科学通报 2026 年 第 71 卷 第 15 期: 3320 ~ 3327
人工智能驱动的数字医疗新实践专题 进 展
《中国科学》杂志社 SCIENCE CHINA PRESS

从临床经验到人工智能: 斜视数字化诊断与智能手术规划系统的研究进展

苗泽群[1,2,3], 王文思[4], 王乐今[1,2,3*]

1. 北京大学人民医院眼科, 北京 100044
2. 眼病与视光医学北京市重点实验室, 北京 100044
3. 北京大学医学部眼视光学院, 北京 100044
4. 北京工业大学信息科学技术学院, 北京 100124
* 联系人, E-mail: Dr.wanglejin@bjmu.edu.cn

2026-04-20 收稿, 2026-05-06 修回, 2026-05-07 接受, 2026-05-08 网络版发表
国家自然科学基金重大科研仪器研制项目(31427801)资助

摘要 斜视是常见眼科疾病, 传统诊断和手术规划高度依赖临床经验, 易受检查者经验、患者配合程度和医疗资源差异影响. 近年来, 人工智能(artificial intelligence, AI)、深度学习、眼动追踪和虚拟现实(virtual reality, VR)等数字化技术为斜视筛查、偏斜角量化、动态功能评估和手术辅助规划提供了新的工具. 本文定位为叙述性综述, 围绕近年代表性研究, 归纳斜视数字化诊断与智能手术规划的主要技术路线、应用场景和证据边界. 在数字化诊断方面, 基于静态图像/视频的深度学习方法可实现斜视筛查、分型和偏斜角估计; 眼动追踪方法有助于捕获间歇性斜视和双眼协调异常; 移动端与嵌入式系统则提高了基层和居家筛查的可及性. 在智能手术规划方面, 机器学习回归、多任务学习、角膜光反射照片分析、生物力学建模和VR模拟等方法正在探索对手术肌肉选择、手术量设计和术后效果预测的辅助价值. 现有研究提示, 部分AI模型在特定数据集和受控条件下显示出较高性能, 但多数证据仍来自单中心、回顾性或内部验证研究, 复杂斜视类型、外部验证、多模态整合、长期预后预测及临床责任边界仍不足. 未来研究应重视标准化数据采集、多中心前瞻性验证、可解释性、安全监管和临床工作流融合, 使AI更稳妥地服务于斜视精准诊疗.

关键词 斜视, 人工智能, 深度学习, 数字化诊断, 手术规划

斜视是一种以双眼视轴相对偏离为主要特征的眼部疾病, 可表现为水平、垂直和/或旋转方向的眼位异常. 亚洲学龄前儿童斜视患病率约为3.55%~5.65%[1,2]. 斜视不仅影响外观, 还可能导致复视、异常头位、立体视觉受损和心理社会功能受损[3,4]. 儿童期斜视手术时机与后续心理健康结局也存在关联, 提示及时、规范诊疗具有重要意义[5].

精确判断斜视类型和偏斜程度是治疗方案制定和手术量设计的基础. 三棱镜加交替遮盖试验(alternate prism cover test, APCT)仍是临床评估斜视度数的重要方法[6,7]. 然而, APCT结果可能受到检查者经验、观察判断、患者配合、屈光矫正状态和检查条件等因素影响, 不同检查者之间存在一定变异[6]. 因此, 建立更加客观、自动化和可重复的斜视评估体系, 是提高诊疗一致性和基层可及性的重要方向. 近年来, AI与数字健康技术快速发展, 为斜视诊疗从经验依赖向数据辅助决策转变提供了新的技术基础. 深度学习模型可从面部照片、眼部图像和遮盖试验视频中提取眼位和角膜光

引用格式: 苗泽群, 王文思, 王乐今. 从临床经验到人工智能: 斜视数字化诊断与智能手术规划系统的研究进展. 科学通报, 2026, 71: 3320–3327
Miao Z, Wang W, Wang L. From clinical experience to artificial intelligence: research progress in digital diagnosis and intelligent surgical planning systems for strabismus (in Chinese). Chin Sci Bull, 2026, 71: 3320–3327, doi: 10.1360/CSB-2026-0534

© 2026《中国科学》杂志社

反射等特征；眼动追踪可记录双眼运动协调过程；VR技术可构建标准化动态测试场景；机器学习和生物力学建模则为手术规划和疗效预测提供新的辅助路径[8~10].

