

·标准与规范探讨·

中国基于眼底照相的人工智能青光眼辅助筛查系统规范化设计及应用指南(2020年)

中华医学会眼科学分会青光眼学组 中国医学装备协会眼科人工智能学组

通信作者:王宁利,首都医科大学附属北京同仁医院北京同仁眼科中心 北京市眼科研究所 北京市眼科学与视觉科学重点实验室 100730, Email: wningli@vip.163.com

【摘要】 随着基于医疗大数据逐步建立的人工智能(AI)辅助筛查技术的发展和普及,基于彩色眼底图片的 AI 辅助筛查系统在青光眼的筛查工作中展现了良好的灵敏度和特异度。为了规范 AI 青光眼辅助筛查系统的设计及临床应用,中华医学会眼科学分会青光眼学组联合相关专家,经过广泛调研和认真讨论,针对 AI 青光眼辅助筛查系统的数据采集、算法模型构建、硬件要求、数据集建立和标注、AI 筛查方案、AI 筛查报告要求、数据安全等制定统一标准,以便指导临床开展工作。(中华眼科杂志,2020,56:423-432)

一、AI 青光眼辅助筛查系统研发和应用的目的及意义

青光眼是一种导致视神经进行性、特征性损伤的疾病^[1],其所导致的视功能损伤往往呈进行性发展,无法恢复。青光眼的患病率和致盲率均较高,是我国乃至全球首位不可逆性致盲眼病^[2]。Quigley 和 Broman^[3]根据世界各地的流行病学资料推算,到 2020 年全世界青光眼人数将达到 7 960 万,届时中国的青光眼患者将达到 2 180 万。在致盲率方面,2010 年全球约有 840 万青光眼患者致盲,而其中中国致盲人数占患者的 1/4 以上^[3]。青光眼的高患病率及高致盲率,将给人类社会造成严重的经济负担。美国青光眼治疗的直接花费达 29 亿美元/年^[4]。而我国青光眼患者数量是美国的 5~6 倍^[5-6],据此估计我国用于青光眼的直接治疗费用无比庞大,且由青光眼致残所造成的间接经济和社会负担更难以估量。因此,青光眼已成为严重危害人类视觉健康和影响人类生活质量的公共卫生问题。

据临床统计,目前原发性闭角型青光眼(primary angle-closure glaucoma, PACG)依然是我国最常见的青光眼类型,这可能是由于统计数据来源

于就诊患者,因此症状更加明显的 PACG 患者多数被统计在内^[7]。新加坡、印度以及北京的流行病学调查结果表明,约 90% 原发性开角型青光眼(primary open-angle glaucoma, POAG)患者未被发现^[8-10],这可能是因筛查中常用的眼压测试结果无法作为 POAG 的诊断或筛查指标^[1]。在亚洲国家中,眼压正常的 POAG 患者可达 70% 以上,其中 61% 发现时青光眼病程已发展至中晚期^[11-13]。同时据 Cockburn 等估算,37% 的 POAG 患者在眼压控制正常后视野损失仍可进展^[14]。因此,在健康人群中筛查青光眼,尤其发病隐匿、后果严重的 POAG 尤为重要。近期的研究结果表明,眼底照相可作为筛查 POAG 的重要方法之一^[15];同时,为了避免 PACG 患者因对疾病认识不足而错过最佳治疗时机,在健康人群中采用眼底照相筛查 PACG 极为必要。进行青光眼筛查和防治已得到国内外许多机构认可^[16-18],因此亟须确立青光眼筛查机制和标准^[19],从而在青光眼较早期就给予患者有效的干预治疗。

青光眼筛查强调早期发现、早期诊断和早期治疗。早期确诊的患者通过合理治疗,绝大部分可以终生保有视功能。然而,目前青光眼筛查中存在缺

DOI: 10.3760/cma.j.cn112142-20200102-00003

收稿日期 2020-01-02 本文编辑 黄翊彬

引用本文:中华医学会眼科学分会青光眼学组,中国医学装备协会眼科人工智能学组.中国基于眼底照相的人工智能青光眼辅助筛查系统规范化设计及应用指南(2020年)[J].中华眼科杂志,2020,56(6):423-432. DOI: 10.3760/cma.j.cn112142-20200102-00003.



乏早期发现的有效手段、不典型及复杂性青光眼诊断难度大、随访困难等问题。传统意义上的医疗资源配置调整、病因学研究投入增加等策略难以改变现状;而且青光眼诊断仍然受到医师主观因素影响,尤其在早期青光眼检查中。同时,我国医疗资源分布差距明显,基层医疗卫生机构的青光眼诊疗水平不理想且诊治水平差异较大。因此,国内医疗行业迫切需要一种新的技术以统一乃至提高各地区的青光眼诊断筛查水平。以远程医疗会诊为代表的新型诊断方式,在一些特定环境下能够缓解基层医院医疗能力不足的问题,但是青光眼筛查要对视神经图片进行专业判读,这需要耗费大量的眼科专业人力资源。普通眼科医师对青光眼彩色眼底图片诊断的时间为每 1 000 张耗时 12~15 h,因此我国目前仅有的 3 万余名眼科医师是无法支持全国 14 亿人口的筛查。但是,若能够利用有效手段将诊断时间大大缩短,则可极大节省筛查成本,提高眼科医师的工作效率,从而使筛查工作覆盖更为广泛的人群。

人工智能(artificial intelligence, AI)技术可能是解决上述困境的一个关键突破口。近年来很多高质量的 AI 算法已成功用于辅助医学诊断,如利用深度学习系统(deep learning system, DLS)采用皮肤癌图片对皮肤癌进行区分^[20],研究 DLS 对乳腺癌淋巴结转移的诊断价值等^[21]。在眼科方面,Abrahamoff 等^[22]利用爱荷华检测项目(iowa detection program, IDP)对 Messidor-2 数据集^[23]中的糖尿病眼底病变(diabetic retinopathy, DR)彩色图片进行学习验证,结果显示 IDP 的灵敏度和特异度分别为 96.8% 和 59.4%。此外,还有利用 AI 技术通过相干光层析成像术(optical coherence tomography, OCT)预测年龄相关性黄斑变性(age related macular degeneration, ARMD)患者对于抗血管内皮生长因子(vascular endothelial growth factor, VEGF)药物治疗量的需求,从而对患者提供更为精准的临床治疗方案^[24]。在青光眼方面, Muhammad 等^[25]利用基于卷积神经网络的混合深度学习方法,评估根据视神经纤维层(retinal nerve fiber layer, RNFL)厚度进行青光眼临床诊断的准确度。2019 年, Liu 等^[26]建立了一种卷积神经网络诊断青光眼的 DLS,结果显示该 DLS 的诊断灵敏度和特异度分别为 96.2% 和 97.7%。同时,又将由深层特征形成的可视化地图叠加在输入图片上,突出显示对诊断尤为重要的模型区域,从而使 DLS 解释青光眼诊断成为可能。

