

· 综 述 ·

DeepSeek 在眼科的研究进展和应用前景

牛世阳 综述 曲利军 审校

哈尔滨医科大学附属第二医院眼科, 哈尔滨 150001

通信作者: 曲利军, Email: qulijun@hrbmu.edu.cn

【摘要】 眼科学正经历着以深度学习技术为代表的人工智能 (AI) 驱动的深刻变革。大型语言模型 (LLM) 作为 AI 发展的前沿, 在处理海量非结构化医疗文本、辅助复杂临床推理方面展现出巨大潜力。DeepSeek 系列模型, 作为先进 LLM 的代表, 凭借其卓越的推理能力、显著的成本效益和开源特性, 为眼科的临床、科研与管理提供了新的机遇。鉴于直接针对 DeepSeek 在眼科应用的研究尚处萌芽阶段, 本综述通过严格界定其作为 LLM 的能力边界, 并以同类先进模型已验证的应用作为基准参照, 系统阐述其在辅助诊断、自然语言处理及作为多模态系统语言中枢等方面的应用潜力, 并深入探讨其在临床实践中面临的模型“幻觉”、数据治理、监管审批等关键挑战, 以期在这个新兴的交叉领域提供早期、严谨的学术指引, 以推动 DeepSeek 技术在眼科的健康、规范与创新发展。

【关键词】 DeepSeek; 人工智能; 眼科学; 深度学习; 大语言模型

基金项目: 中华医学会医学教育分会、中国高等教育学会医学教育专业委员会医学教育研究课题 (2012-RC-27); 北京医学奖励基金会研究项目 (YXJL-2021-0815-0547); 黑龙江省教育科学“十二五”规划重点课题 (GBB1211035)

DOI: 10.3760/cma.j.cn115989-20250607-00184

Research progress and application prospects of DeepSeek in ophthalmology

Niu Shiyang, Qu Lijun

Department of Ophthalmology, the Second Affiliated Hospital of Harbin Medical University, Harbin 150001, China

Corresponding author: Qu Lijun, Email: qulijun@hrbmu.edu.cn

【Abstract】 Ophthalmology is undergoing a profound transformation driven by artificial intelligence (AI), represented by deep learning technologies. As the frontier of AI advancement, large language models (LLMs) demonstrate substantial potential in processing massive unstructured medical text and assisting with complex clinical reasoning. The DeepSeek series models, as exemplary advanced LLMs, provide novel opportunities for ophthalmic clinical practice, scientific research, and administrative management through their exceptional reasoning capabilities, remarkable cost-effectiveness, and open-source nature. Given research specifically targeting DeepSeek's ophthalmological applications remains nascent, this review systematically delineates its potential in diagnostic assistance, natural language processing, and functioning as the linguistic core for multimodal systems, which is achieved by precisely defining its LLM capabilities and benchmarking against validated applications of comparable advanced models. This article thoroughly discusses critical challenges including model "hallucinations", data governance, and regulatory approvals in clinical practice, aiming to provide early, rigorous academic guidance for this emerging interdisciplinary field and promote the healthy, standardized, and innovative advancement of DeepSeek technology in ophthalmology.

【Key words】 DeepSeek; Artificial intelligence; Ophthalmology; Deep learning; Large language models

Fund program: Medical Education Branch of Chinese Medical Association, Medical Education Professional Committee of China Association of Higher Education Research Topic of Medical Education (2012-RC-27); Beijing Medical Award Foundation (YXJL-2021-0815-0547); Key Project of Heilongjiang Province Education Science "Twelfth Five-Year" Plan (GBB1211035)

DOI: 10.3760/cma.j.cn115989-20250607-00184



现代眼科学面临疾病负担加重、精准诊疗需求提升及多源、海量数据处理困难等多重挑战。人工智能 (artificial intelligence, AI), 特别是深度学习技术, 已在推动眼科临床和科研革新中展现出巨大发展潜力^[1-2]。以 DeepSeek 系列为代表的大型语言模型 (large language models, LLM) 的兴起, 正为医学 AI 领域开辟新范式。与专注于图像识别的计算机视觉 (computer vision, CV) 模型不同, LLM 的核心优势在于对自然语言的深度理解、逻辑推理与内容生成, 其在处理眼科复杂的文本数据 (如电子病历、科研文献、临床试验报告) 和辅助诊断推理方面展现出独特潜力^[3-4]。

在众多 LLM 中, DeepSeek 模型的出现具有特殊意义。其技术先进性源于混合专家、链式思维推理与强化学习的有机结合, 为其卓越的推理能力奠定了基础。更重要的是, DeepSeek 采用了宽松的 MIT 许可证, 使其成为一个透明、可定制且低成本的开源模型。这一特性与主流的闭源商业模型形成鲜明对比, 极大地降低了医疗机构采用先进 AI 技术的门槛, 有助于在资源有限的环境中推动医学 AI 的创新与普惠发展^[5]。然而, DeepSeek 的开源特性在带来透明度和可及性的同时, 也增加了模型被滥用、恶意修改以及用于传播错误信息的风险, 对现有的监管框架和伦理规范提出了全新挑战^[6]。

目前关于 DeepSeek 在眼科领域应用的研究较为有限, 该领域尚处于探索阶段。在一个新兴技术与传统医学交叉的初期, 及时、清晰地界定其能力边界、明确其潜在的应用路径并预见挑战, 对于指导后续研究至关重要。同时需指出, OpenAI 系列模型是最早在眼科领域被系统评估与验证的 LLM 之一, 相关研究已在知识问答、病例推理、患者教育及多语言交互等场景构建了较为稳固的参考基线, 亦为本文以同类模型为基准对 DeepSeek 进行参照提供了重要依据^[7-10]。因此, 本综述立足于 DeepSeek 自身的技术特性, 通过对同类先进模型进行基准参照, 对 DeepSeek 在眼科的应用潜力进行系统性的分析与展望。

1 DeepSeek 技术概要

DeepSeek 是由深度求索公司开发的一系列 LLM, 旨在实现卓越性能与高成本效益的统一^[3]。本节将从核心技术架构、开源特性与成本、以及能力边界与多模态路径 3 个方面对其进行概述。其机构成立及系列模型发展历程见图 1。

DeepSeek-R1 (2025.01.20) 代表了一项关键的技术突破, 其核心在于展示了 LLM 的推理能力可以通过纯粹的强化学习进行激励, 而无需依赖大规模的人类标记推理轨迹。该模型建立在 DeepSeek-V3 Base 之上, 创新性地采用了群组相对策略优化作为强化学习框架。其奖励信号仅基于最终预测答案相对