本文为叙述性综述，聚焦斜视数字化诊断与智能手术规划领域的代表性研究. 文章首先归纳静态图像/视频分析、眼动追踪和移动端/嵌入式筛查等诊断技术路线，随后讨论智能手术规划、生物力学建模和VR模拟等方向，并重点分析现有证据的适用边界、技术短板和临床转化挑战.

1 斜视数字化诊断方法的研究进展

斜视数字化诊断的输入模态包括静态图像、遮盖试验视频、角膜光反射照片、眼动轨迹和移动端采集数据等. 按照数据来源和算法任务, 现有研究大致可分为基于静态图像/视频的深度学习方法、基于眼动追踪的动态评估方法，以及面向基层场景的移动端与嵌入式系统. 图1概括了斜视诊疗从传统经验到数字化采集、AI算法建模、临床功能输出和应用场景部署的整体框架.

2 基于静态图像与视频的深度学习方法

基于面部图像或眼部图像的深度学习方法是当前斜视AI诊断研究中较常见的技术路线. 卷积神经网络(convolutional neural network, CNN)及其变体可自动提取眼位、角膜光反射点、瞳孔轮廓和眼部关键点等视觉特征，从而完成斜视筛查、亚型分类或偏斜角度估计.

Rodriguez等人[11]开发了基于Raspberry Pi 4、USB摄像头和3D打印下巴托的低成本嵌入式筛查系统，并将NASNetLarge模型转换至TensorFlow Lite框架用于实时推理. 该研究提示, 轻量化深度学习模型可在低成