总体而言,青光眼具有高患病率、高致盲率、发

展不可逆等特点,因此采用 AI 辅助诊断技术实现对青光眼视神经损伤进行早期筛查尤为重要。同时,由于基层医疗卫生机构可作为健康体检以及各种临床疾病的首诊机构,故率先在基层医疗卫生机构推广普及 AI 辅助诊断技术可以很好解决医疗资源分布不均、医师水平差异较大及效率低等问题。本指南旨在规范 AI 青光眼辅助筛查系统的设计,规范其在青光眼筛查中的使用,规范诊断结果的解读,以期未来在统一诊断标准下实现青光眼的早诊、早治及终身随访;并且希望随着 AI 青光眼辅助筛查系统在基层医疗卫生机构普及,更大范围和力度避免因早期漏诊而致青光眼患者视功能受损甚至致盲的情况发生。

二、AI 青光眼辅助筛查系统的构建和准确性要求

(一)数据集来源

建议至少收集 3 家不同等级医疗机构或图片采集中心的彩色眼底图片作为来源,以构建 AI 青光眼辅助筛查系统的数据集。数据集图片的来源可为各类公开数据集(如各级别综合医院、眼科专科医院、流行病学调查数据、互联网公开数据)。收集资料的原则应该遵循多层级、多中心、多设备来源和多种族搜集,使收集的临床资料尽可能代表真实世界中临床数据资料的质量和多元化的分布状态。同时,建议尽可能收集并完善其他眼科影像资料(OCT、相干光层析血管成像术、视野等)以及临床病历记录(眼压、用药、病程、手术、预后等)用以后续分析。获取数据集数据必须经过相关单位的伦理审查委员会批准。

(二)训练集数量

根据国家药品监督管理局医疗器械技术审评中心发布的《深度学习辅助决策医疗器械软件审批要点》,训练集应当保证样本分布的均衡性^[27]。因此,用于青光眼的 AI 辅助筛查模型训练集中的彩色眼底图片应至少涵盖符合本指南提出的图片质量控制标准且符合《青光眼视神经损害的远程筛查标准(征求意见稿)》^[28]定义的疑似青光眼或早期青光眼 2 000 例,中晚期青光眼 2 000 例,临床正常眼底 2 000 例,其他眼底疾病(非青光眼)2 000 例,图片质量不符合青光眼视神经损伤筛查标准要求 2 000 例。其中,疑似青光眼、早期青光眼和中晚期青光眼为青光眼视神经损伤阳性,临床正常眼底、其他疾病眼底为青光眼视神经损伤阴性。

训练集中的彩色眼底图片至少来自 3 个不同

等级的医疗机构或图片采集中心,每个机构每类图片不少于 400 例。若 AI 辅助筛查系统不指定或捆绑眼底照相机品牌和型号,则彩色眼底图片需来自至少 3 种以上不同品牌的眼底照相机,每种眼底照相机每类图片不少于 300 例。建议每类彩色眼底图片数量应按照被采集者性别均衡分配;每类彩色眼底图片数量在各年龄段均衡分布或按照青光眼各年龄段发病率高低分布。用于模型训练的数据计量需满足相同数据去重要求,即来自单眼的多张彩色眼底图片,其数量统计为 1 例数据。

(三)调优集数量

调优集,即验证集,是用于算法超参数调优和模型选择的数据集。根据国家药品监督管理局医疗器械技术审评中心发布的《深度学习辅助决策医疗器械软件审批要点》,调优集应当保证样本分布符合临床实际情况,且与训练集无交集^[27]。

用于青光眼 AI 辅助筛查模型参数调优的数据集至少涵盖符合本指南提出的图片质量控制标准的彩色眼底图片至少 1 200 例,其中符合《青光眼视神经损害的远程筛查标准(征求意见稿)》^[28]定义的疑似青光眼或早期青光眼至少 150 例,中晚期青光眼至少 150 例,临床正常眼底至少 750 例,其他眼底疾病至少 150 例。在数据计量中,来自单眼的多张数据计为 1 例数据。

(四)训练集及调优集标注

建议将训练集和调优集影像资料的标注工作分成两个阶段,即人工阅片标识阶段及专家鉴定阶段。在第 1 阶段中组建人工阅片小组进行标识工作,参与小组组建的眼科医师必须具备眼科专业中级以上职称,并经过统一标准的青光眼阅片培训后,取得相应评估资质方可上岗。每张彩色眼底图片的标识工作由不少于 3 名具有以上资质的眼科医师独立完成,并采用多数同意原则形成标识结果。在第 2 阶段中,建议聘请至少 1 名眼科专业副高级以上职称的眼科专家进行标注复核,并出具最终标注结果。

(五)标准测试集

根据国家药品监督管理局医疗器械技术审评中心发布的《深度学习辅助决策医疗器械软件审批要点》,测试集应当保证样本分布符合临床实际情况,且与训练集和调优集两两无交集^[27]。

用于青光眼 AI 辅助筛查模型测试的数据集至少涵盖符合本指南提出的图片质量控制标准的彩色眼底图片至少 5 000 例,其中符合《青光眼视神经损害的远程筛查标准(征求意见稿)》^[28]定义的疑似

青光眼或早期青光眼至少 500 例,中晚期青光眼至少 500 例,临床正常眼底至少 3 500 例,其他眼底疾病至少 500 例。

标准测试集的多中心和多设备标准要求同训练集,需要来自至少 5 个中心(至少 1 个中心与训练集非同源),每个中心每类彩色眼底图片不少于 100 例数据。在数据计量中,来自单眼的多张数据计为 1 例数据。每类图片 500 例数据中,至少有 200 例数据含有至少 3 张不同拍摄角度或位置的彩色眼底图片。

(六)标准测试集标注

建议使用与训练集相同的标注方法对标准测试集进行数据标注。但在第 2 阶段中,若眼科专家复核意见与人工阅片小组结果不一致,则由青光眼专家进行最终审核,并出具最终标注结果。

(七)算法模型构建

建议使用但不仅限于使用深度神经网络作为 AI 辅助诊断模型。该模型负责给出有或无青光眼视神经损伤的诊断参考意见以及相关病灶特征。在进行深度神经网络训练的过程中,训练集数据及相应标注会依次送达,并对模型参数进行调整,以获得最优的训练结果。在模型训练过程中,可以对原始彩色眼底图片进行随机左右翻转、旋转、平移和剪裁,调节彩色眼底图片色彩和明暗程度等图片增强的方式,增强样本多样性。此外,可以在训练过程中对图片引入随机噪声或其他干扰信息,增强模型对于噪声和对抗样本攻击的鲁棒性。在训练结束后使用调优集进行超参的调节,以选择最优的模型及超参。在调优结束后,模型可以在标准测试集上进行测试及性能评估。