真实答案的正确性, 而不对推理过程本身施加任何约束。这种方法的重大突破在于成功促进了高级推理模式的涌现, 例如自我反思、验证和动态策略调整。训练结果显示, 模型在未经明确教导如何推理的情况下, 自发演化出了更复杂的解决问题策略, 例如, 该模型在编码竞赛以及研究生水平的生物学、物理学和化学问题上展现了卓越的性能, 在可验证任务上超越了依赖传统人类示范进行监督学习的同类模型^[11]。

1.1 技术架构与核心能力

DeepSeek 的技术先进性源于其高效的架构设计。首先, 它采用混合专家架构, 通过稀疏激活机制在保证模型性能的同时显著降低了计算量, 提升了训练和推理效率, 使其在有限的硬件资源下具备高可行性^[12-14]。其次, 引入多头潜在注意力 (multi-head latent attention, MLA) 机制, 优化注意力计算, 通过压缩键和值矩阵, 有效降低了推理过程中的内存占用^[15]。在核心能力层面, DeepSeek 具备强大的长上下文处理能力, 支持高达 128K 令牌牌的上下文窗口, 能够完整分析复杂的临床文档; 同时, DeepSeek-R1 模型以推理为核心进行设计, 通过强化学习等方法着重提升了其在复杂逻辑任务上的表现, 这为其在眼科辅助诊断中的应用奠定了技术基础。

1.2 开源特性与成本

DeepSeek 最引人注目的特点是其开源特性及显著的成本优势。其训练计算预算远低于同类顶尖模型 (见表 1), 结合宽松的 MIT 许可证, 为医疗 AI 领域带来了潜在的范式转变, 打破了先进技术由少数大型机构垄断的局面。其应用层面成本优势也得到具体研究的量化支持。Mikhail 等^[3]对使用 DeepSeek-R1 和 OpenAI o1 处理 300 例眼科临床病例所需的 API 成本进行了透明化测算。该计算基于 2 个模型处理相同输入提示 (约 9.7 万令牌牌) 和各自生成输出 (DeepSeek-R1 约 21.4 万 tokens, OpenAI o1 约 10.6 万 tokens) 所需的总令牌量, 结合各自官方发布的 API 计费标准发现, 完成同样的工作量, OpenAI o1 的预估成本约为 7.78 美元, 而 DeepSeek-R1 仅为 0.52 美元, 成本降低约 15 倍^[3]。这种经济上的可行性与技术开放策略, 极大地降低了医疗机构采纳和部署先进 AI 的门槛, 并允许在本地服务器运行模型以保障数据隐私, 为该技术在眼科的广泛应用奠定了坚实基础^[16]。

1.3 能力边界与多模态路径

在评估 DeepSeek 的应用潜力时, 必须严格界定其作为 LLM 的能力边界。其核心为语言模型, 专长于处理文本信息, 不具备原生像素级图像分析能力^[17]。因此, 其在眼科影像领域的应用, 最现实且有效的路径并非直接解读图像, 而是作为多模态系统中的“语言中枢”, 与专用的 CV 模型协同工作^[18]。

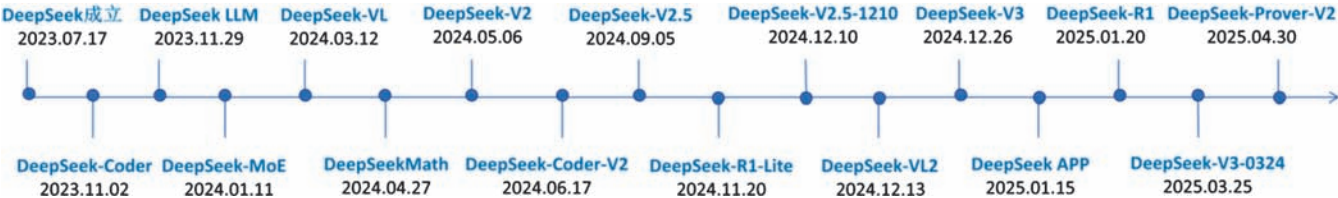


图 1 深度求索公司成立及 DeepSeek 系列模型发展时间线

表 1 DeepSeek-R1 与其他 LLM 的成本比较
(部分数据源自预印本研究^[3])

模型	指标	估算值
DeepSeek-R1	模型训练估算成本	29.4 万美元 ^[11]
OpenAI GPT-4	模型训练估算成本	7 800 万美元
Meta Llama 3.1 405B	模型训练估算成本	6 000 万美元
Google Gemini Ultra	模型训练估算成本	1.91 亿美元
DeepSeek-R1	300 例眼科病例 API 成本	0.52 美元
OpenAI o1	300 例眼科病例 API 成本	7.78 美元

2 DeepSeek 在眼科的应用现状与潜力

2.1 辅助诊断与临床推理

DeepSeek 强大的推理能力使其在辅助诊断方面具有巨大潜力^[15]。多项研究从知识掌握、独立诊断性能、人机协作价值等多个维度对其进行了初步评估。

在高水平专业知识掌握方面,一项包含近 6 000 道眼科专业题目的大规模基准测试显示,DeepSeek-R1 的准确率高达 88.8%,与表现最佳的 OpenAI o1 共同处于第一性能梯队,证明了其强大的知识储备^[7]。

在多个独立诊断任务中,DeepSeek-R1 展现了稳定的性能,但其与人类专家尚有差距。例如在处理 125 例病例时,其诊断表现与 GPT-4o 相当^[8];在另一项覆盖眼科多个亚专科的研究中,其总体诊断准确率为 79.2%,但在青光眼诊断上准确率较低(54.5%)^[19];在角膜疾病诊断任务中,DeepSeek-R1 准确率为 65.0%,但显著低于人类专家的 92.5%^[20];而在泪道疾病问答领域,其事实准确性(60%)已与最新版 ChatGPT(56%)相当,并展现出自主修正错误的能力^[9](表 2)。这些研究共同表明,DeepSeek 已处于 LLM 的第一性能梯队,但其作为独立诊断工具的可靠性仍需进一步验证。

更重要的是,DeepSeek 在人机协作方面的赋能作用也更加突出。一项针对危重症医学领域复杂危重病例诊断的前瞻性研究显示,当获得 DeepSeek-R1 辅助后,住院医师诊断准确率从 27% 升至 58%,同时诊断时间也缩短一半以上。这项研究证明,DeepSeek 是一个高效赋能工具,能够显著提升年轻医师处理复杂病例的能力和效率,揭示了其在临床决策支持和医学教育中的应用价值^[21]。