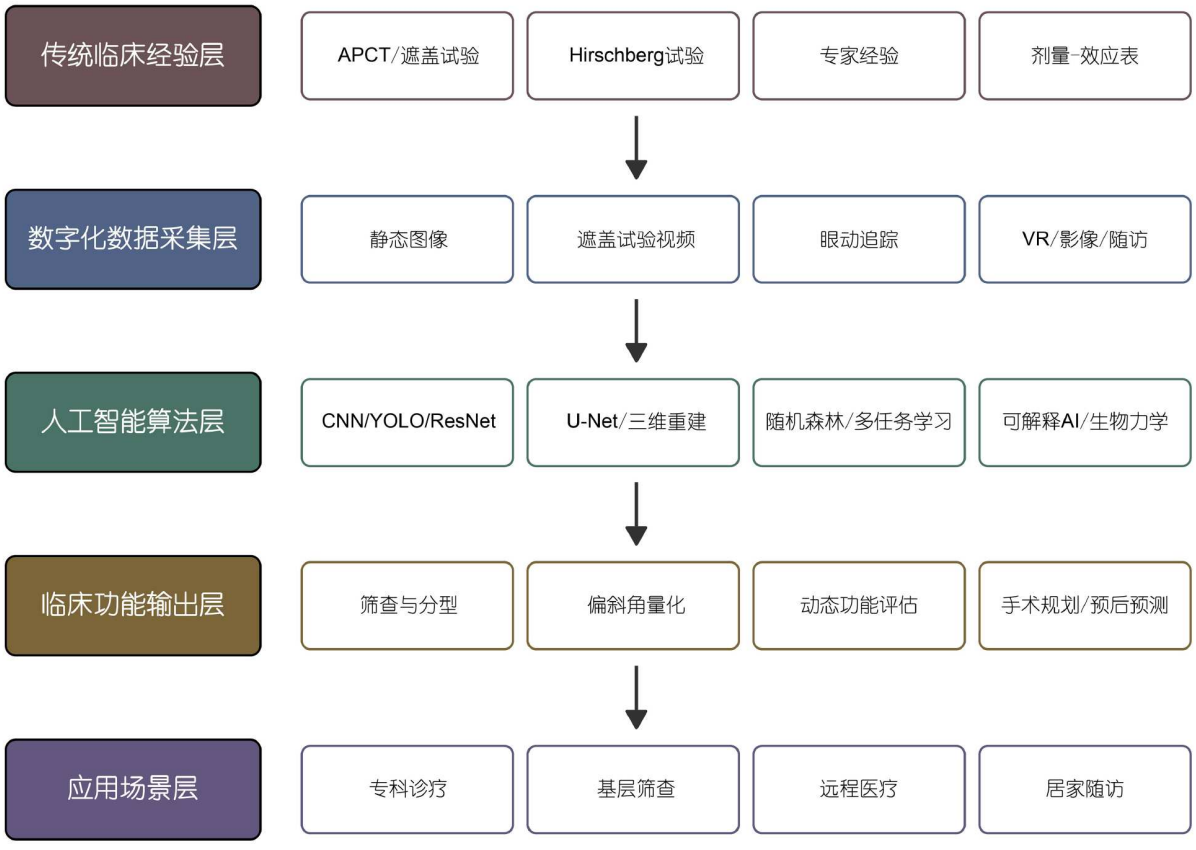


图 1 (网络版彩色)从临床经验到人工智能的斜视数字化诊疗体系框架. APCT, 三棱镜加交替遮盖试验; AI, 人工智能; CNN, 卷积神经网络; VR, 虚拟现实

Figure 1 (Color online) Framework of digital strabismus diagnosis and treatment from clinical experience to artificial intelligence. APCT, alternate prism cover test; AI, artificial intelligence; CNN, convolutional neural network; VR, virtual reality

本硬件上完成初步视觉错位筛查，但其主要任务仍偏向筛查而非复杂分型和临床诊断替代.

Garg等人[12]提出StrabNet-CQ框架, 整合YOLOv8、ResNet101和ResNet18, 用于斜视检测、亚型分类和眼部关键点量化. 该系统在600张眼部图像上实现94%的斜视检出准确率和90%的分型准确率, 量化参数与人工棱镜测量值相关系数为r=0.733. 该研究展示了“分类+量化”一体化模型的潜力, 但数据规模、外部验证和复杂斜视覆盖仍需加强.

视频模态可记录遮盖试验过程中的动态眼位变化. Wang等人[13]开发了智能手机“数字斜视尺”(digital ruler of strabismus, DRS), 利用三维重建和深度学习分析30s遮盖试验视频. 在335名参与者、459段视频中, DRS与APCT相比的平均绝对误差为4.51棱镜度(prism diopter, PD), 水平偏斜测量组内相关系数(intraclass correlation coefficient, ICC)为0.98, 并可分析间歇性斜视的恢复时间和速度. 该研究优势在于多中心前瞻性设计和动态量化, 但仍需评估不同拍摄质量、复杂眼位异常和真实世界居家场景下的稳定性.

另有研究将交替遮盖试验与Faster-RCNN、Efficient-UNet等算法结合, 通过检测面部和眼部区域、分割角膜和眶区并计算角膜位移辅助判断斜视, 显示出将传统检查流程转化为算法流程的可行性[14]. 在证据综合层面, Yang等人[15]对AI斜视筛查研究进行系统评价和荟萃分析, 提示AI模型在筛查任务中总体表现较好, 但研究异质性、外部验证不足和数据模态差异仍限制了跨研究比较. Rajavi等人[8]亦指出, AI与数字健康工具在斜视筛查和诊断中具有潜力, 但需要更多高质量临床验证.

3 基于眼动追踪的动态评估方法

静态图像方法适合标准注视位下的眼位分析, 但对间歇性斜视和双眼协调的动态异常捕获有限. 眼动追踪设备可连续记录眼球运动轨迹, 为识别短暂眼位偏移、融合恢复速度和双眼运动协调性提供信息.