(八)算法模型评估指标

1. 青光眼视神经损伤分类评估指标

根据国家药品监督管理局医疗器械技术审评中心发布的《深度学习辅助决策医疗器械软件审批要点》,AI 算法模型应基于测试集进行算法评估,明确算法假阴性与假阳性、重复性与再现性、鲁棒性等的评估要求^[27]。基于此项要求,青光眼视神经损伤的 AI 模型评估指标应至少涵盖灵敏度、特异度、曲线下面积(area under curve, AUC)以及模型的可重复性和鲁棒性评估。

$$\text{灵敏度} = \frac{\text{真阳性}}{\text{真阳性} + \text{假阴性}}$$

$$\text{特异度} = \frac{\text{真阴性}}{\text{真阴性} + \text{假阳性}}$$

真阳性即实际为阳性且正确判定为阳性的数据;真阴性即实际为阴性且正确判定为阴性的数据;假阳性即实际为阴性且错误判定为阳性的数据;假阴性即实际为阳性且错误判定为阴性的数据。灵敏度指标衡量实际阳性且被判定为阳性样本的比例;特异度指标衡量实际阴性且被判定为阴性样本的比例。

受试者操作特征(receiver operating characteristic, ROC)曲线,即利用AI模型预测彩色眼底图片青光视神经损伤的概率值,在0~1之间,以0.01为间隔得到不同的阈值(操作点),以此阈值对测试集中彩色眼底图片是否有青光视神经损伤进行二分类,进而得到该操作点对应的灵敏度和特异度,并且以特异度为横轴、灵敏度为纵轴进行绘图,即得到ROC曲线。ROC曲线与横轴0~1间的面积即AUC。

计算灵敏度、特异度和AUC时,在5 000例的标准测试集中随机抽取4 000例数据进行测试,每例数据随机抽取1张彩色眼底图片进行计算,随机抽样次数不少于5次,并对3个指标的置信区间进行统计。

模型可重复性是针对标准测试集中含多张彩色眼底图片的样本例,计算每例数据下AI模型的输出,至少重复3组,并计算3组结果两两之间的Kappa系数,所得到的Kappa系数的平均值。Kappa系数的计算基于混淆矩阵,具体计算公式如下:

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

$$p_o = \frac{\sum_{i=0}^c M_{ii}}{n}$$

$$p_e = \frac{\sum_{i=0}^c \left(\sum_{j=0}^c M_{ij} \times \sum_{j=0}^c M_{ji} \right)}{n^2}$$

其中, c 代表 c 个类别的预测任务, M 代表2组预测间的混淆矩阵,为 $c \times c$ 尺寸的矩阵, n 代表每组预测中样本的数量。公式中, p_o 为观测一致率, p_e 为期望一致率,即2次预测结果由于偶然机会造成的一致率。Kappa系数的理论范围值在-1~1之间,但通常在0~1之间。Kappa值在0.00~0.20之间表示几乎不存在一致性,0.21~0.39间表示极低一致性,0.40~0.59间表示较弱一致性,0.60~0.79间表示中等一致性,0.80~0.90间表示强一致性,0.90~1.00间表示一致性接近完美^[29]。

模型鲁棒性是指AI面对非正常干扰或输入的健壮性。建议测试人员和用户关注模型的鲁棒性,对正常彩色眼底图片采集过程中可能出现的各种

图片变化和噪声进行必要的归纳,如图片旋转、剪切、照明条件、对比度变化等因素,评估这些变化对算法性能的影响。

检验算法模型指标的具体要求是根据大量眼科临床医师基于彩色眼底图片进行青光眼诊断的成功率,综合考虑各项因素,建议AI辅助青光眼筛查系统的验证结果需要特异度达到85%、灵敏度达到90%以上才可投入临床使用。

2. 视杯视盘直径比值(C/D)指标评估

C/D为0~1之间的数值,C/D评估指标包括预测值与标定值之间的均方根误差(即标准差值)和R平方值(即决定系数)。

均方根误差可以衡量预测值与真值之间的偏差,反映测量的精确度。均方根误差越接近0,表明模型对于C/D的预测效果越好。

$$\text{均方根误差} = \sqrt{\frac{1}{\text{总样本数}} \sum (\text{预测值} - \text{标定值})^2}$$

R平方值,也叫决定系数,是回归预测值与标定值之间拟和程度的统计系数。R平方值介于0~1之间,越接近0,表明模型的预测结果越接近随机;越接近1,表明模型回归预测C/D的拟和效果越好。

$$R^2 = 1 - \frac{\sum (\text{预测值} - \text{标定值})^2}{\sum (\text{标定值平均值} - \text{标定值})^2}$$

三、AI青光眼辅助筛查系统的临床应用标准

(一)AI青光眼辅助筛查系统的架构和功能

AI青光眼辅助筛查系统主要面向专科诊疗经验不足或非眼科专业操作人员,经过一定培训达到标准后使用。为了进行高效筛查,及时、准确发现并提示潜在的具有青光眼风险的人群,AI青光眼辅助筛查系统应当具备的相应功能包括受检者管理、图片显示、AI评估结果、专科医师拟诊意见、报告功能。系统的部署应当适应相应医疗机构的网络情况,在保证数据安全的前提下,铺设AI青光眼辅助筛查系统。

1. 受检者管理:在系统中可以进行受检者信息的录入、修改和查看,包括受检者个人信息(如姓名、性别、年龄等)和相关检查信息(建议提供视力、眼压、血压、血糖浓度等可以辅助诊断眼底疾病的信息,尤其对于具有相关病史或主诉的患者)。

2. 图片显示:在系统中可以浏览彩色眼底图片,图片质量需保持与眼底照相机获得的原始图片同等水平,且具有缩放、移动图片的功能,以帮助专科医师进行分析。

3. 图片质量评估:对输出的彩色眼底图片质量按照本指南的图片质量控制标准进行评估,给出图片质量评价结果。

4. AI 评估结果:给出经过 AI 算法分析的有或无青光眼视神经损伤风险的评估结果以及相关参数(给出算法识别的视杯与视盘的轮廓和比值、盘沿宽度参数、RNFL 层缺损情况等),AI 初步风险评估报告应在 1 min 内给出。

5. 专科医师拟诊意见:为保证筛查结果的有效性,在向受检者出具报告前,专科医师须复核筛查结果,并给出拟诊意见。专科医师审核的拟诊意见应在 24 h 内给出。

6. 报告功能:筛查结果将以纸质或电子版的形式提供给受检者,具体内容须符合本指南的 AI 青光眼辅助筛查报告要求。

7. AI 青光眼辅助筛查系统部署方式:对于具有相应诊疗资质眼科医师的筛查机构(如医院体检中心),可以使用本地部署的 AI 青光眼辅助筛查系统,在本机构局域网构建筛查和审核体系。对于不具备诊疗资质眼科医师的筛查机构,如社区卫生服务中心,推荐使用云端远程部署的 AI 青光眼辅助筛查系统,在符合本指南的 AI 青光眼辅助筛查数据安全相关内容的前提下,利用互联网与具备眼科诊疗资质的中心指导机构构建筛查和远程审核体系。

