



面向基层的人工智能眼底彩色照相黄斑区域病变体征筛查系统规范化设计及应用指南

中华医学会眼科学分会眼底病学组 人工智能研发应用专家指导组

通信作者: 黎晓新, Email: dr_lixiaoxin@163.com; 张明, Email: zhangmingscu@126.com;

许言午, Email: ywxu@ieee.org

【摘要】 随着人工智能 (AI) 概念的推广和技术的发展, 在医疗领域逐渐出现基于AI算法的疾病筛查系统。这类系统可用于疾病的大规模筛查, 以缓解基层医疗卫生压力。近年来, AI算法在眼底彩色照相黄斑区域的病变体征分析和识别中展现了良好的性能, 随着时间的推移, 未来定会出现适用于基层筛查的眼底病变体征筛查系统。因此, 为了规范基于AI算法构建的黄斑区域病变体征筛查系统的设计及临床应用, 中华医学会眼科学分会眼底病学组联合相关专家, 经过调研问题、讨论生产证据、召开指南研讨会等环节, 针对基层筛查系统的黄斑区域和病变体征定义、AI应用场景、算法模型构建、数据集的建立和标注、架构和功能设计和图像数据采集等方面制定统一标准, 以便指导筛查工作的开展。

【关键词】 指南; 人工智能; 黄斑; 眼底病变; 筛查

DOI: 10.3760/cma.j.cn511434-20220616-00364

The standardized design and application guidelines: a primary-oriented artificial intelligence screening system of the lesion sign in the macular region based on fundus color photography

Ocular Fundus Diseases Group of Chinese Ophthalmological Society, Expert Group for Artificial Intelligence Research, Development, and Application

Corresponding author: Li Xiaoxin, Email: dr_lixiaoxin@163.com; Zhang Ming, Email: zhangmingscu@126.com; Xu Yangwu, Email: ywxu@ieee.org

【Abstract】 With the popularity and development of artificial intelligence (AI), disease screening systems based on AI algorithms are gradually emerging in the medical field. Such systems can be used for primary screening of diseases to relieve the pressure on primary health care. In recent years, AI algorithms have demonstrated good performance in the analysis and identification of lesion signs in the macular region of fundus color photography, and a screening system for fundus lesion signs applicable to primary screening is bound to emerge in the future. Therefore, to standardize the design and clinical application of the macular region lesion sign screening systems based on AI algorithms, the Ocular Fundus Diseases Group of Chinese Ophthalmological Society, in collaboration with relevant experts, has developed this guideline after investigating issues, discussing production evidence, and holding guideline workshops. This guideline aims to establish uniform standards for the definition of the macular region and lesion signs, AI adoption scenarios, algorithm model construction, datasets establishment and labeling, architecture and functions design, and image data acquisition for the screening system to guide the implementation of the screening work.

【Key words】 Guidebooks; Artificial intelligence; Macula lutea; Fundus lesions; Screening

DOI: 10.3760/cma.j.cn511434-20220616-00364

一 人工智能 (AI) 黄斑区域病变体征筛查系统研发和应用的目的是意义

黄斑区主要与精细视觉及色觉等视功能相关, 黄斑区出现病变, 是视力下降的主要因素之一。由于黄斑区域的敏感性, 该区域的损伤可造成中心视力的下降, 对患者的生活质量影响较大^[1]。早期眼科筛查是降

低病变风险、避免视力丧失的积极手段。数据显示, 全国约有4.48万眼科医生, 这意味着平均1个眼科医生要为约3.125万人提供医疗服务^[2]。由此可见, 供需的极度不平衡, 使我国眼科医生在眼底疾病筛查中面临较大的工作压力, 同时也存在一部分患者无法得到及时的筛查及治疗的情况。因此, 借助AI、自动化、信息化技术实现大基数人口的眼底筛查是必然的发展趋势^[3]。

眼底彩色照相（以下简称为眼底彩照）是一种方便、非接触式的眼科检查，可观察视神经、视网膜、脉络膜以及屈光间质中是否存在病变，是目前筛查眼底疾病最具成本效益的成像方式^[4]。黄斑区域内可发生的疾病种类众多，表现多样，如老年性黄斑病变、中心性浆液性脉络膜视网膜病变、中心性渗出性脉络膜视网膜病变及视网膜静脉阻塞等^[5]。眼底彩照能提供的病变信息有限，临床上难以基于单一眼底彩照检查给出准确的疾病诊断结论。但在基层筛查应用场景中，筛查系统仅需给出黄斑区域内是否出现可疑病变体征的结论，若出现一定程度的病变体征，则认为受检人存在黄斑区域内眼疾的患病风险，需到医院进行进一步的检查。由此可见，基层筛查可仅基于眼底彩照进行，并且可只在病变体征维度进行分析。因此，本指南拟规范面向基层的眼底彩照黄斑区域病变体征筛查系统的设计及应用。由于非散瞳眼底彩照可满足检查区域的呈现，因此一般情况下，本指南讨论的眼底彩照为非散瞳情况下采集的。