Zhao等人[16]报道了结合可穿戴眼动追踪设备和AI算法的间歇性斜视筛查方法. 该系统在双眼大角度连续运动过程中采集红外眼动图像, 并基于瞳孔-眦角向量提取16项具有生理学解释的眼动特征, 经随机森林模型分类后在70名受试者中取得较高筛查准确率. 该方法的优势在于动态评估和特征可解释性, 但目前样本量有限, 且设备标准化、佩戴舒适度和多场景验证仍是推广前需要解决的问题.

Miao等人[10]进一步探索了VR环境下结合眼动追踪的眼位测量. VR可提供可控、可重复的动态视觉刺激环境, 有助于观察不同注视方向和视觉任务下的眼位变化, 但设备成本、临床流程嵌入和儿童患者配合度仍限制其常规应用.

4 移动端与嵌入式系统: 提高筛查可及性的路径

基层和资源有限地区面临专业眼科设备不足、受训人员有限和转诊链条不完善等问题. 移动端和嵌入式系统的核心价值不在于替代专科诊断, 而在于以较低成本实现风险初筛、远程管理和随访提醒.

Wu等人[17]开发了基于智能手机图像的斜视筛查和管理AI平台, 展示了移动端工具在患者管理中的潜力. Rodriguez等人[11]的嵌入式系统则体现了边缘计算和低成本硬件结合的可行性. 此类系统未来需要与标准化拍摄流程、远程质控和转诊机制结合, 才能真正提高基层筛查效率.

图2进一步概括了不同技术路线的数据输入、核心算法、临床输出和应用边界. 总体而言, 静态图像/视频方法适用于筛查、分型和角度估计; 眼动追踪更适合动态功能评估; 移动端系统更强调可及性; 智能手术规划则面向更复杂的临床决策.

5 不同诊断研究的证据特征与局限性

由于不同研究的临床任务、输入模态、参照标准和评价终点存在明显差异, 单纯将准确率、敏感性、特异性、角度误差、ICC和相关系数等指标并列比较, 容易造成统计口径不一致和临床解释偏差. 因此, 表1以代表性研究为基础, 分项梳理其研究任务、数据模态、适用范围、验证方式和主要局限.

6 斜视智能手术规划系统的研究进展

数字化诊断主要关注斜视识别、分型和偏斜角量化; 智能手术规划进一步面向手术肌选择、手术量设计和术后效果预测等临床决策问题. 传统手术规划依赖经验性剂量-效应表和术者个人经验, 存在一定术者间差异. AI方法可从既往病例中学习术前特征与手术方案之间的关联, 但其输出应被理解为辅助建议, 而不能替代专科医师对患者个体情况的综合判断.

在机器学习回归方面, Speidel等人[18]纳入767例水



图 2 (网络版彩色)斜视人工智能诊疗研究的主要技术路线及应用边界. AI, 人工智能; CNN, 卷积神经网络; MLP, 多层感知机

Figure 2 (Color online) Major technical routes and application boundaries of artificial intelligence in strabismus diagnosis and treatment. AI, artificial intelligence; CNN, convolutional neural network; MLP, multi-layer perceptron

平非麻痹性斜视手术患者, 使用多层感知机(multi-layer perceptron, MLP)预测手术量. 模型对第一和第二手术肌肉的预测均方根误差分别为0.54和0.71 mm, 且可解释AI分析提示术前偏斜角是重要特征. 该研究说明神经网络可能辅助手术量决策, 但作者也强调扩大数据基础后才可能进一步提高准确性.

在深度学习多任务学习方面, Wang等人[9]开发并内部验证了水平斜视术前规划模型, 尝试同时预测手术肌选择和手术量等关联任务. 该类模型的优势在于能够学习多个临床任务之间的共享表征, 但其结果容易受到单中心术者习惯、病例选择和训练数据分布影响, 尚需多中心外部验证和前瞻性临床评估.