(二)彩色眼底图片采集的硬件要求

AI 青光眼辅助筛查系统主要针对基层医疗卫生机构或非眼科专业筛查机构使用,因此建议使用半自动或全自动免散大瞳孔平面彩色眼底照相机,以缩短基层医疗卫生机构人员的培训周期,提高筛查普及率。各项指标参数要求:(1)对焦方式:可以支持自动对焦;(2)曝光方式:可以支持自动曝光;(3)视场角度:不小于 45° ;(4)免散大瞳孔可拍摄的最小瞳孔直径:可达到 3.3 mm;(5)屈光补偿范围:不小于 ± 15 D;(6)内固视点:眼底照相机的内置固视点至少可以调节至 3 个位置,分别对应以视盘为视野中心、以视盘和黄斑连线的中点为视野中心的位置,且可根据需要指定固视点;(7)图片存储格式:支持无损压缩的 TIFF 格式或无损压缩的 PNG 格式,支持压缩 JPEG 格式(压缩比例不大于 10:1),支持 DICOM 格式;(8)图片拍摄像素:不低于 1 800 万像素。

(三)彩色眼底图片的采集标准

彩色眼底图片的拍摄和采集标准种类繁多。对于青光眼筛查而言,建议选用远程眼科眼底图片质量标准作为采集标准较为合适和便捷^[30],因此提

倡使用目前远程眼科医疗中使用的每眼单张后极部 45° 眼底图片作为 AI 青光眼辅助筛查系统的图片采集标准。

1. 拍摄位置:从解剖位置而言,健康人黄斑位置在视盘颞侧 3~4 mm,黄斑中心凹则在视盘水平中线下 0.8 mm 处。因此,作为眼底图片重点结构的视盘和黄斑区,必须要求其中心在图片的垂直居中位置,而且单张眼底图片的中心必须位于视盘和黄斑之间(图 1)。建议将单张眼底图片的视盘、黄斑区、上下方视网膜血管弓 4 个区域可进行清晰判读作为图片合格的标准(图 2)。

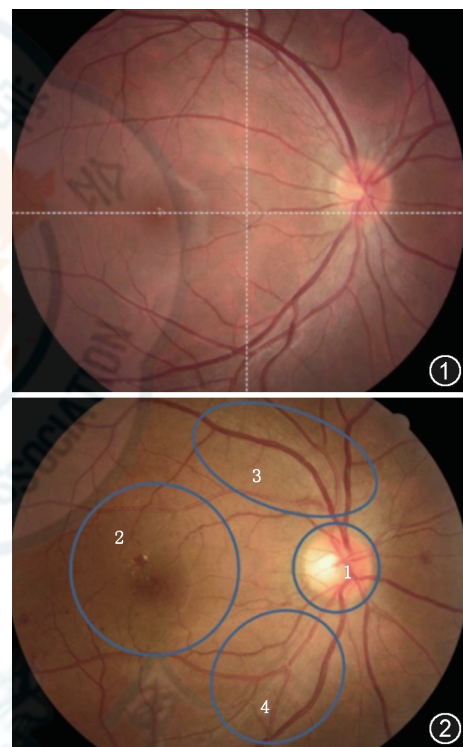


图 1 人工智能青光眼辅助筛查系统要求彩色眼底图片拍摄位置示意 图 2 人工智能青光眼辅助筛查系统判定彩色眼底图片质量的 4 个区域(蓝圈)

2. 图片质量控制标准:每名受检者双眼至少各拍摄 1 张彩色眼底图片供 AI 青光眼辅助筛查系统分析和专科医师阅片及保存随访。彩色眼底图片应严格符合本指南制定的质量控制标准。

标准彩色眼底图片的定义:眼底照相机采用单视野拍摄法,以黄斑和视盘连线的中点为拍摄视野的中心,成像至少涵盖 45° 视网膜区域,焦点在视盘和黄斑之间,并能够清晰观察到视盘、黄斑区、上下方视网膜血管弓的第一分支的眼底图片,且能够涵盖视盘周边 2 个视盘直径的区域。

眼底数据采集过程中需要拍摄医师严格按照清晰图片的拍摄标准进行操作,避免上传质量不合格的图片。考虑到图片质量对 AI 判读的影响,本指南将彩色眼底图片质量分为合格、不合格和可接受 3 类(图 3):(1)合格,即图片不存在质量问题,且拍摄位置符合本指南的彩色眼底图片采集标准对拍摄位置的要求。(2)不合格,包含严重曝光异常、严重屈光间质混浊、大片污损、信息缺失及无关图片等。(3)可接受,即存在轻微的曝光问题(边缘轻度漏光)、不影响判读的小污损、轻微的失焦或模糊图片等。

彩色眼底图片常见质量问题及相应调整规范(图 4):(1)曝光异常,包括图片过亮、过暗、出现漏光及整体过度曝光。调节照相机曝光参数,调整受检者坐姿,重新确认拍摄环境,降低检查室亮度后重新拍摄。(2)图片失焦模糊:确定是否由仪器问题或受检者眼球移动引起,调节仪器焦点重新拍摄。(3)虹膜反光:降低检查室亮度,固定视点,重新拍摄。(4)睫毛遮挡:提示受检者睁大眼睛,提高眼睑位置,重新拍摄。(5)镜头污损:检查并清理镜头,重新拍摄。(6)信息缺失:确认图片信息存储完整,重新拍摄并储存图片。(7)外眼、非眼底图片:防止上传镜头盖未打开图片、误触拍摄的环境图片以及外眼检查图片等无关图片。

(四)AI 青光眼辅助筛查方案

1. 彩色眼底图片可信度评估:图片经质量控制

后,为了使筛查结果更加准确,建议在 AI 分析前增加额外环节对图片的可信度进行评估,使图片更加规范。参考《我国糖尿病视网膜病变筛查的图像采集及阅片指南(2017 年)》^[31],同时根据青光眼的特殊眼底表现,本指南制定了青光眼彩色眼底图片采集达标标准(表 1),分为 3 项内容:(1)拍摄位置:图片的各区域是否处于正确位置,有无偏转、移位;(2)清晰度:图片中的重要区域是否清晰可见;(3)可读范围:图片是否完整拍摄到镜头捕捉区域中的全部眼底结构(是否有睫毛、异物遮挡等)。汇总 3 项内容,给出最终的可信度评估结果:(1)可信度良好:3 项均达标;(2)可信度一般,包括 3 项勉强达标,1 项达标且 2 项勉强达标,2 项达标且 1 项勉强达标;(3)可信度差:3 项中有 1 项不达标。