随着基于医疗大数据的逐步建立以及AI筛查技术的发展和运用，很多高质量的AI算法已成功应用于医学辅助诊断，如利用图像对皮肤癌进行判别^[6]，分析乳腺癌淋巴结转移^[7]等。在眼科方面，2017年，Gargeya和Leng^[8]利用75 137例糖尿病患者的眼底彩照训练AI模型，使其检测糖尿病视网膜病变的灵敏度和特异性分别达到了94%和98%。此外，基于眼底彩照的AI模型在青光眼的筛查工作中也展现了良好的性能^[9]。

目前在世界范围内，黄斑区域病变体征的AI识别及筛查尚未建立统一的数据来源、模型构建、临床评价等方面的标准，因此限制了AI筛查系统在该领域的临床应用。本指南制定小组在中华医学会眼科学分会青光眼学组和中国医学装备协会眼科人工智能学组发布的《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）》^[9]以及中国医药教育协会智能医学专业委员会智能眼科学组和国家重点研发计划“眼科多模态成像及人工智能诊疗系统的研发和应用”项目组发布的《基于眼底照相的糖尿病视网膜病变人工智能筛查系统应用指南》^[10]的基础上，参考了中国食品药品检定研究院组织起草的《人工智能医疗器械行业标准》^[11-13]和中国质量检验协会发布的《眼底彩照标注与质量控制规范》（T/CAQI 166-2020）^[4]，制定了可供系统研发工程师和基层医生等相关人员参考的筛查系统设计与应用指南。具体而言，本指南旨在对黄斑区域及相关病变体征定义、AI应用场景、算法模型构建、数据集的建立和标注、筛查系统架构和

功能的设计、图像数据的采集等方面提供参考性意见，以期规范AI系统在黄斑区域病变体征基层筛查中的应用，推动我国眼底病早期筛查水平的整体提升。本指南内容未来会根据医疗器械领域新法规、新标准以及黄斑区域内疾病诊疗新指南的更新进行及时迭代。

二 黄斑区域和病变体征

（一）黄斑区域

根据年龄相关性眼底病变研究^[15]，在讨论老年性黄斑病变时，将黄斑区域定义为以中心小凹为中心、半径为2倍视盘直径（DD）的圆形区域。其黄斑区域与视盘边缘几乎相切（图1）。此外，在解剖学上，黄斑区定义为颞侧上下血管弓内5.5 mm范围内的区域^[16]。

综合上述黄斑区域定义，本指南定义黄斑区域为中点在黄斑中心小凹的圆形区域，其半径为2 DD、黄斑中心小凹到上下血管弓最小距离以及黄斑中心小凹到视盘边缘最小距离中的最小值（图2）。

（二）黄斑区域病变体征

在黄斑区域病变体征的早期筛查场景中，我们不但需要检测高风险疾病相关体征；同时需要关注潜在风险体征和潜在病发区域。本指南对黄斑区域内病变体征进行定义，疑似出现以下任意一项体征时建议转诊；若确定出现至少一项下述体征，则确定转诊。

1、直径大于125 μm 的玻璃膜疣（drusen）^[17]。玻璃膜疣是胶样或透明的小体，是视网膜色素上皮（RPE）细胞异常代谢产物在视网膜上的异常沉积所致。125 μm 相当于视盘颞下方边缘处静脉管径大小。临床上将玻璃膜疣分为4型^[18]：（1）硬性疣：表现为边缘清晰的黄色圆点；（2）软性疣：表现为边缘不清晰互相融合的病灶，伴有RPE色素改变；（3）混合性疣：兼有上述两型的临床表现；（4）退化性疣：眼底可见不规则形如“地图”或“晕轮状”萎缩区，并逐渐扩张，可因钙化而呈现为白色。

2、地图样萎缩（GA）^[17]。其病理表现为RPE、脉络膜毛细血管及其上光感受器细胞丢失，眼底可见边界清晰的单灶或多灶脉络膜视网膜萎缩区，其中脉络膜组织的可见度增加，周围常见大小不等的玻璃膜疣。

3、伴有至少一个直径大于125 μm 的出血或渗出的新生血管^[17]。黄斑新生血管可分为3类^[19]：（1）1型黄斑新生血管，其起源于脉络膜的新生血管，终止于RPE下；（2）2型黄斑新生血管，其起源于脉络膜的新生血管，穿过RPE，在视网膜神经纤维层下生长，眼底彩照可见橘红色样病灶；（3）3型黄斑新生血管，其起源于视网膜中间层的新生血管，向RPE生长，是从视