Mao等人[19]开发了基于角膜光反射照片的AI平台, 由斜视筛查、偏斜角评估和手术建议三个深度学习子系统构成. 该研究在回顾性和前瞻性数据集中均报告了较高筛查性能, 并显示角膜光反射照片在水平斜视评估和外斜视手术建议中的潜力.

生物力学建模和VR模拟为智能手术规划提供机制驱动的补充路径. 斜视手术的本质是改变眼外肌力学平衡, 单纯依赖数据关联的模型在复杂斜视、再手术或异常解剖场景下可能泛化不足. 结合影像重建、有限元建模、眼动追踪和VR模拟, 可为理解个体眼外肌力学状态、模拟不同方案效果和培训年轻医生提供工具[10,20]. 然而, 这些方向仍面临建模成本高、参数标准化不足和临床验证有限等问题. 表2汇总了斜视智能手术规划及扩展技术的代表性研究.

7 创新性、应用边界与科学意义

从技术角度看, 斜视AI研究正在从单一二分类筛查向“筛查—分型—量化—规划”的复合任务演进. 静态图像模型有助于标准化眼位判断, 视频和眼动追踪拓展了动态功能评估, 多任务学习和生物力学建模则开始面向更复杂的手术决策问题.

从临床角度看, 这些工具的主要价值在于辅助医

表 1 斜视数字化诊断代表性研究的证据特征、验证方式与局限性[a)]

Table 1 Evidence characteristics, validation methods and limitations of representative studies on digital diagnosis of strabismus

代表性研究	研究任务	数据模态/设备	样本量/数据量	参照标准/验证方式	主要结果与局限性
Yang等人[15]	AI斜视筛查荟萃分析	多种图像、视频和眼动数据	24项研究, 至少8484名患者和40394个眼部测量数据	各原始研究临床标准; 系统评价与荟萃分析	汇总敏感性和特异性较高; 但原始研究异质性、外部验证不足
Garg等人[12]	斜视检出、分型与量化	眼部图像; YOLOv8、ResNet101、ResNet18	600张眼部图像	人工标注/棱镜测量; 内部验证	检出准确率94%, 分型准确率90%; 样本量和外部验证有限
Wang等人[13]	智能手机自动测量斜视	遮盖试验视频; 三维重建和深度学习	335名参与者, 459段视频	APCT; 前瞻性多中心队列	平均绝对误差4.51 PD, 水平偏斜ICC 0.98; 居家视频质量和复杂病例仍需验证
Zhao等人[16]	间歇性斜视动态筛查	可穿戴红外眼动追踪	70名受试者	临床诊断; 室内条件验证	筛查准确率较高; 设备依赖性和真实世界泛化仍需评估
Rodriguez等人[11]	低成本视觉错位筛查	Raspberry Pi、USB摄像头、TensorFlow Lite	图像数据集及小样本验证	人工/临床标注; 交叉验证或验证集	便携、低成本; 主要用于初筛, 分型和角度量化能力有限
Wu等人[17]	智能手机筛查与管理	智能手机图像; AI平台	原文数据集	临床诊断; 内部验证	适合移动端管理探索; 多中心真实世界验证仍需加强
Wang等人[14]	遮盖试验结合深度学习检测	遮盖试验图像/视频; Faster-RCNN、Efficient-UNet	会议研究数据集	临床检查; 内部验证	AUC 0.901, 敏感性96.91%, 特异性83.33%; 会议研究信息有限

a) AI, artificial intelligence, 人工智能; APCT, alternate prism cover test, 三棱镜加交替遮盖试验; AUC, area under the receiver operating characteristic curve, 受试者工作特征曲线下面积; CNN, convolutional neural network, 卷积神经网络; ICC, intraclass correlation coefficient, 组内相关系数; PD, prism diopter, 棱镜度

表 2 斜视智能手术规划及扩展技术的代表性研究

Table 2 Representative studies on intelligent surgical planning and extended technologies for strabismus