此外,DeepSeek 在跨语言推理中也展现出独特优势。一项研究指出,其在中文中的表现优于英文:在中文环境中,它显著优于 Gemini 2.0 Pro、OpenAI o1 和 OpenAI o3-mini,在英文环境中,它显著优于 OpenAI o1 和 OpenAI o3-mini,并与 Gemini 2.0 Pro 的表现相当^[10](表 2)。

综合这些初步性能数据,可以预见在经过专门眼科数据微调 and 严格临床验证后,未来 DeepSeek 有望在辅助诊断任务中发挥更重要的作用。

2.2 自然语言处理与知识管理

作为 LLM,DeepSeek 在处理和理解眼科自然语言方面展现出强大的能力,其应用主要体现在面向医疗专业人员的知识管理和面向患者的沟通教育 2 个层面。

面向医疗专业人员的应用方面,DeepSeek 能够赋能眼科临床工作的多个环节。在基础数据处理上,它有望高效地从电子健康记录(electronic health record, EHR)中提取关键信息并辅助生成临床文书^[22];在知识应用上,它能辅助医师分析海量文献与数据库,深化对眼部疾病的认知;在决策支持上,一项研究表明,经过微调的 DeepSeek-R1 能够从 EHR 中准确预测处方药物,展现了其在辅助临床决策中的潜力^[23]。

表 2 DeepSeek-R1 在眼科相关任务中的准确率

任务	题目/病例数量	准确率
眼科题目问答	5 888 道题目	88.8% ^[7]
泪道疾病问答	25 道题目	60.0% ^[9]
总体诊断	53 例病例	79.2% ^[19]
角膜疾病诊断	20 例病例	90.0% ^[19] 、65.0% ^[20]
青光眼诊断	11 例病例	54.5% ^[19]
神经眼科疾病诊断	22 例病例	81.8% ^[19]
中文眼科推理	130 道题目	86.2% ^[10]
英文眼科推理	130 道题目	80.8% ^[10]

面向患者的沟通与教育方面,研究证实 LLM 能够将专业的眼科资料转化为通俗易懂的简明摘要,且保持较高的语义准确性^[24]。DeepSeek 亦有望应用于生成清晰、准确的患者教育材料,以改善医患沟通^[25-26]。多项研究围绕“如何在保证信息准确性的前提下实现有效沟通”这一核心问题,对 DeepSeek 的性能进行了评估。在信息质量维度,DeepSeek 表现出强大的实力。在处理常见眼科患者咨询时,其在内容完整性、准确性等方面的综合表现优于 ChatGPT-4o^[27]。在依据《美国眼科学会指南》生成糖尿病视网膜病变患者教育指南时,DeepSeek 在关键信息的完整性和眼科检查建议的分类上表现最佳。然而,在沟通效果维度,结果则更为复杂。上述糖尿病视网膜病变指南生成研究指出,DeepSeek 的回答中存在使用“光学相干断层扫描”等专业术语以及未能准确传达妊娠期糖尿病筛查的特定细节的问题,可能影响患者理解^[28]。一项针对 LLM 解释复杂医学知识的研究发现,DeepSeek-R1 生成的文本可读性表现突出,同时也指出其信息质量因引用文献有限而有待进一步提升^[29]。另一项研究揭示了不同 LLM 在提供诊断和治疗建议、以及作为患者沟通潜在来源时回答质量上的互补优势:DeepSeek-R1 在信息的准确性、清晰度和可信度上评分更高,而 GPT o1-preview 可能在情感交互和共情表达上更具优势^[30]。

综上所述,DeepSeek 在赋能专业人员和辅助患者沟通方面均展现出巨大潜力。未来的研究方向不仅在于提升其面向患者沟通时在细节准确性、语言通俗化上的表现,还在于探索如何整合其专业的信息处理能力与富有人文关怀的交互能力,以开发出真正高效且人性化的医疗 AI 工具。

2.3 作为“语言中枢”的多模态系统集成框架

尽管 DeepSeek 无法直接分析图像,但其作为“语言中枢”,在多模态系统中扮演着不可或缺的角色。以 DeepSeek 为核心的应用框架设想为一个三层工作流:(1)图像分析层:由 CV 模

型,如 RETFound^[31],负责解读眼底照相、光学相干断层扫描等眼科影像,并输出结构化的文本描述或量化指标;(2)整合与推理层:以 DeepSeek 为核心,接收来自图像层的文本信息,并结合 EHR 中的患者病史、临床检查结果等数据进行综合分析和逻辑推理;(3)报告生成层:由 DeepSeek 执行,最终生成一份全面的诊断报告、鉴别诊断列表和个性化治疗建议。

最近,一项名为 OphthUS-GPT 的研究为上述框架设想提供了强有力的实践证据。该研究为眼科 B 型超声开发了一个创新的两阶段 AI 系统:第一阶段,由一个基于 BLIP 的 CV 模型分析 B 型超声图像并自动生成诊断报告;第二阶段,基于 DeepSeek 的 LLM 对这份报告进行智能问答,为医生和患者提供决策支持。在该系统中,DeepSeek 扮演了关键的“语言中枢”和“决策支持”角色。性能评估结果显示,在智能问答环节,DeepSeek 在准确性、完整性和用户满意度等多个维度上均表现优异,性能与 GPT-4o 和 OpenAI o1 相当。OphthUS-GPT 的成功,不仅为眼科 B 型超声的自动化诊断提供了创新解决方案,更为构建一个从图像分析到智能决策的全自动化多模态 AI 系统,并清晰地界定其中 LLM(如 DeepSeek)的角色,提供了极具价值的范例^[32]。

以 DeepSeek 为中心的框架下,现有关于 CV 模型在白内障^[33-36]、青光眼^[37-40]、糖尿病视网膜病变^[41,42]、年龄相关性黄斑变性^[43]等疾病中的研究成果,以及用于视网膜血管分割、视盘分割、病灶分割等任务的 AI 模型^[44-50],均可作为与 DeepSeek 集成的“前端工具”,为其提供高质量的输入信息。这种集成模式的本质是专业分工与优势互补:让 CV 模型专注于视觉模式识别,而让 DeepSeek 等 LLM 专注于复杂逻辑推理与语言生成。因此,未来推动多模态眼科 AI 系统发展的关键,在于提升前端 CV 模型将眼科影像准确、完整地“翻译”为结构化文本的能力,确保为后端的“语言中枢”提供最高质量的“养料”。