每次拍摄完成后均须对图片的可信度进行评估:(1)可信度良好的图片直接进入筛查步骤;(2)可信度一般的图片应针对可能出现的图片质量问题(图 4)进行合理调整后重新拍摄,取效果最好(评价最高)的图片进行分析并得出诊断结果,但须在诊断处注明“图片质量偏差,结果仅供参考”; (3)可信度差的图片进行针对性调整后重新拍摄,取效果最好(评价最高)的图片根据 1 和 2 条进行相应处理。若 3 次拍摄可信度均差,则为拍摄失败,无法进行眼底图片筛查。

2. 制定筛查方案及分级指导标准:建立在健康

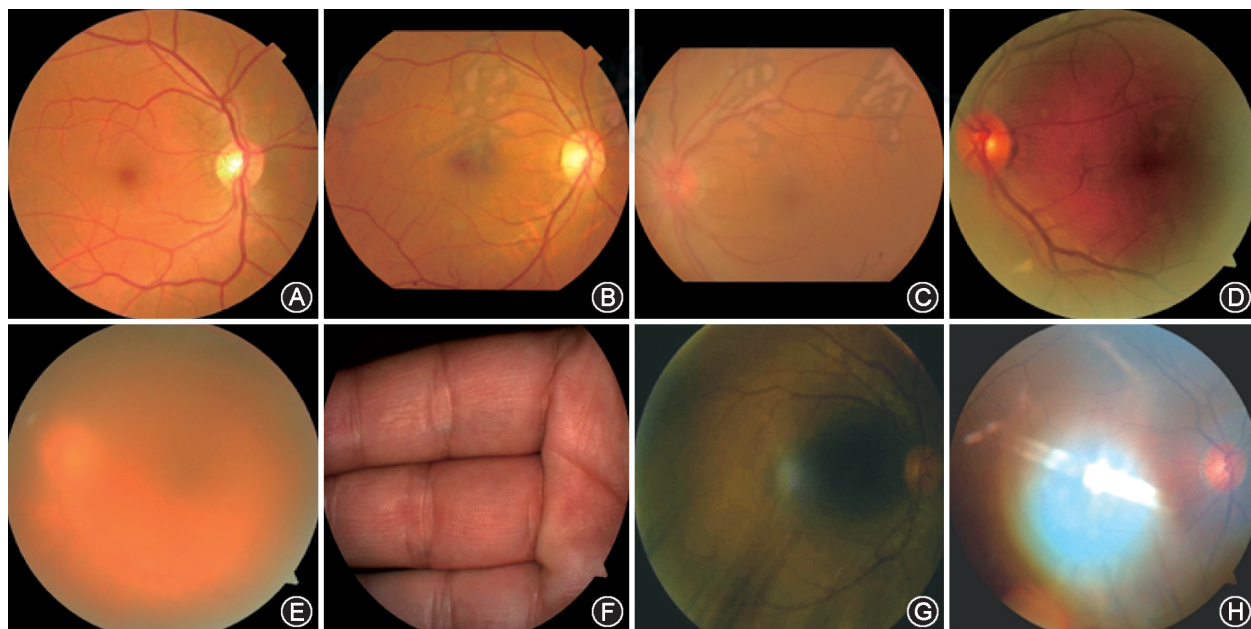


图 3 人工智能青光眼辅助筛查系统彩色眼底图片质量分类示意 A 示合格;B 示可接受,图片具有不影响判读的小污损;C 示可接受,图片受到屈光间质轻度混浊影响;D 示可接受,图片边缘漏光;E 示不合格,图片受到屈光间质重度混浊影响;F 示不合格,非眼底图片;G 示不合格,图片严重过度曝光;H 示不合格,图片有严重污损

表 1 人工智能青光眼辅助筛查系统彩色眼底图片可信度评估中不同内容的达标标准

内容	达标	勉强达标	不达标
拍摄位置	视盘中心距离正常位置<1个视盘直径	视盘中心距离正常位置1~2个视盘直径	视盘中心距离正常位置>2个视盘直径
清晰度	视盘、黄斑区及血管清晰可见	视盘清晰可见,黄斑区及血管模糊不清	视盘、黄斑区及血管模糊不清
可读范围	整个彩色眼底图片	占彩色眼底图片≥80%	占彩色眼底图片<80%

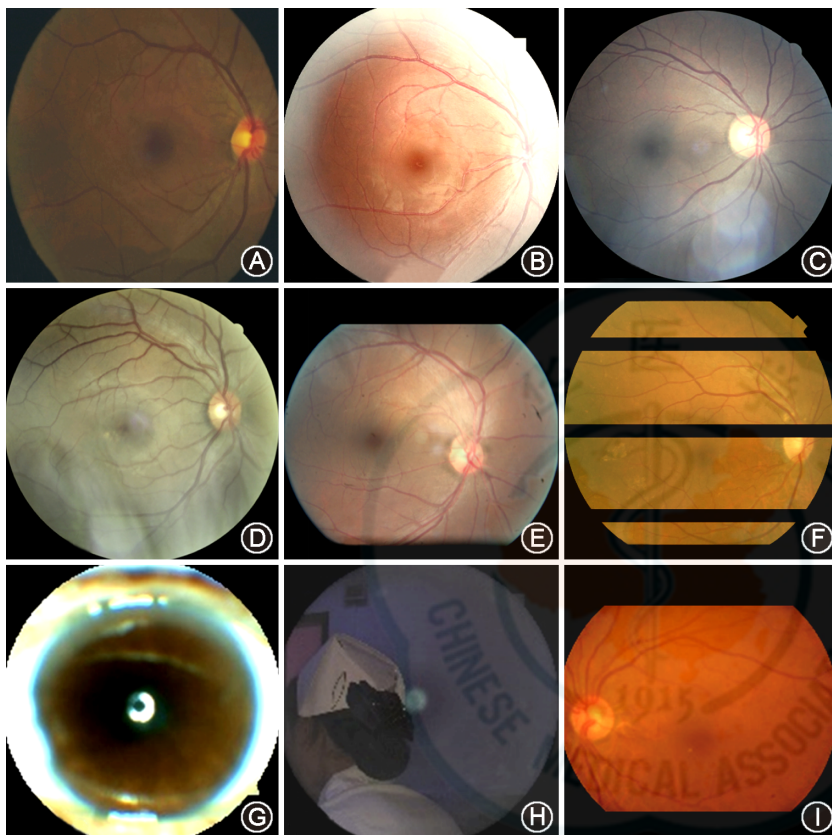


图 4 彩色眼底图片常见问题示意 A 示曝光不足,图片整体或周边过暗;B 示漏光,图片局部或整体过度曝光;C 示虹膜反光;D 示睫毛遮挡;E 示镜头污损;F 示信息缺失;G 示外眼图片;H 示无关图片;I 示失焦,图片模糊