网膜循环向下生长至外层视网膜的血管复合体,表现为视网膜下液和渗出、视网膜退化、视网膜脱离、视网膜下新生血管膜。表现为黄斑中心小凹或中心小凹旁某一象限内有一不规则的类圆形病灶,呈灰白色或黄白色,病灶周围可见出血或反光晕。此外,视网膜新生血管表现为起源自视盘表面及视网膜的小静脉,沿视网膜表面生长,在有玻璃体粘连的部位可长入玻璃体内,并含有数量不等的纤维组织。

4、存在至少一个渗出灶直径大于 $125\ \mu\text{m}$ 的渗出(exudation)^[20]。硬性渗出表现为边界清晰的腊样色斑点或斑块,可数个或成堆出现;软性渗出又称为棉绒斑或棉絮状白斑,表现为边界不清的灰白色斑。

5、存在至少一个出血灶直径大于 $125\ \mu\text{m}$ 的出血(hemorrhage)^[20]。按出血所在的视网膜层次的不同,可分为视网膜前出血、视网膜浅层出血、视网膜深层出血、RPE下出血以及多层次同时受累。视网膜前出血表现为鲜红色,多呈“舟状”,形成典型液面。视网膜浅层出血多呈鲜红色“火焰状”,可表现为Roth斑,即出血性病灶中央可见白色小点,为毛细血管破裂时产生的血小板-纤维蛋白血栓^[21];视网膜深层出血呈点状、圆形的暗红色小斑片,沿着神经走行纵向扩散。视网膜下出血多成暗红色,RPE下出血多成棕红色。

6、瘢痕(scar)^[20]。渗出物在视网膜间隙刺激RPE与胶质细胞增生,最终形成纤维化瘢痕组织,表现为边界清晰的类圆形黄褐色斑块^[22]。

7、色素变动(pigment mottling)。其可分为色素增生(hyperpigmentation)与色素缺失(depigmentation)两种。色素增生即累及黄斑区域且有影响视力可能的色素增生体征^[20],表现为有突或分支的黑点或白点,逐渐聚集为墨黑色或白色的“蜘蛛状”或骨细胞样斑点,可为不规则条状。色素缺失表现为界限清楚的淡白色或淡黄色斑块。

8、黄斑裂孔(macular hole)^[23]。黄斑部呈一圆形或椭圆形的边缘锐利的孔洞,偶见不规则形,裂孔大小不一。

9、黄斑前膜(epiretinal membrane)^[24]。眼底黄斑区及附近视网膜表面可见一层半透明、菲薄或增厚发灰的无血管性增生膜,视网膜出现褶皱,附近视网膜小血管纤曲。

10、黄斑水肿(macular edema)^[17]。根据糖尿病视网膜病变早期治疗研究^[25],有临床意义的黄斑水肿表现为:(1)视网膜水肿增厚,范围在黄斑中心 $500\ \mu\text{m}$ 区域以内;(2)硬性渗出(附近视网膜增厚),侵

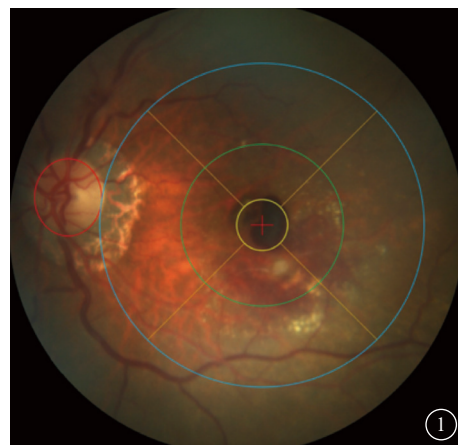


图1 黄斑区域示意图

犯黄斑中心 $500\ \mu\text{m}$ 区域以内;(3)视网膜增厚范围大,位于黄斑任一象限,但有部分侵犯黄斑中心区1DD之内。囊样黄斑水肿^[26-27]表现为黄斑中心小凹反光反射弥散或消失,视网膜增厚,水肿区呈不同程度反光增强或“绸缎样”反光。晚期黄斑水肿呈“蜂窝状”或囊状外观,囊壁视网膜厚薄不均匀,可见“蜂窝状”内部分隔及血管暗影。