2.4 前沿应用探索

2.4.1 赋能药物研发 DeepSeek 具有强大的信息处理能力,在赋能眼科药物研发方面展现出巨大潜力。其应用贯穿了从基础研究到临床转化的多个环节。在信息层面,它能快速整合并分析海量科研文献和多组学数据(基因组学、蛋白质组学等),以辅助识别新的药物靶点^[4,51-54]。在分子层面,研究表明 DeepSeek-R1 在蛋白-配体相互作用分析、蛋白片段预测等任务中表现突出,能够生成具有生物学意义的结构化结果,为优化候选药物提供了可能^[55]。在临床转化层面,依托其数据建模能力,DeepSeek 亦有望用于高风险人群的早期识别和个性化治疗方案的制定,从而加速眼科精准医学的发展。

2.4.2 功能性生物信号分析 LLM 的应用范式正从解读静态的临床数据(如病历和影像报告),拓展至分析动态的功能性生物信号,这为评估与视觉相关的神经认知功能开辟了新路径。在一项探索利用眼动追踪数据检测阅读障碍的研究中,DeepSeek-V3 作为被评估的主流 LLM 之一,为该前沿方向提供了性能基准。尽管其在该项特定任务中的性能并非最优,但其验证了 LLM 处理此类动态生物信号、辅助神经眼科及相关功

能障碍(如阅读障碍)诊断的巨大潜力;该研究也揭示了不同 LLM 对提示工程策略的反应存在差异,表明未来需要对模型进行更具针对性的优化,以适应不同的应用场景^[56]。

3 应用前景与未来方向

DeepSeek 自 2025 年初在中国部分三级甲等医院试点推广以来,其应用逐步渗透到临床决策支持、EHR 管理、患者随访、行政管理、科研协作及医院运营等多个环节,初步展现了其在构建医院智能化生态系统的潜力^[57-58]。其强大的推理能力有望辅助临床医师进行更精准的鉴别诊断,并制定更具个体化的治疗方案。在科研层面,该模型亦能加速文献回顾、数据分析与新知识发现的进程。未来,DeepSeek 在眼科领域的应用有望在以下关键方向上取得更深入的发展与突破:

(1)眼科专用 LLM 的构建 通过高质量的眼科专业数据集对 DeepSeek 等通用 LLM 进行微调,构建更精准、更可靠的领域专用模型,减少“幻觉”发生率并提升在特定任务上的性能^[33]。(2)融合可穿戴监测 DeepSeek 或与眼部健康监测可穿戴设备集成,实现持续实时监测、早期预警与干预。(3)基于纵向文本数据的疾病进展预测 利用 LLM 分析患者多年的 EHR、随访记录等长周期文本数据,挖掘与疾病(如青光眼、年龄相关性黄斑变性)进展相关的文本模式,构建更精准的预后预测模型^[59-60]。(4)临床工作流程的 AI 助手 将 LLM 无缝嵌入 EHR 系统,作为医生的“智能助手”,实时提供信息查询、文献支持、鉴别诊断提示和临床路径建议,实现高效的人机协同。(5)高级多模态信息融合与自动化报告生成 深化 LLM 与各类 CV、数据分析模型的集成,实现从多源数据(影像、基因等)到高度整合、结构化、富有洞见的临床报告的全流程自动化^[61],生成高质量影像报告、病程记录等,减轻医师文书负担,提升文书规范性与完整性。(6)AI 赋能的手术规划与培训 在手术规划阶段,LLM 可分析海量既往病例数据和临床指南,为手术方案制定提供文本性的风险评估和策略建议;在培训方面,LLM 可用于创建交互式虚拟病例,模拟临床决策过程^[62-65]。(7)加速眼科新药的研发进程 应用 LLM 分析复杂的生物医学数据和临床试验文档,有望加速新药靶点发现、药物筛选优化及疗效与不良反应预测,缩短研发周期。

4 挑战与考量

尽管 DeepSeek 在眼科应用前景广阔,但其在临床实践中的广泛部署与深度融合仍面临从技术到伦理的多维度挑战:(1)数据治理与隐私安全 训练和优化如 DeepSeek 这类大规模 AI 模型,依赖于海量高质量、多样化且标注精准的眼科数据集;数据的获取、整合、标准化以及在整个生命周期中的隐私保护与信息安全保障是其成功应用的首要前提与核心挑战^[66-67]。(2)模型的可靠性与可解释性 LLM 固有的“幻觉”问题,即生成不实信息的现象,在医疗高风险决策中是不可接受的;同时,其“黑箱”特性也阻碍了临床信任的建立和责任的界定^[68-69];然而,部分 LLM 模型的推理风格为缓解这一困境提供了可能,如 DeepSeek-R1 倾向于输出详细且全面的中间推理

过程,这种“展示思考过程”的特性虽不能完全消除“黑箱”,但有助于临床医师审查其决策逻辑,从而在一定程度上提升其临床可信度^[7]。(3)泛化能力与鲁棒性验证 模型需在不同患者群体、疾病亚型、机构及设备条件下稳定可靠,并能有效应对真实世界数据质量不一、噪声干扰等问题^[70];(4)监管审批与标准制定 将 LLM 作为医疗器械或眼科辅助诊断工具应用于临床,须经严格监管审批;亟需完善适应技术发展的审评标准、验证路径及监管框架。需注意,闭源商业模型(如 OpenAI)与开源可本地化模型(如 DeepSeek)在数据流转路径、模型可审计性与备案资料可获取性方面存在差异,因此合规策略、验证材料准备与网络安全评估流程并不相同,应在产品化早期即与监管部门及信息安全团队协同规划。(5)临床工作流程整合与人机协同 如何将基于 DeepSeek 的智能工具无缝嵌入现有临床流程并与 EHR 等高效协同,探索有效人机协作模式,明确分工,避免增加额外负担^[71]。此外,模型的推理效率也是一个关键因素。尽管 DeepSeek-R1 在推理深度上表现优异,但其较慢的推理速度可能会影响在需要快速决策的临床场景(如门诊、急诊)中的医患体验和采纳意愿,这是未来优化和部署时必须考虑的现实问题^[7]。(6)伦理规范与社会影响 AI 应用引发伦理问题,包括训练数据偏见导致的决策不公、医疗差错责任界定、患者信任度下降,以及对医疗成本、资源分配和可及性的影响^[72-73]。(7)大规模前瞻性临床验证 在 LLM 工具被广泛推广应用于临床常规实践之前,必须通过设计严谨、样本量充足的大规模前瞻性临床试验,全面评估其在真实医疗环境下的安全性、有效性、临床效用及成本效益^[74]。(8)学术研究与内容生成的规范应用 AI 生成内容(论文、报告)的原创性、逻辑深度及严谨性或存在不足,应用时须恪守科研诚信,审慎质控与人工核验^[75]。