体检中进行基于眼底照相的 AI 青光眼辅助筛查的新模式。首先确定筛查人群为进行眼底照相的健康体检人群,同时与有眼底照相检查项目的健康体检机构建立 AI 青光眼辅助筛查项目,并对结果建立确认机制。针对筛查可疑结果的复核机制:(1)确认图片是否符合质量控制标准,不符合标准者应重新拍摄。(2)符合标准的图片经过 AI 青光眼辅助筛查系统评估为可疑青光眼视神经损伤时,图片提交专业眼科医师进行人工复核。若复核意见相同,则确认筛查结果,并建议转诊进行进一步确诊。(3)若人工复核认为可疑青光眼视神经损伤不成立,即人工复核意见与 AI 青光眼辅助筛查系统评估结果不一致,则结果纳入“不确定”范畴,须并入远程阅片转诊进行进一步确诊。(4)需要强调的是,所有的 AI 青光眼辅助筛查报告均应告知体检者筛查结果

仅供参考,并须经具有相应资质的专业医师确认并签字后,方可具备医学效力。

3. 基于眼底照相的 AI 青光眼辅助筛查的临床指导原则:根据青光眼视神经病变的主要表现^[32],参照 2015 年《青光眼视神经损害的远程筛查标准(征求意见稿)》^[28],依据以下表现对青光眼视神经损伤进行判断:(1)垂直 C/D>0.6;(2)双眼垂直 C/D 相差>0.2;(3)视盘颞下方或颞上方盘沿宽度窄于鼻侧,但盘沿颜色正常;(4)与视盘颞下方或颞上方盘沿变窄相对应的局部楔形视网膜神经纤维层缺损;(5)视盘边界处线状出血。

在此基础上,基于眼底照相的 AI 青光眼辅助筛查结果的分类及对应处理原则:(1)未见青光眼视神经损伤,即无上述 5 条表现,建议每年进行 1 次眼底健康检查;(2)可疑青光眼视神经损伤,即出现上述 5 条中任意 1 条表现,应转诊眼科医师尤其青光眼专业医师。

需要指出的是,通过单张彩色眼底图片,通常可以初步判断患者有无青光眼视神经损伤^[33]。以下情形需要进行更多检查加以确诊,不属于本指南所述筛查范畴:(1)单张眼底图片存在青光眼视神经损伤,受检者需要进一步确诊;(2)少部分受检者通过单张眼底图片得到可疑青光眼视神经损伤的诊断;(3)单张眼底图片无青光眼视神经损伤表现,不意味一定不存在青光眼,可能为确实无青光眼,原发性闭角型青光眼、继发性青光眼及原发性先天性青光眼的早期,高眼压症等。

(五) AI 青光眼辅助筛查报告要求

为了使筛查报告信息尽量完整而又不过于复杂,建议筛查报告应包括以下信息:(1)受检者的基本信息,如病历号或检查编号、姓名、年龄、性别、受检眼别等;(2)眼科相关信息,如视力、眼压、眼科疾

病史和治疗史(手术、用药)等;(3)图片采集设备信息,如采集日期、设备型号、成像范围、图片存储格式等;(4)AI 青光眼辅助筛查系统相关信息,如版本号、图片分析日期等;(5)图片评估信息,如图片质量、图片可信度等;(6)筛查结果,如受检眼的诊断、患病或健康概率、青光眼特征性眼底改变描述等;(7)权责说明。

(六)AI 青光眼辅助筛查数据的存储

用于青光眼辅助诊疗的彩色眼底图片数据是来自眼底照相机拍摄的对受检者真实诊疗活动的记录,数据存储的安全和标准对于受检者病情的诊断、随访均具有重要意义。图片的存储需要满足数据安全和高效管理的要求,还须方便医院信息系统、医学影像信息系统、AI 辅助筛查系统使用。具体要求:(1)DICOM 格式或 JPEG 格式存储;采用 DICOM 格式存储时,应包含受检者信息;采用 JPEG 存储时,受检者信息应能与图片对应。(2)需要保存 AI 模型计算结果,并能关联受检者的检查信息和图片。(3)本地部署的 AI 青光眼辅助筛查系统和云端部署的 AI 辅助筛查系统,均须备份数据,在存储容量达到一定阈值时告警,以便系统管理员进行存储更新。(4)AI 青光眼辅助筛查系统与医院信息系统、医学影像信息系统对接,需要满足相关医疗协议和规范。

(七)AI 青光眼辅助筛查数据的安全

受检者的个人信息涉及隐私,具有特殊敏感性。在传输、使用数据时,必须确保医疗数据的安全性,保障数据不泄露和被恶意篡改,确保数据使用能够得到审计,保证数据访问有权限控制。参照我国医疗数据信息安全相关法律及规范要求,应做到:(1)数据使用须进行权限管理,相关工作人员和医师需要使用账户和密码进行身份验证,不同人员具备不同权限,如浏览、修改、审核、打印等;(2)数据使用需要保留日志,日志管理有授权机制,且日志需要备份;(3)数据传输需要加密,保证数据不被非法旁路获取;(4)数据存储需要加密并防篡改存储;(5)具备数据脱敏机制,数据传输给 AI 引擎前需要对相关信息进行脱敏,如姓名、手机号、身份证号、工作单位、家庭成员、联系人、收入、邮箱、生物信息等;(6)AI 科研单位和安装使用的医疗机构须有数据安全责任人,并制定数据安全的应急机制。

形成指南意见的专家组成员:

中华医学会眼科学分会青光眼学组

王宁利 首都医科大学附属北京同仁医院北京同仁眼科

中心(组长,执笔)

余敏斌 中山大学中山眼科中心(副组长)

刘旭阳 暨南大学附属深圳眼科医院(现在厦门大学附属厦门眼科中心,副组长)

陈君毅 复旦大学附属眼耳鼻喉科医院眼科(副组长)
(以下委员按姓氏拼音排序)

才瑜 北京大学第一医院眼科

蔡鸿英 天津市眼科医院

戴超 陆军军医大学西南医院西南眼科医院

范肃洁 邯郸市眼科医院

方严 安徽眼科研究所 淮南市第一人民医院眼科

郭文毅 上海交通大学医学院附属第九人民医院眼科

黄丽娜 中南大学爱尔眼科学院 深圳爱尔眼科医院

梁亮 三峡大学第一临床医学院眼科

梁远波 温州医科大学附属眼视光医院杭州院区

林丁 长沙爱尔眼科医院

潘晓晶 山东第一医科大学(山东省医学科学院)山东省眼科研究所 青岛眼科医院

申家泉 山东省立医院眼科

石晶明 中南大学湘雅二医院眼科

孙红 南京医科大学第一附属医院眼科

唐广贤 石家庄市第一医院石家庄市第一眼科医院

汪建涛 天津医科大学眼科医院(现在暨南大学附属深圳眼科医院)

