为医生日常诊疗的一部分。因此, 如何使 AI 系统与现有医院信息系统无缝对接, 是推动医疗 AI 应用的关键。以影像 AI 为例, 现有影像 AI 系统大多是独立运行的, 医生使用时需要额外上传影像数据并等待 AI 分析结果, 这样的工作流程显然低效且易出错。DeepSeek 作为基础模型之一, 在提升 AI 诊断效率方面有巨大潜力。为了使 DeepSeek 等大模型真正落地并服务于临床, 必须进行流程适配, 确保 AI 能够自动从医院信息系统中提取患者影像 (需要模型有多模态能力)、电子病历等数据, 实时分析并给出结果, 从而减少人为干预的过程, 提高工作效率。此外, AI 系统还应具备自我更新和优化能力。随着患者就诊过程不断推进, 数据也在实时更新, AI 系统应能够根据最新的临床数据进行再学习和优化, 保持与临床实践的同步性。

### 4.2 优化“医生 - AI - 患者”的交互

医疗 AI 不能仅仅是一个“黑箱工具”, 其应与医生、患者形成良好互动。在辅助决策场景中, 医生依赖 AI 提供的信息作出诊疗决策, 也希望了解 AI 作出决策的依据以及推理过程。这就要求医疗 AI 系统不仅仅提供结论性答案, 还需要具备较强的可解释性和透明度, 能够清晰地展示决策背后的数据支持和分析过程。例如, 在癌症诊断过程中, 大模型可以通过多模态数据分析 (包括影像、病理、基因等) 为医生提供初步诊断意见, 但医生往往希望理解 AI 如何从影像学特征、基因标志物等多维度数据中提取出关键信息, 以及如何结合患者临床表现给出最终诊断建议。这种“对话式”的交互方式不仅增强医生信任度, 也帮助医生在 AI 的基础上进行更加个性化和综合性的诊疗决策。对于患者而言, AI 系统还需要提供友好的交互界面和功能, 使其能够理解 AI 推荐的诊疗方案并在必要时进行调整。这不仅提升了患者的就医体验, 也有助于增强患者对医疗 AI 的信任。

### 4.3 数据闭环与持续学习

当前, 很多医疗 AI 模型是静态训练的, 即利用历史数据进行训练, 并在模型部署后不再更新。然而医学知识和患者健康状况是动态发展的, 医疗 AI 必须具备持续学习和优化的能力, 才能在不断变化的临床环境中保持有效性。为了应对医疗领域知识的动态变化, DeepSeek 等模型可以通过构建数据闭环机制进行持续学习。具体来说, 在临床实践中, AI 系统可以根据实时反馈和新的病例数据进行再训练, 优化模型性能。例如, 某患者在接受 AI 辅助诊断后, 医生根据患者具体情况给出了个性化治疗方案, 并在随访过程中反馈了治疗效果。这些新的治疗数据可以作为 AI 系统的输入, 帮助其不断学习和改进推理路径。此外, DeepSeek 等大模型还可以通过联邦学习等技术, 在保证数据隐私和安全的前提下, 汇集不同医院的临床数据, 推动整个医疗系统智能化升级。