5 小结与展望

本文系统性探讨了以 DeepSeek 为代表的 LLM 在眼科领域的研究现状、应用潜力与未来路径,强调其核心价值在于卓越的自然语言处理、逻辑推理与知识整合能力。在研究有限的现状下,本文通过基准参照的方法,对 DeepSeek 在辅助诊断、自然语言处理、作为多模态系统的语言中枢以及在前沿探索中的应用框架进行了科学推演。然而,将这一巨大潜力转化为可靠的临床实践,仍需攻克数据治理、模型可靠性、监管审批和伦理规范等多维度挑战。未来研究应聚焦于构建眼科领域专用的 LLM、发展可信 AI 技术、探索高效且人性化的人机协同工作流程,通过大规模前瞻性临床试验验证其安全性和有效性。通过产学研医各方的紧密协作,推进技术研发、临床验证与规范化部署,以 DeepSeek 为代表的先进 AI 技术,有望为提升全球眼科医疗服务的质量与效率、加速眼科学的创新突破,最终改善患者的视觉健康与生活质量,发挥真正的、严谨的、革命性的作用。

利益冲突 所有作者均声明不存在任何利益冲突

参考文献

[1] Wang S, He X, Jian Z, et al. Advances and prospects of multi-modal

- ophthalmic artificial intelligence based on deep learning: a review [J]. *Eye Vis (Lond)*, 2024, 11(1): 38. DOI: 10. 1186/s40662-024-00405-1.
- [2] 陈有信, 杨景元. 人工智能大模型在眼科临床的应用 [J]. *中华眼科杂志*, 2024, 60(3): 303-306. DOI: 10. 3760/cma. j. cn112142-20231211-00285.
- Chen YX, Yang JY. Application of large artificial intelligence models in clinical ophthalmology [J]. *Chin J Ophthalmol*, 2024, 60(3): 303-306. DOI: 10. 3760/cma. j. cn112142-20231211-00285.
- [3] Mikhail D, Farah A, Milad J, et al. Performance of DeepSeek-R1 in ophthalmology: an evaluation of clinical decision-making and cost-effectiveness [J]. *Br J Ophthalmol*, 2025, 109(9): 976-981. DOI: 10. 1136/bjo-2025-327360.
- [4] Sevgi M, Ruffell E, Antaki F, et al. Foundation models in ophthalmology: opportunities and challenges [J]. *Curr Opin Ophthalmol*, 2025, 36(1): 90-98. DOI: 10. 1097/ICU. 0000000000001091.
- [5] Ye J, Bronstein S, Hai J, et al. DeepSeek in healthcare: a survey of capabilities, risks, and clinical applications of open-source large language models [J/OL]. *arXiv e-prints*, 2025 [2025-05-26]. <https://doi.org/10.48550/arXiv.2506.01257>.
- [6] Zeng D, Qin Y, Sheng B, et al. DeepSeek's "low-cost" adoption across china's hospital systems: too fast, too soon? [J]. *JAMA*, 2025, 333(21): 1866-1869. DOI: 10. 1001/jama. 2025. 6571.
- [7] Zou M, Srinivasan S, Lo TWS, et al. Benchmarking next-generation reasoning-focused large language models in ophthalmology: a head-to-head evaluation on 5,888 items [J/OL]. *arXiv e-prints*, 2025 [2025-05-26]. <https://doi.org/10.48550/arXiv.2504.11186>.
- [8] Sandmann S, Hegselmann S, Fujarski M, et al. Benchmark evaluation of DeepSeek large language models in clinical decision-making [J]. *Nat Med*, 2025, 31(8): 2546-2549. DOI: 10. 1038/s41591-025-03727-2.
- [9] Ali MJ. DeepSeekTM and lacrimal drainage disorders: hype or is it performing better than ChatGPT? [J]. *Orbit*, 2025: 1-7. DOI: 10. 1080/01676830. 2025. 2501656.
- [10] Xu P, Wu Y, Jin K, et al. DeepSeek-R1 outperforms Gemini 2.0 Pro, OpenAI o1, and o3-mini in bilingual complex ophthalmology reasoning [J]. *Adv Ophthalmol Pract Res*, 2025, 5(3): 189-195. DOI: 10. 1016/j. aopr. 2025. 05. 001.
- [11] Guo D, Yang D, Zhang H, et al. DeepSeek-R1 incentivizes reasoning in LLMs through reinforcement learning [J]. *Nature*, 2025, 645(8081): 633-638. DOI: 10. 1038/s41586-025-09422-z.
- [12] Mohana Sundaram A, Sathanantham ST, Ivanov A, et al. DeepSeek's readiness for medical research and practice: prospects, bottlenecks, and global regulatory constraints [J]. *Ann Biomed Eng*, 2025, 53(7): 1754-1756. DOI: 10. 1007/s10439-025-03738-7.
- [13] Peng Y, Malin BA, Rousseau JF, et al. From GPT to DeepSeek: significant gaps remain in realizing AI in healthcare [J]. *J Biomed Inform*, 2025, 163: 104791. DOI: 10. 1016/j. jbi. 2025. 104791.
- [14] Gao T, Jin J, Ke ZT, et al. A comparison of DeepSeek and other LLMs [J/OL]. *arXiv e-prints*, 2025 [2025-05-26]. <https://doi.org/10.48550/arXiv.2502.03688>.
- [15] Hayder WA. Highlighting DeepSeek-R1: architecture, features and future implications [J]. *Int J Comput Sci Mobile Computing*, 2025. DOI: 10. 47760/IJCSMC. 2025. V14I02. 001.
- [16] Temsah A, Alhasan K, Altamimi I, et al. DeepSeek in healthcare: revealing opportunities and steering challenges of a new open-source artificial intelligence frontier [J/OL]. *Cureus*, 2025, 17(2): e79221 [2025-05-26]. <https://pubmed.ncbi.nlm.nih.gov/39974299/>. DOI: 10. 7759/cureus. 79221.
- [17] Shi D, Zhang W, Chen X, et al. Eyefound: a multimodal generalist foundation model for ophthalmic imaging [J/OL]. *arXiv e-prints*, 2024 [2025-05-26]. <https://doi.org/10.48550/arXiv.2405.11338>.
- [18] Betzler BK, Chen H, Cheng CY, et al. Large language models and their impact in ophthalmology [J/OL]. *Lancet Digit Health*, 2023, 5(12): e917-e924 [2025-05-28]. <https://pubmed.ncbi.nlm.nih.gov/38000875/>. DOI: 10. 1016/S2589-7500(23)00201-7.