王峰 哈尔滨医科大学附属第一医院眼科

王涛 首都医科大学附属北京同仁医院北京同仁眼科中心

王军明 华中科技大学同济医学院附属同济医院眼科

王凯军 浙江大学医学院附属第二医院眼科中心

王玉宏 厦门大学附属厦门眼科中心

吴慧娟 北京大学人民医院眼科

夏晓波 中南大学湘雅医院眼科

谢琳 陆军军医大学大坪医院眼科(现在重庆医科大学附属第三医院眼科)

杨新光 西安市第四医院眼科(现在西安交通大学医学院附属广仁医院眼科)

原慧萍 哈尔滨医科大学附属第二医院眼科

张纯 北京大学第三医院眼科

张旭 南昌大学附属眼科医院

张忠志 中国医科大学附属第一医院眼科

郑雅娟 吉林大学第二医院眼科

钟华 昆明医科大学第一附属医院眼科

周和政 解放军中部战区总医院眼科

周崎 中国医学科学院 北京协和医学院 北京协和医院眼科

朱益华 福建医科大学附属第一医院眼科

卓业鸿 中山大学中山眼科中心

李树宁 首都医科大学附属北京同仁医院北京同仁眼科

中心(非学组委员,秘书)

张秀兰 中山大学中山眼科中心(非学组委员,秘书)

中国医学装备协会眼科人工智能学组

甄毅 首都医科大学附属北京同仁医院北京同仁眼科中心 北京市眼科研究所(组长)

刘含若 首都医科大学附属北京同仁医院北京同仁眼科中心 北京市眼科研究所(副组长,执笔)

边桂彬 中科院自动化研究所(副组长)

李涛 南开大学信息技术科学院(副组长)

郭玮 杭州依图科技有限公司(副组长)

(以下委员按姓氏拼音排序)

白玉婧 深圳硅基仿生科技有限公司

曹绪胜 首都医科大学附属北京同仁医院北京同仁眼科中心

戴云 中国科学院光电技术研究所

丁晓伟 苏州体素信息科技有限公司

侯志尚 上海交通大学生物医学工程学院

姜泓羊 北京至真互联网技术有限公司

康宏 南开大学信息技术科学院

李劲 微软研究院

李潇 北京大学信息技术高等研究院

李经纬 华为技术有限公司

李宏智 杭州依图科技有限公司

李庆学 北京农林科学院

刘志伟 杭州依图科技有限公司

毛羽 首都医科大学附属北京同仁医院北京同仁眼科中心 北京市眼科研究所

宋世燕 华为技术有限公司

王恺 南开大学信息技术科学院

王苗苗 华为技术有限公司

王宇飞 华为技术有限公司

谢镭 华为技术有限公司

辛晨 首都医科大学附属北京同仁医院北京同仁眼科中心 北京市眼科研究所

颜子夜 杭州依图医疗技术有限公司

甄浩 上海星济信息科技有限公司

臧睿 明灏科技北京公司

张静 华为技术有限公司

张秀峰 华为技术有限公司

赵豪欣 中科院光电技术研究所

赵永强 北京华识知识产权代理有限公司

卓俊伟 华为技术有限公司

参与讨论的其他专家(按姓氏拼音排序)

常佳 深圳市健康医疗人工智能重点实验室

冬雪川 北京明达同泽科技有限公司

杜一帆 首都医科大学附属北京同仁医院北京同仁眼科中心(执笔)

方金武 中国信息通信研究院华东分院

高军 广东科鉴检测工程技术有限公司

高忠科 天津大学电气自动化与信息工程学院

李慧琦 北京理工大学信息与电子学院

李佳戈 中国食品药品检定研究院医疗器械检定所光机电室(执笔)

李建军 首都医科大学附属北京同仁医院北京同仁眼科中心 北京市眼科研究所

廖洪恩 清华大学医学院

刘江 南方科技大学计算机与工程系

刘明宇 广东省医疗器械产业技术创新战略联盟

刘重生 上海市医疗器械检测所有源室

鲁白 清华大学信息与电子学院

马锴 腾讯天衍实验室

孟德宇 西安交通大学数学与统计学院

任秋实 北京大学工学院生物医学工程系

孙炳刚 广东省精准医学应用学会

汪东生 首都医科大学附属北京同仁医院北京同仁眼科中心

王子龙 苏州体素信息科技有限公司

吴文达 腾讯医疗健康(深圳)有限公司(执笔)

徐迈 北京航空航天大学信息与电子学院

张红 中国电子标准化研究院

张莉 首都医科大学附属北京同仁医院北京同仁眼科中心 北京市眼科研究所

张道强 南京航空航天大学计算机科学与技术学院

张定康 广东省标准化研究院

张雪丽 中国信通院云计算与大数据研究所

赵盟 电气与电子工程师学会

郑海荣 中国科学院深圳先进技术研究院

郑冶枫 腾讯天衍实验室(执笔)

钟代笛 重庆大学生物工程学院

周传清 深圳湾实验室生物医学工程研究所

周子捷 飞利浦人工智能解决方案中心

声明 本文为专家意见,为临床医疗服务提供指导,不是在各种情况下都必须遵循的医疗标准,也不是为个别特殊个人提供的保健措施;本文内容与相关产品的生产和销售厂商无经济利益关系

参 考 文 献

- [1] Weinreb RN, Khaw PT. Primary open-angle glaucoma[J]. Lancet, 2004, 363(9422): 1711-1720. DOI: 10. 1016/S0140-6736(04)16257-0.
- [2] Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review[J]. JAMA, 2014, 311(18): 1901-1911. DOI: 10. 1001/jama. 2014. 3192.
- [3] Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020[J]. Br J Ophthalmol, 2006, 90(3): 262-267. DOI: 10. 1136/bjo. 2005. 081224.
- [4] Rein DB, Zhang P, Wirth KE, et al. The economic burden of