## 5 医学大模型的技术创新与实现路径

### 5.1 多模态数据融合

医学大模型的核心挑战之一是如何有效融合不同来源的多模态数据。数据类型的异质性、数据结构的差异性以及数据之间的潜在关联性,使得多模态数据融合成为一个技术难题。医疗 AI 需要突破传统的单一数据源处理方式,通过多模态数据融合,提取出全面、准确的医学信息。多模态大模型在这方面进行了一些创新,通过多模态数据的预处理与特征提取,对不同类型数据进行“统一表示”。例如,对影像数据、病历文本、检验检查数据等进行预处理,将其转化为统一的特征向量进行进一步分析。这种融合方式有效提升了模型在多模态数据下的性能,避免了信息丢失和数据偏倚的问题。然而,多模态数据融合的挑战不仅仅局限于表型空间。表型空间通常包含患者临床数据,如影像、电子病历、生命体征等,这些数据提供了患者疾病表现的外在信息。然而,要实现精准医疗,医学大模型还需要融合干预空间和生物学空间的数据。干预空间涉及医疗干预信息,包括诊疗决策、药物治疗、手术干预等。每种医疗干预对患者的影响是动态的,且常常在短期和长期之间产生复杂变化。因此,如何对干预决策和后续效果进行建模,是医学大模型需要解决的一个关键问题。研究者希望大模型通过持续的临床数据反馈和患者随访数据整合,不仅在表型空间中进行信息融合,也试图通过对医疗干预的长期效果进行建模,增强对患者治疗路径的优化。生物学空间是指患者基因组数据、分子层面的疾病特征以及细胞水平的生物学过程。基因组学、蛋白质组学等生物学数据具有高维度、高复杂性,其处理和分析在医疗 AI 中面临巨大挑战。将基因数据与临床数据(如影像、电子病历等)深度结合,不仅能通过基因突变信息进行疾病风险预测,还能基于患者的个体基因信息优化治疗方案,为精准医疗提供支持。通过有效的多模态数据融合,医学大模型不仅能在表型空间提供准确诊断,还能在干预空间和生物学空间有效建模,推动精准

医疗发展。这种多维度的融合,不仅提升了模型预测精度,也为医疗决策提供了更为丰富的数据支持。

### 5.2 因果推理的引入

与传统模式识别和统计推断不同,医疗 AI 需要具备因果推理能力。医生在作出诊断和治疗决策时,往往不仅依赖于相关性分析,还需要考虑疾病发生的因果机制。而目前大多数医疗 AI 系统主要基于模式匹配和相关性推理,缺乏对因果关系的深入理解。DeepSeek - R1 以及部分带有推理能力的大模型,正在尝试通过引入因果推理的方法,提升模型的医学推理能力。理想的医学大模型能够结合循证医学和因果推理模型,通过构建因果图和推理网络,模拟疾病发生的因果关系。例如,大模型不仅仅分析影像数据与基因数据之间的相关性,还能够基于患者临床表现和治疗反应,推理疾病发展的潜在原因,并给出合理的治疗建议。

### 5.3 模型可解释性的提升

医疗 AI 的决策不仅要准确,还要具备高度的可解释性。可解释性是医疗 AI 能够广泛应用于临床的前提。尤其是在肿瘤、心脑血管等重大疾病的诊疗中,医生必须理解 AI 作出某一诊断或治疗推荐的依据,从而作出最终决策。通过引入符号推理方法和注意力机制,在生成最终诊断或治疗建议时,能够明确指出其推理路径和支持证据,提升模型可解释性。例如,在肿瘤诊断中,模型不仅提供癌症的概率预测,还能指出影像中最关键的区域、基因突变和临床数据,帮助医生理解 AI 推理过程。

### 5.4 权威的医学知识图谱

权威的医学知识对于保障医疗 AI 的严谨性和可靠性至关重要。在构建医学大模型时,结合权威的医学知识图谱是确保模型能够提供科学、准确和可靠结论的基础<sup>[5]</sup>。医学知识图谱不仅包括疾病知识、药物作用、治疗指南等信息,还涵盖了临床实践中积累的宝贵经验和知识。

在可信 DeepSeek 医学应用的构建过程中,医学

知识图谱将发挥重要作用。通过与权威的医学知识库对接, DeepSeek 能够结合最新的治疗方案和药物应用指导, 优化诊疗决策。在肿瘤诊断和治疗的过程中, DeepSeek 能够通过医学知识图谱对患者病情进行精准比对, 依据科学证据推荐最优治疗方案。医学知识图谱的应用增强了 DeepSeek 在真实临床场景中的可用性和可信度, 确保了 AI 决策的医学严谨性。