- [19] Hussain ZS, Delsoz M, Elahi M, et al. Performance of DeepSeek, Qwen 2.5 MAX, and ChatGPT assisting in diagnosis of corneal eye diseases, glaucoma, and neuro-ophthalmology diseases based on clinical case reports[J/OL]. medRxiv, 2025 [2025-05-28]. <https://pubmed.ncbi.nlm.nih.gov/40166547/>. DOI: 10. 1101/2025. 03. 14. 25323836. [Online ahead of print]
- [20] Jiao C, Rosas E, Asadigandomani H, et al. Diagnostic performance of publicly available large language models in corneal diseases: a comparison with human specialists[J/OL]. Diagnostics (Basel), 2025, 15(10): 1221 [2025-05-28]. <https://pubmed.ncbi.nlm.nih.gov/40428214/>. DOI: 10. 3390/diagnostics15101221.
- [21] Wu X, Huang Y, He Q. A large language model improves clinicians' diagnostic performance in complex critical illness cases[J]. Crit Care, 2025, 29(1): 230. DOI: 10. 1186/s13054-025-05468-7.
- [22] Gurav P. Natural language processing in electronic health records: a review[J]. Artif Intell Health, 2024, 1(1): 16-31. DOI: 10. 36922/aih. 2147.
- [23] Alghamdi H, Mostafa A. Advancing EHR analysis: predictive medication modeling using LLMs[J]. Inf Syst, 2025, 131: 102528. DOI: 10. 1016/j.is. 2025. 102528.
- [24] Tailor PD, D'Souza HS, Castillejo Becerra CM, et al. Utilizing AI-generated plain language summaries to enhance interdisciplinary understanding of ophthalmology notes: a randomized trial [J/OL]. medRxiv, 2024 [2025-05-28]. <https://doi.org/10.1101/2024.09.12.24313551>. [Online ahead of print]
- [25] Postaci SA, Dal A. The ability of large language models to generate patient information materials for retinopathy of prematurity: evaluation of readability, accuracy, and comprehensiveness[J]. Turk J Ophthalmol, 2024, 54(6): 330-336. DOI: 10. 4274/tjo.galenos. 2024. 58295.
- [26] AlSammarräie A, Househ M. The use of large language models in generating patient education materials: a scoping review [J]. Acta Inform Med, 2025, 33(1): 4-10. DOI: 10. 5455/aim. 2024. 33. 4-10.
- [27] Sallam M, Alasfoor IM, Khalid SW, et al. Chinese generative AI models (DeepSeek and Qwen) rival ChatGPT-4 in ophthalmology queries with excellent performance in Arabic and English[J/OL]. Narra J, 2025, 5(1): e2371 [2025-05-28]. <https://pubmed.ncbi.nlm.nih.gov/40352182/>. DOI: 10. 52225/narra. v5i1. 2371.
- [28] Cao F, Hou S. Comparative analysis of artificial intelligence tools in generating diabetic retinopathy guidelines[J]. Adv Ophthalmol Pract Res, 2025, 5(3): 180-181. DOI: 10. 1016/j. aopr. 2025. 04. 004.
- [29] Zhou M, Pan Y, Zhang Y, et al. Evaluating AI-generated patient education materials for spinal surgeries: comparative analysis of readability and DISCERN quality across ChatGPT and deepseek models[J]. Int J Med Inform, 2025, 198: 105871. DOI: 10. 1016/j. ijmedinf. 2025. 105871.
- [30] Zhou H, Wang Z, Wang R, et al. DeepSeek Versus GPT: evaluation of large language model Chatbots' responses on orofacial clefts[J]. J Craniofac Surg, 2025, 36(6): 2197-2201. DOI: 10. 1097/SCS. 00000000000011399.
- [31] Ayhan MS, Ong AY, Ruffell E, et al. In-context learning for data-efficient classification of diabetic retinopathy with multimodal foundation models[J/OL]. medRxiv, 2025 [2025-05-28]. <https://doi.org/10.1101/2025.03.09.25323618>. [Online ahead of print]
- [32] Gan F, Chen L, Qin W, et al. OphthUS-GPT: multimodal AI for automated reporting in ophthalmic B-scan ultrasound[J/OL]. medRxiv, 2025. <https://doi.org/10.1101/2025.03.03.25323237>. [Online ahead of print]
- [33] Zheng C, Ye H, Guo J, et al. Development and evaluation of a large language model of ophthalmology in Chinese[J]. Br J Ophthalmol, 2024, 108(10): 1390-1397. DOI: 10. 1136/bjo-2023-324526.
- [34] Panthier C, Zeboulon P, Rouger H, et al. CATALYZE: a deep learning approach for cataract assessment and grading on SS-OCT images[J]. J Cataract Refract Surg, 2025, 51(3): 222-228. DOI: 10. 1097/j. jcrs. 0000000000001598.