代表性研究	技术方向	输入数据	算法/模型	输出任务	主要价值与局限性
Mao等人[19]	诊断—量化—手术建议平台	角膜光反射照片	三个深度学习子系统	筛查、角度评估、外斜视手术建议	实现从图像到辅助建议的流程探索; 主要适用于特定任务和数据分布
Wang等人[9]	水平斜视术前规划	术前临床变量	多任务学习模型	手术肌选择、后徙/缩短量预测	可学习关联任务; 单中心内部验证, 需外部验证
Speidel等人[18]	手术剂量预测	术前偏斜角、屈光等变量	MLP与可解释AI	手术剂量预测	预测误差较小; 复杂病例和长期预后预测不足
Miao等人[10]	VR眼位测量与模拟	VR场景下眼动数据	VR结合眼动追踪	动态眼位测量、手术模拟探索	提供标准化动态环境; 设备和流程限制明显
Liu和Wu[20]	AI在斜视诊疗中的应用展望	多模态临床数据	大模型、多模态融合、智能诊疗平台	个性化管理与精准化辅助	强调未来方向; 仍需原始临床研究验证

生提高评估客观性、记录量化指标、降低基层筛查门槛并减少部分重复性工作. 尤其在基层和远程医疗场景中, 移动端筛查工具可帮助识别高风险患者并促进转诊. 然而, 当前证据不足以支持AI系统独立完成诊断或手术方案制定, 临床专家仍应承担最终判断和责任.

从学科交叉角度看, 该领域融合眼科学、计算机视觉、机器学习、生物力学、人机交互和医学信息学. 未来若能建立标准化多中心数据库, 并将图像、视频、眼动、影像、术中记录和术后随访等信息有机整合, 可能为斜视机制研究和精准治疗提供更丰富的数据基础.

8 挑战与展望

第一, 数据代表性不足仍是核心瓶颈. 现有模型多基于单中心、回顾性或受控条件下采集的数据, 常见水平斜视病例占比较高, 而垂直斜视、旋转斜视、麻痹性斜视、限制性斜视、既往手术后斜视和合并眼部或神经系统疾病的复杂病例纳入不足. 这将限制模型在真实世界场景中的泛化能力.

第二, 数据采集和参照标准尚不统一. 不同研究在拍摄距离、光照、注视目标、头位控制、遮盖试验流程和眼动设备参数方面存在差异; 参照标准也可能包括APCT、Hirschberg试验、专家标注或术者经验方案.

若不明确报告参照标准和验证流程，模型性能之间难以直接比较.

第三, 外部验证和前瞻性研究不足. 部分模型在内部验证中表现较好，但其性能可能反映训练机构的病例结构、拍摄流程和专家习惯. 未来需要跨地区、跨设备、跨年龄段的外部验证，以及前瞻性研究评估AI辅助是否真正改善诊断一致性、手术决策质量和患者结局.

第四, 智能手术规划的临床边界尤其需要谨慎. 手术量不仅由术前偏斜角决定，还受到眼外肌弹性、既往手术史、融合功能、屈光状态、年龄、术者技术和术后组织反应等因素影响. 当前模型更多是对既往专家方案的统计学习，尚不能充分解释复杂病例中的生物力学和愈合差异.

第五, 临床落地涉及工作流、责任和监管问题. AI系统需要嵌入门诊检查、电子病历、术前讨论和随访流程，同时明确误诊、漏诊或不恰当手术建议时的责任边界. 数据隐私、算法公平性、医疗器械审批、模型更新和持续质控也必须纳入临床转化设计.

未来研究应从“单模型性能提升”转向“临床可验证、可解释、可监管的系统建设”. 可行方向包括标准化多模态数据库、多中心前瞻性验证、面向复杂斜视的分层建模、结合生物力学机制的个体化预测、长期预后和再手术风险评估，以及以医生为中心的人机协同决策界面.