### 5.5 全栈人工智能技术的融合

随着人工智能技术的不断发展, 语音交互、数字人、智能体等技术逐渐成为 AI 系统的重要组成部分。全栈人工智能技术的融合, 不仅能提升 AI 的交互性和智能化水平, 还能进一步拓宽其应用场景, 增强医疗 AI 系统的实用性<sup>[6]</sup>。DeepSeek 与全栈人工智能技术的结合, 使 AI 不仅仅作为一个静态工具存在, 而是成为一个能够与医生和患者进行高效互动的智能体。在医生诊疗过程中, DeepSeek 可以通过语音识别技术与医生进行“对话”, 解释其诊断和治疗建议, 并根据医生反馈不断优化模型。在患者治疗过程中, 数字人技术可以提供更加人性化的服务, 帮助患者更好地理解治疗过程, 解答患者疑虑, 提升患者就医体验。此外, 智能体技术能够将 DeepSeek 的强大推理能力与医院其他智能系统深度对接, 实现智能协作和资源优化。通过智能体的协调, 医生可以在更短时间内获得全面的病历分析和治疗方案推荐, 提高决策效率。

## 6 DeepSeek 对数智医院与医疗服务模式的重构

随着医疗 AI 技术不断发展, 未来医疗服务模式将发生深刻变革。DeepSeek 等大模型不仅能提高医生诊疗效率和准确性, 还能重构数智医院基座, 进一步优化医疗资源配置和提升患者就医体验。

### 6.1 数智医院建设

数智医院是基于大数据、人工智能等技术的现代化医院模式, 其核心特点是“智能化”和“数据

驱动”。在数智医院中, DeepSeek 等大模型将成为核心技术之一。通过全面整合医院临床数据、影像数据、病理数据等, 数智医院能够实现精准的诊疗决策、个性化的治疗方案以及高效的资源调配。大模型通过多模态数据融合和持续学习机制, 将为数智医院提供实时、智能化的决策支持, 帮助医生更好地管理患者健康、监控疾病进展和评估治疗效果。随着 AI 技术的发展, 数智医院将能够实现智能化的全生命周期健康管理, 从预防、诊断、治疗到随访, 全面提升医疗服务的质量和效率。

### 6.2 医疗服务模式重构

传统医疗服务模式往往依赖于医生的经验和临床直觉, 尽管医生的专业知识是不可替代的, 但在面对复杂疾病时, 传统诊疗模式存在信息不对称和效率低下的问题。通过引入 DeepSeek 等大模型, 将重构医疗服务模式。AI 不仅能够在诊断中为医生提供决策支持, 还能够在治疗过程中进行持续监控和优化。随着患者使用 AI 工具日益频繁, 医生的专业度与 AI 的对比将变得更加复杂。传统的医生角色通常是依赖个人经验进行决策, 而随着 AI 的加入, 医生的角色将更多地转向管理复杂的 AI 决策过程, 确保决策符合患者的实际情况。医生的核心价值将从单纯的知识传授转向更高层次的综合判断和治疗决策, AI 提供的技术支持将成为医生决策的工具, 而医生则专注于把握患者个体化治疗需求。此外随着 AI 的普及, 患者与医生的关系也将发生变化。患者可能更多地依赖 AI 工具获取初步的健康信息, 医生则通过深度的医学知识和专业判断对 AI 的推荐进行调整, 从而提升整体治疗效果。AI 将不会取代医生, 而是成为医生的得力助手, 帮助医生更好地理解复杂病情并提供个性化治疗。

## 7 结语

随着 AI 技术的不断发展, 医学大模型将在未来的医疗实践中发挥越来越重要的作用。DeepSeek 作为代表性的大模型之一, 正在推动医疗 AI 从研

(下转第 13 页)