- [35] Son KY, Ko J, Kim E, et al. Deep learning-based cataract detection and grading from slit-lamp and retro-illumination photographs: model development and validation study[J]. Ophthalmol Sci, 2022, 2(2): 100147. DOI: 10. 1016/j. xops. 2022. 100147.
- [36] 王婷, 汪瑞昕, 林浩添. 基于人工智能的白内障新诊疗模式[J]. 中华实验眼科杂志, 2021, 39(9): 832-836. DOI: 10. 3760/cma. j. cn115989-20200608-00405.
- Wang T, Wang RX, Lin HT. New diagnosis and treatment pattern for cataract based on artificial intelligence[J]. Chin J Exp Ophthalmol, 2021, 39(9): 832-836. DOI: 10. 3760/cma. j. cn115989-20200608-00405.
- [37] Chuter B, Huynh J, Hallaj S, et al. Evaluating a foundation artificial intelligence model for glaucoma detection using color fundus photographs[J]. Ophthalmol Sci, 2025, 5(1): 100623. DOI: 10. 1016/j. xops. 2024. 100623.
- [38] Koornwinder A, Zhang Y, Ravindranath R, et al. Multimodal artificial intelligence models predicting glaucoma progression using electronic health records and retinal nerve fiber layer scans[J]. Transl Vis Sci Technol, 2025, 14(3): 27. DOI: 10. 1167/tvst. 14. 3. 27.
- [39] Hussain S, Chua J, Wong D, et al. Predicting glaucoma progression using deep learning framework guided by generative algorithm[J]. Sci Rep, 2023, 13(1): 19960. DOI: 10. 1038/s41598-023-46253-2.
- [40] 王媛媛, 曾献军. 人工智能在青光眼诊疗中的应用进展: 从传统眼科检查到 MRI 技术[J]. 磁共振成像, 2025, 16(1): 198-203. DOI: 10. 12015/issn. 1674-8034. 2025. 01. 032.
- Wang YY, Zeng XJ. Progress in the application of artificial intelligence in the diagnosis and treatment of glaucoma: from traditional eye examination to MRI technology[J]. Chin J Magn Reson Imaging, 2025, 16(1): 198-203. DOI: 10. 12015/issn. 1674-8034. 2025. 01. 032.
- [41] Dai L, Sheng B, Chen T, et al. A deep learning system for predicting time to progression of diabetic retinopathy[J]. Nat Med, 2024, 30(2): 584-594. DOI: 10. 1038/s41591-023-02702-z.
- [42] 董力, 李赫妍, 魏文斌. 人工智能在糖尿病视网膜病变中的应用进展[J]. 中华眼科医学杂志(电子版), 2024, 14(1): 57-61. DOI: 10. 3877/cma. j. issn. 2095-2007. 2024. 01. 010.
- Dong L, Li HY, Wei WB. Advances on the application progress of artificial intelligence in diabetic retinopathy[J]. Chin J Ophthalmol Med (Electronic Edition), 2024, 14(1): 57-61. DOI: 10. 3877/cma. j. issn. 2095-2007. 2024. 01. 010.
- [43] Schmidt-Erfurth U, Waldstein SM, Klmscha S, et al. Prediction of individual disease conversion in early AMD using artificial intelligence[J]. Invest Ophthalmol Vis Sci, 2018, 59(8): 3199-3208. DOI: 10. 1167/iovs. 18-24106.
- [44] Liu Z, Sunar MS, Tan TS, et al. Deep learning for retinal vessel segmentation: a systematic review of techniques and applications[J]. Med Biol Eng Comput, 2025, 63(8): 2191-2208. DOI: 10. 1007/s11517-025-03324-y.
- [45] 张金泽, 李嘉雄, 王耿媛, 等. 基于多路径网络的多病种精准视网膜血管网络分割方法的建立及应用[J]. 中华实验眼科杂志, 2024, 42(12): 1120-1126. DOI: 10. 3760/cma. j. cn115989-20240731-00215.
- Zhang JZ, Li JX, Wang GY, et al. Development and application of an accurate retinal vascular network segmentation method for multiple diseases based on a multi-path network[J]. Chin J Exp Ophthalmol, 2024, 42(12): 1120-1126. DOI: 10. 3760/cma. j. cn115989-20240731-00215.
- [46] Linde G, Rodrigues de Souza W Jr, Chalakkal R, et al. A comparative evaluation of deep learning approaches for ophthalmology[J]. Sci Rep, 2024, 14(1): 21829. DOI: 10. 1038/s41598-024-72752-x.
- [47] Ortiz-Feregrino R, Tovar-Arriaga S, Pedraza-Ortega JC, et al. Retinal lesion segmentation using transfer learning with an encoder-decoder CNN[J]. Rev Mex Ing Biomed, 2022, 43(2): 53-63. DOI: 10. 17488/rmib. 43. 2. 4.
- [48] Wang G, Huang Y, Ma K, et al. Automatic vessel crossing and bifurcation detection based on multi-attention network vessel segmentation and directed graph search[J]. Comput Biol Med, 2023,