- major adult visual disorders in the United States[J]. Arch Ophthalmol, 2006, 124(12): 1754-1760. DOI: 10. 1001/archophth. 124. 12. 1754.
- [5] Liang YB, Friedman DS, Zhou Q, et al. Prevalence of primary open angle glaucoma in a rural adult Chinese population: the Handan eye study[J]. Invest Ophthalmol Vis Sci, 2011, 52(11): 8250-8257. DOI: 10. 1167/iov. 11-7472.
 - [6] 赵家良, 睢瑞芳, 贾丽君, 等. 北京市顺义县 50 岁及以上人群中青光眼患病率和正常眼压的调查[J]. 中华眼科杂志, 2002, 38(6): 335-339. DOI: 10. 3760/j. issn: 0412-4081. 2002. 06. 004.
 - [7] 林明楷, 葛坚. 青光眼住院病人的构成比变化特点[J]. 眼科学报, 1997, 13(2): 96-99.
 - [8] Tielsch JM, Sommer A, Katz J, et al. Racial variations in the prevalence of primary open-angle glaucoma. the Baltimore eye survey[J]. JAMA, 1991, 266(3): 369-374.
 - [9] Foster PJ, Oen FT, Machin D, et al. The prevalence of glaucoma in Chinese residents of Singapore: a cross-sectional population survey of the Tanjong Pagar district[J]. Arch Ophthalmol, 2000, 118(8): 1105-1111. DOI: 10. 1001/archophth. 118. 8. 1105.
 - [10] Dandona L, Dandona R, Srinivas M, et al. Open-angle glaucoma in an urban population in southern India: the Andhra Pradesh eye disease study[J]. Ophthalmology, 2000, 107(9): 1702-1709. DOI: 10. 1016/s0161-6420(00) 00275-x.
 - [11] Zhao J, Solano MM, Oldenburg CE, et al. Prevalence of normal-tension glaucoma in the Chinese population: a systematic review and meta-analysis[J]. Am J Ophthalmol, 2019, 199(3): 101-110. DOI: 10. 1016/j. ajo. 2018. 10. 017.
 - [12] Yamamoto S, Sawaguchi S, Iwase A, et al. Primary open-angle glaucoma in a population associated with high prevalence of primary angle-closure glaucoma: the Kumejima study[J]. Ophthalmology, 2014, 121(8): 1558-1565. DOI: 10. 1016/j. ophtha. 2014. 03. 003.
 - [13] 李建军, 徐亮, 张蓉秀, 等. 倍频视野检测在 40 岁以上人群中筛查青光眼的价值[J]. 中华眼科杂志, 2005, 41(3): 221-225. DOI: 10. 3760/j. issn: 0412-4081. 2005. 03. 007.
 - [14] Cockburn DM. Does reduction of intraocular pressure (IOP) prevent visual field loss in glaucoma? [J]. Am J Optom Physiol Opt, 1983, 60(8): 705-711. DOI: 10. 1097/00006324-198308000-00009.
 - [15] Xu L, Wang Y, Li Y, et al. Causes of blindness and visual impairment in urban and rural areas in Beijing: the Beijing eye study[J]. Ophthalmology, 2006, 113(7): 1134. DOI: 10. 1016/j. ophtha. 2006. 01. 035.
 - [16] Heijl A, Leske MC, Bengtsson B, et al. Reduction of intraocular pressure and glaucoma progression: results from the early manifest glaucoma trial[J]. Arch Ophthalmol, 2002, 120(10): 1268-1279. DOI: 10. 1001/archophth. 120. 10. 1268.
 - [17] Tielsch JM, Katz J, Singh K, et al. A population-based evaluation of glaucoma screening: the Baltimore eye survey[J]. Am J Epidemiol, 1991, 134(10): 1102-1110. DOI: 10. 1093/oxfordjournals. aje. a116013.
 - [18] 徐亮. 青光眼筛查是否必要和可行? [J]. 中华眼科杂志, 2003, 39(8): 454-456.
 - [19] 徐亮, 杨桦, 赵秀军, 等. 青光眼快速筛查的初步研究[J]. 中华眼科杂志, 2001, 37(1): 16-20. DOI: 10. 3760/j. issn: 0412-4081. 2001. 01. 005.
 - [20] Esteva A, Kuprel B, Novoa RA, et al. Corrigendum: Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 546(7660): 686. DOI: 10. 1038/nature22985.
 - [21] Ehteshami Bejnordi B, Veta M, Johannes van Diest P, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer[J]. JAMA, 2017, 318(22): 2199-2210. DOI: 10. 1001/jama. 2017. 14585.
 - [22] Abràmoff MD, Folk JC, Han DP, et al. Automated analysis of retinal images for detection of referable diabetic retinopathy[J]. JAMA Ophthalmol, 2013, 131(3): 351-357. DOI: 10. 1001/jamaophthalmol. 2013. 1743.
 - [23] Decencière E, Zhang X, Cazuguel G, et al. Feedback on a publicly distributed image database: the Messidor database[J]. Image Anal Stereol, 2014, 33(3): 231-234. DOI: 10.5566/ias.1155.
 - [24] Bogunovic H, Waldstein SM, Schlegl T, et al. Prediction of Anti-VEGF treatment requirements in neovascular AMD using a machine learning approach[J]. Invest Ophthalmol Vis Sci, 2017, 58(7): 3240-3248. DOI: 10. 1167/iov. 16-21053.
 - [25] Muhammad H, Fuchs TJ, De Cuir N, et al. Hybrid deep learning on single wide-field optical coherence tomography scans accurately classifies glaucoma suspects [J]. J Glaucoma, 2017, 26(12): 1086-1094. DOI: 10. 1097/IJG. 0000000000000765.
 - [26] Liu H, Li L, Wormstone IM, et al. Development and validation of a deep learning system to detect glaucomatous optic neuropathy using fundus photographs [J]. JAMA Ophthalmol, 2019, 137(12): 1353-1360. DOI: 10. 1001/jamaophthalmol. 2019. 3501.
 - [27] 国家药品监督管理局医疗器械技术审评中心. 深度学习辅助决策医疗器械软件审评要点及相关说明 [EB/OL]. [2019-06-28]. <https://www. cmde. org. cn/CL0004/19342. html>.
 - [28] 李建军, 徐亮, 王亚星, 等. 青光眼视神经损害的远程筛查标准(征求意见稿)[J]. 眼科, 2015, 24(3): 152.
 - [29] McHugh ML. Interrater reliability: the kappa statistic[J]. Biochem Med (Zagreb), 2012, 22(3): 276-282.
 - [30] 李建军, 徐亮, 彭晓燕, 等. 远程眼科单张眼底像质量标准(征求意见稿)[J]. 眼科, 2015, 24(1): 11-12. DOI: 10. 13281/j. cnki. issn. 1004-4469. 2015. 01. 005.
 - [31] 中华医学会眼科学分会眼底病学组, 中国医师协会眼科医师分会眼底病专业委员会. 我国糖尿病视网膜病变筛查的图像采集及阅片指南(2017 年)[J]. 中华眼科杂志, 2017, 53(12): 890-896. DOI: 10. 3760/cma. j. issn. 0412-4081. 2017. 12. 003.
 - [32] 徐亮. 青光眼视神经损害的三要素及其盘沿丢失的识别 [J]. 中华眼科杂志, 2006, 42(3): 196-198. DOI: 10. 3760/j. issn: 0412-4081. 2006. 03. 002.
 - [33] 李建军, 徐亮, 杨桦, 等. 远程青光眼筛查时单张眼底像的价值 [J]. 国际眼科纵览, 2015, 39(3): 216. DOI: 10. 3706/cma. j. issn. 1673-5803. 2015. 03. 015.