9 总结

斜视数字化诊断与智能手术规划正处于从技术探索走向临床验证的关键阶段. 基于静态图像/视频的深度学习方法、眼动追踪动态评估、移动端与嵌入式筛查系统，以及机器学习手术规划、生物力学建模和VR模拟等方向，均为提高斜视诊疗客观性、可及性和个体化水平提供了新思路. 然而, 现有证据仍受到样本代表性不足、外部验证有限、多模态整合不足和临床责任边界不清等因素制约. 未来应以多中心、前瞻性、真实世界研究为基础，推动AI从单点工具逐步发展为安全、可解释、可监管的临床辅助系统.

参考文献

1 Chen X, Fu Z, Yu J, et al. Prevalence of amblyopia and strabismus in Eastern China: results from screening of preschool children aged 36–72 months. Br J Ophthalmol, 2016, 100: 515–519

2 McKean-Cowdin R, Cotter S A, Tarczy-Hornoch K, et al. Prevalence of amblyopia or strabismus in Asian and non-hispanic white preschool children. Ophthalmology, 2013, 120: 2117–2124

3 Hofsli M, Torp-Pedersen T, Lønkvist C S, et al. Strabismus surgery alleviates functional and psychosocial problems. Ugeskr Laeger, 2023, 185: v07220432

4 Dohlman J C, Hunter D G, Heidary G. The impact of strabismus on psychosocial equity. Semin Ophthalmol, 2023, 38: 52–56

5 Hidinger I, Kong L, Ely A. Associations of strabismus surgery timing in childhood with mental health: a retrospective cohort study. J Am Assoc Pediatr Ophthalmol Strabismus, 2024, 28: 103929

6 de Jongh E, Leach C, Tjon-Fo-Sang M J, et al. Inter-examiner variability and agreement of the alternate prism cover test (APCT) measurements of strabismus performed by 4 examiners. Strabismus, 2014, 22: 158–166

7 Johns H A, Manny R E, Fern K, et al. The intraexaminer and interexaminer repeatability of the alternate cover test using different prism neutralization endpoints. Optom Vis Sci, 2004, 81: 939–946

8 Rajavi Z, Hatami F, Khosravi Mirzaei S, et al. Advancements in strabismus diagnosis: a comprehensive systematic review of artificial intelligence and digital health applications. Exp Eye Res, 2025, 258: 110501

9 Wang J, Wu X, Ou S. Development and internal validation of a prescriptive multi-task learning model for horizontal strabismus surgery planning. BMC Ophthalmol, 2026, 26: 37

10 Miao Y, Jeon J Y, Park G, et al. Virtual reality-based measurement of ocular deviation in strabismus. Comput Methods Programs Biomed, 2020, 185: 105132

11 Rodriguez D S, Rivera Gomez A E, Serrezuela R R. Development of an embedded diagnostic tool for visual misalignment screening. HardwareX, 2025, 23: e00692

12 Garg S, Sunkarapalli A, Ghosh D, et al. StrabNet-CQ: an integrated deep learning framework for automated strabismus classification and quantification using ocular landmark detection. BMC Ophthalmol, 2025, 25: 595

13 Wang R, Lyu J, Yang Y, et al. A smartphone-based digital ruler to automatically measure strabismus in ophthalmologist-level: a prospective, multicenter cohort study. NEJM AI, 2025, 2: e2500657