- [2024-12-25]. <https://arxiv.org/abs/2302.13971>.
- 11 BAI J, BAI S, CHU Y, et al. Qwen technical report [EB/OL]. [2024-12-25]. <https://arxiv.org/abs/2309.16609>.
  - 12 阮彤, 卞侯昂, 余广涯, 等. 医学大语言模型研究与应用综述 [J]. 中国卫生信息管理杂志, 2023, 20 (6): 853-861.
  - 13 刘泓泽, 王耀国, 唐圣晟, 等. 医学大语言模型的应用现状与发展趋势研究 [J]. 中国数字医学, 2024, 19 (8): 1-7, 13.
  - 14 ZHANG H, CHEN J, JIANG F, et al. HuatuoGPT, towards taming language model to be a doctor [C]. Singapore: Findings of the Association for Computational Linguistics; EMNLP 2023, 2023.
  - 15 CHEN Y, WANG Z, XING X, et al. BianQue: balancing the questioning and suggestion ability of health LLMs with multi-turn health conversations polished by ChatGPT [EB/OL]. [2024-12-25]. <https://arxiv.org/abs/2310.15896>.
  - 16 LIAO Y, JIANG S, WANG Y, et al. MING-MOE: enhancing medical multi-task learning in large language models with sparse mixture of low-rank adapter experts [EB/OL]. [2024-12-25]. <https://arxiv.org/abs/2404.09027>.
  - 17 宋雪君. 电子病历在医疗质量控制管理中的应用与研究 [J]. 现代计算机, 2023, 29 (6): 88-90.
  - 18 冯园园, 苏源, 来庆玲, 等. 基于 3 级质控管理的电子病历系统流程优化研究 [J]. 医学信息学杂志, 2023, 44 (4): 73-77.
  - 19 周芮, 陈彦东, 宿明. 基于大数据和 AI 的电子病历自动质控系统构建 [J]. 医学信息, 2022, 35 (4): 13-16.
  - 20 刘晓娇, 黄春芳, 朱玉婷, 等. 基于临床数据中心的单病种质控管理平台的实践 [J]. 中国医疗设备, 2023, 38 (2): 90-94, 130.
  - 21 蔡子杰, 方荟, 刘建华, 等. 基于大型语言模型指令微调的心理健康领域联合信息抽取 [J]. 中文信息学报, 2024, 38 (8): 112-127.
  - 22 DING N, QIN Y, YANG G, et al. Parameter-efficient fine-tuning of large-scale pre-trained language models [J]. Nature machine intelligence, 2023, 5 (3): 220-235.
  - 23 LONGPRE S, HOU L, VU T, et al. The flan collection: designing data and methods for effective instruction tuning [C]. Honolulu: The 40th International Conference on Machine Learning (ICML23), 2023.
  - 24 WEI J, BOSMA M, ZHAO V Y, et al. Finetuned language models are zero-shot learners [C]. Virtual: The Tenth International Conference on Learning Representations, 2022.
  - 25 WIES N, LEVINE Y, SHASHUA A. The learnability of in-context learning [C]. New Orleans: Advances in Neural Information Processing Systems, 2023.

(上接第 7 页)

究阶段向临床应用迈进。通过多模态数据融合、因果推理的引入、模型可解释性的提升, 大模型不仅能提升诊断效率和准确性, 还将为数智医院的建设和医疗服务模式的重构提供强有力的技术支持。

**作者贡献:** 阎海荣负责论文构思与撰写; 江瑞、张学工参与医学大模型技术创新讨论并提出建议; 王存库参与医疗服务模式重构讨论并提出建议。

**利益声明:** 所有作者均声明不存在利益冲突。

## 参考文献

- 1 刘哲, 石钰, 林延带, 等. 智能医学的现状与未来 [J]. 科学通报, 2023, 68 (10): 1165-1181.
- 2 DeepSeek-V3 technical report [EB/OL]. [2025-01-10]. [https://github.com/deepseek-ai/DeepSeek-V3/blob/main/DeepSeek\\_V3.pdf](https://github.com/deepseek-ai/DeepSeek-V3/blob/main/DeepSeek_V3.pdf).
- 3 DeepSeek-R1: incentivizing reasoning capability in LLMs via reinforcement learning [EB/OL]. [2025-01-10]. [https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek\\_R1.pdf](https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf).
- 4 郭华源, 刘盼, 卢若谷, 等. 人工智能大模型医学应用研究 [J]. 中国科学: 生命科学, 2024, 54 (3): 482-506.
- 5 肖仰华, 徐一丹. 大规模生成式语言模型在医疗领域的应用: 机遇与挑战 [J]. 医学信息学杂志, 2023, 44 (9): 1-11.
- 6 相东升, 赖宇, 陈浩, 等. 智能时代下的空间数据科学: 基础模型研究与行业应用 [J]. 无线电工程, 2025, 55 (1): 184-195.