- 155 : 106647. DOI:10.1016/j.comphomed.2023.106647.
- [49] Machado J, Marta A, Mestre P, et al. Data augmentation with generative methods for inherited retinal diseases: a systematic review [J]. Appl Sci, 2025, 15(6) : 3084. DOI:10.3390/app15063084.
- [50] Yang Z, Wang D, Zhou F, et al. Understanding natural language: potential application of large language models to ophthalmology [J]. Asia Pac J Ophthalmol (Phila), 2024, 13(4) : 100085. DOI:10.1016/j.apjo.2024.100085.
- [51] Ehlers JP, Clark J, Uchida A, et al. Longitudinal higher-order OCT assessment of quantitative fluid dynamics and the total retinal fluid index in neovascular AMD [J]. Transl Vis Sci Technol, 2021, 10(3) : 29. DOI:10.1167/tvst.10.3.29.
- [52] Ehlers JP, Zahid R, Kaiser PK, et al. Longitudinal assessment of ellipsoid zone integrity, subretinal hyperreflective material, and subretinal pigment epithelium disease in neovascular age-related macular degeneration [J]. Ophthalmol Retina, 2021, 5(12) : 1204-1213. DOI:10.1016/j.oret.2021.02.012.
- [53] M Bran A, Cox S, Schilter O, et al. Augmenting large language models with chemistry tools [J]. Nat Mach Intell, 2024, 6(5) : 525-535. DOI:10.1038/s42256-024-00832-8.
- [54] Chakraborty C, Bhattacharya M, Pal S, et al. AI-enabled language models (LMs) to large language models (LLMs) and multimodal large language models (MLLMs) in drug discovery and development [J]. J Adv Res, 2025, DOI:10.1016/j.jare.2025.02.011.
- [55] Farhat F. DeepSeek and the future of drug discovery: a correspondence on artificial intelligence integration [J/OL]. Intell Med, 2025, 5(2) : 164-165 [2025-05-28]. https://mednexus.org/doi/10.1016/j.imed.2025.03.001.
- [56] Nguyen QT, Nguyen H, Tang QH, et al. Learning disorder detection using eye tracking: Are large language models better than machine learning? [M]//Proceedings of the 2025 Symposium on Eye Tracking Research and Applications. New York: Association for Computing Machinery, 2025 : 1-10.
- [57] Chen J, Zhang Q. DeepSeek reshaping healthcare in China's tertiary hospitals [J/OL]. ArXiv e-prints, 2025 [2025-05-28]. https://doi.org/10.48550/arXiv.2502.16732.
- [58] Chen J, Miao C. DeepSeek deployed in 90 Chinese tertiary hospitals: how artificial intelligence is transforming clinical practice [J]. J Med Syst, 2025, 49(1) : 53. DOI:10.1007/s10916-025-02181-4.
- [59] Tordjman M, Liu Z, Yuce M, et al. Comparative benchmarking of the DeepSeek large language model on medical tasks and clinical reasoning [J]. Nat Med, 2025, 31(8) : 2550-2555. DOI:10.1038/s41591-025-03726-3.
- [60] 《人工智能在近视防治中的应用专家共识(2024)》专家组, 国际转化医学会眼科学专委会, 中国医药教育协会眼科影像与智能医疗分会, 等. 人工智能在近视防治中的应用专家共识(2024) [J]. 中华实验眼科杂志, 2024, 42(8) : 689-697. DOI:10.3760/cma.j.cn115989-20231207-00200.
- Expert Workgroup of Consensus on the application of artificial intelligence in myopia (2024), Ophthalmology Committee of International Association of Translational Medicine, Ophthalmic Imaging and Intelligent Medicine Branch of Chinese Medicine Education Association. Consensus on the application of artificial intelligence in myopia (2024) [J]. Chin J Exp Ophthalmol, 2024, 42(8) : 689-697. DOI:10.3760/cma.j.cn115989-20231207-00200.
- [61] Faray de Paiva L, Luijten G, Puladi B, et al. How does DeepSeek-R1 perform on USMLE? [J/OL]. medRxiv, 2025 [2025-05-28]. https://doi.org/10.1101/2025.02.06.25321749.
- [62] 王峰, 蓝诚红, 彭静, 等. 人工智能在眼科手术中的应用 [J]. 国际眼科纵览, 2024, 48(5) : 393-399. DOI:10.3760/cma.j.issn.1673-5803.2024.05.011.
- Wang F, Lan CH, Peng J, et al. The application of artificial intelligence in ophthalmic surgery [J]. Int Rev Ophthalmol, 2024, 48(5) : 393-399. DOI:10.3760/cma.j.issn.1673-5803.2024.05.011.
- [63] Morley D, Evans M. Multi-device pupil, limbus, and eyelid segmentation using deep learning [J]. Invest Ophthalmol Vis Sci, 2022, 63(7) : 2062-F0051.
- [64] Morley D, Evans M. Scheimpflug image segmentation using deep learning [J]. Invest Ophthalmol Vis Sci, 2024, 65(7) : 2390.
- [65] Zaldivar R, Zaldivar R, Adamek P, et al. Descriptive analysis of footplate position after myopic implantable collamer lens implantation using a very high-frequency ultrasound robotic scanner [J]. Clin Ophthalmol, 2022, 16 : 3993-4001. DOI:10.2147/OPTH.S393223.
- [66] Choudhury A, Shahsavari Y, Shamszade H. User intent to use DeepSeek for health care purposes and their trust in the large language model: multinational survey study [J/OL]. JMIR Hum Factors, 2025, 12 : e72867 [2025-05-29]. https://pubmed.ncbi.nlm.nih.gov/40418796/. DOI:10.2196/72867.
- [67] Zhang W, Lei X, Liu Z, et al. Safety evaluation of DeepSeek models in Chinese contexts [J/OL]. arXiv e-prints, 2025 [2025-05-29]. https://doi.org/10.48550/arXiv.2502.11137.
- [68] Oshika T. Artificial intelligence applications in ophthalmology [J]. JMA J, 2025, 8(1) : 66-75. DOI:10.31662/jmaj.2024-0139.
- [69] Chowdhury M N-U-R, Haque A, Ahmed I. DeepSeek vs. ChatGPT: a comparative analysis of performance, efficiency, and ethical AI considerations [J/OL]. TechRxiv, 2025 [2025-05-29]. https://www.techrxiv.org/users/690185/articles/1265410-DeepSeek-vs-chatgpt-a-comparative-analysis-of-performance-efficiency-and-ethical-ai-considerations. DOI:10.36227/techrxiv.173929663.35290537/v1.
- [70] Smiju IS, Adinath DR. Advancements in AI-Powered NLP models: a critical analysis of ChatGPT and DeepSeek [J/OL]. SSRN, 2025 [2025-05-29]. http://dx.doi.org/10.2139/ssrn.5125445.
- [71] De Carvalho Souza ME, Weigang L, Grok, Gemini, ChatGPT and DeepSeek: comparison and applications in conversational artificial intelligence [J]. Intellig Artif, 2025, 2(1) : 1-7. DOI:10.5281/zenodo.14885243.
- [72] Yaghy A, Yaghy M, Shields JA, et al. Large language models in ophthalmology: potential and pitfalls [J]. Semin Ophthalmol, 2024, 39(4) : 289-293. DOI:10.1080/08820538.2023.2300808.
- [73] 《眼科人工智能临床应用伦理专家共识(2023)》专家组, 中国医药教育协会数字影像与智能医疗分会, 中国医药教育协会智能医学专业委员会. 眼科人工智能临床应用伦理专家共识(2023) [J]. 中华实验眼科杂志, 2023, 41(1) : 1-7. DOI:10.3760/cma.j.cn115989-20220905-00414.
- Expert Workgroup of Expert consensus for ethics of clinical application of artificial intelligence in ophthalmology (2023), Digital Imaging and Intelligent Medicine Branch of China Medical Education Association, Intelligent Medicine Special Committee of China Medical Education Association. Expert consensus for ethics of clinical application of artificial intelligence in ophthalmology (2023) [J]. Chin J Exp Ophthalmol, 2023, 41(1) : 1-7. DOI:10.3760/cma.j.cn115989-20220905-00414.
- [74] 陈有信, 冯时, 赵清. 眼科人工智能研究的相关问题 [J]. 中华眼底病杂志, 2022, 38(2) : 89-92. DOI:10.3760/cma.j.cn11434-20220210-00077.
- Chen YX, Feng S, Zhao Q. Issues related to artificial intelligence research in ophthalmology [J]. Chin J Ocul Fundus Dis, 2022, 38(2) : 89-92. DOI:10.3760/cma.j.cn11434-20220210-00077.
- [75] Aydin O, Karaarslan E, Erenay FS, et al. Generative AI in academic writing: a comparison of DeepSeek, Qwen, ChatGPT, Gemini, Llama, Mistral, and Gemma [J/OL]. arXiv e-prints, 2025 [2025-05-29]. https://doi.org/10.48550/arXiv.2503.04765.

(收稿日期:2025-06-07 修回日期:2025-08-28)

(本文编辑:张宇 骆世平)