14 Wang T, Jin M, Xu C, et al. Mobile strabismus automatic detection method combining alternating cover test and deep learning. In: Proceedings of the 3rd International Conference on Health Big Data and Intelligent Healthcare. Zhuhai, 2024. 141–146
15 Yang Z, Wu D, Li J, et al. Systematic review and meta-analysis of artificial intelligence models for strabismus screening: methodological insights and future directions. Int J Surg, 2025, 111: 8447–8461
16 Zhao Z, Meng H, Li S, et al. High-accuracy intermittent strabismus screening via wearable eye-tracking and AI-enhanced ocular feature analysis. Biosensors, 2025, 15: 110
17 Wu D, Li Y, Zhang H, et al. An artificial intelligence platform for the screening and managing of strabismus. Eye, 2024, 38: 3101–3107
18 Speidel A J, Fetzer B, Wullbrand M, et al. Investigation of machine learning methods for predicting surgical parameters in strabismus surgery. Discov Artif Intell, 2026, 6: 54
19 Mao K, Yang Y, Guo C, et al. An artificial intelligence platform for the diagnosis and surgical planning of strabismus using corneal light-reflection photos. Ann Transl Med, 2021, 9: 374
20 Liu L Q, Wu D W. Paying attention to the application of artificial intelligence in strabismus diagnosis and treatment (in Chinese). Chin J Exp Ophthalmol, 2024, 42: 1079–1083 [刘陇黔, 吴达文. 关注人工智能在斜视诊疗中的应用. 中华实验眼科杂志, 2024, 42: 1079–1083]

Summary for “从临床经验到人工智能: 斜视数字化诊断与智能手术规划系统的研究进展”

From clinical experience to artificial intelligence: research progress in digital diagnosis and intelligent surgical planning systems for strabismus

Zequn Miao[1,2,3], Wensi Wang[4] & Lejin Wang[1,2,3*]

[1] *Department of Ophthalmology, Peking University People's Hospital, Beijing 100044, China*
[2] *Beijing Key Laboratory of Ocular Disease and Optometry Science, Beijing 100044, China*
[3] *College of Optometry, Peking University Health Science Center, Beijing 100044, China*
[4] *Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China*
* Corresponding author, E-mail: Dr.wanglejin@bjmu.edu.cn

Strabismus is a common ocular disorder characterized by horizontal, vertical, and/or torsional misalignment of the visual axes. Accurate assessment of the type and magnitude of ocular deviation is essential for treatment selection and surgical planning. However, conventional diagnosis and surgical decision-making still rely heavily on clinical experience, manual observation, patient cooperation, and local availability of trained specialists. These factors may lead to variability in measurement and limit access to timely screening and referral in primary care settings. In recent years, artificial intelligence, deep learning, eye tracking, virtual reality, and other digital health technologies have provided new opportunities for objective and quantitative assessment of strabismus. This narrative review summarizes representative advances in digital diagnosis and intelligent surgical planning for strabismus, with an emphasis on clinical tasks, data modalities, algorithmic approaches, validation status, and application boundaries. For digital diagnosis, image- and video-based deep learning models have been explored for screening, subtype classification, and estimation of ocular deviation. Wearable or infrared eye-tracking systems can capture dynamic binocular coordination and may be particularly useful for intermittent strabismus or subtle abnormalities that are difficult to identify from a single static image. Smartphone-based and embedded systems may improve accessibility in community screening and remote follow-up, although their roles are currently better defined as risk stratification tools rather than replacements for specialist diagnosis. For surgical planning, machine learning regression, multi-task learning, artificial intelligence platforms based on corneal light-reflection photographs, biomechanical modeling, and virtual reality simulation have been investigated to support muscle selection, dosage prediction, and surgical training. These studies suggest that artificial intelligence may help standardize parts of the diagnostic and planning workflow. Nevertheless, current evidence remains limited by small or single-center datasets, retrospective design, internal validation, heterogeneous reference standards, and underrepresentation of complex strabismus, reoperations, vertical or torsional deviations, and patients with combined ocular or neurological conditions. Therefore, the performance reported in controlled studies should not be directly interpreted as readiness for autonomous clinical decision-making. Future research should prioritize standardized multimodal data acquisition, multicenter external validation, prospective real-world evaluation, explainable models, integration with clinical workflows, long-term outcome prediction, and clear ethical, regulatory, and responsibility frameworks. Artificial intelligence is more likely to become a safe and useful adjunct to strabismus care when it is developed as part of a clinician-centered, interpretable, and verifiable decision-support system.

strabismus, artificial intelligence, deep learning, digital diagnosis, surgical planning

doi: 10.1360/CSB-2026-0534