11、视网膜脱离(retinal detachment)^[28-29]。其包括孔源性视网膜脱离、渗出性视网膜脱离及牵拉性视网膜脱离。孔源性视网膜脱离中,浅脱离的视网膜表现为色泽变淡,与正常的橘红色呈明显界限;局部高度脱离的视网膜呈灰白色“球形”突入玻璃体腔;广泛脱离的视网膜起伏不平,其上血管纤曲爬行;陈旧的视网膜脱离上下可见增生膜,呈不规则条索状,视网膜表面或全层皱襞形成。渗出性视网膜脱离表面较光滑,无牵拉皱褶。牵拉性视网膜脱离表面可见增生膜或粘连的机化组织。

12、眼内占位性病变(intraocular space-occupying lesions)。该病变表现多样,临床上需结合其他辅助检查协助诊断。

13、其他。非上述类别体征的黄斑区异常。

上述体征与黄斑区域常见病变的对应关系如表1所示。

值得注意的是,本指南给出了上述13种黄斑区域病变体征。筛查系统设计人员在定义系统可筛查病变体征时,请根据产品设定,按实际需求参考本指南及医学专家建议。

三 AI算法在黄斑区域病变体征筛查中的应用场景

(一) 病变体征检测

AI算法可基于输入的眼底彩照,提供图像中是否存在本指南第2部分第2节“黄斑区域病变体征”定义

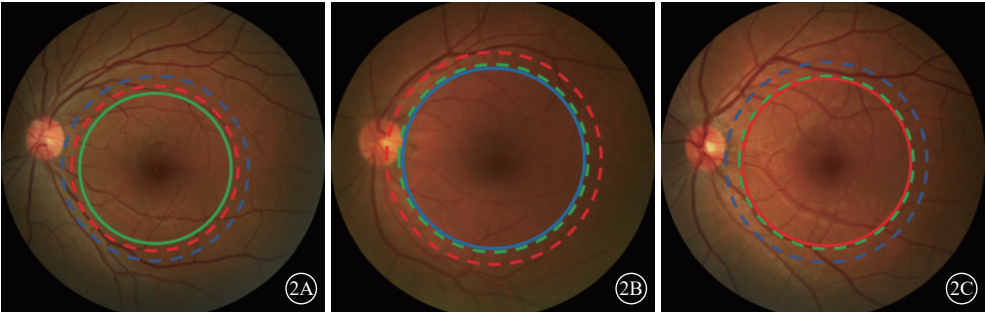


图2 黄斑区域确定方法示意图。实线显示的圆形区域为本指南中定义的黄斑区域。2A示黄斑区域为绿色圆区域，半径为2个视盘直径；2B示黄斑区域为蓝色圆区域，黄斑中心小凹到视盘边缘的最小距离；2C示黄斑区域为红色圆区域，黄斑中心小凹到上下血管弓（主静脉血管）的最小距离

表1 指南中定义的黄斑区域病变体征与常见眼底疾病的对应关系

体征	涉及黄斑区域病变
玻璃膜疣	AMD（萎缩型）
地图样萎缩	AMD（萎缩型）
新生血管	AMD（渗出型）、中心性渗出性脉络膜视网膜病变、阻塞性疾病、病理性近视、DR
渗出	AMD（渗出型）、中心性渗出性脉络膜视网膜病变、DR、高血压性视网膜病变
出血	AMD（渗出型）、中心性渗出性脉络膜视网膜病变、DR、高血压性视网膜病变、息肉样脉络膜血管病变、视网膜血管性疾病
瘢痕	AMD（渗出型）、中心性渗出性脉络膜视网膜病变、其他疾病
色素变动	AMD（萎缩型）、RPE撕裂、中心性渗出性脉络膜视网膜病变、视网膜色素变性
黄斑裂孔	高度近视、囊样黄斑水肿、视网膜变性类疾病、日蚀性视网膜病变
黄斑前膜	视网膜血管病、DR、视网膜静脉阻塞、后葡萄膜炎、Bechet病、Eales病、玻璃体积血、Von-Hippel病、眼内肿瘤、眼球外伤、眼内炎
黄斑水肿	AMD、视网膜静脉阻塞、视网膜色素变性、黄斑前膜、DR、CSC、中心性渗出性视网膜脉络膜病变
视网膜脱离	PDR、视网膜静脉周围炎、视网膜脉络膜肿瘤、葡萄膜炎、脉络膜炎等
眼内占位性病变	脉络膜骨瘤、脉络膜黑色素瘤、脉络膜血管瘤、转移癌、寄生虫等
其他	非上述类别体征的黄斑区异常

注：AMD：老年性黄斑变性；DR：糖尿病视网膜病变；PDR：增生型DR；CSC：中心性浆液性脉络膜视网膜病变；RPE：视网膜色素上皮

的病变体征的结论。同一张眼底彩照上可检出多种病变体征，因此该任务是一个多标签的分类任务，即标签之间不互斥^[30]。AI模型可对输入图像同时进行是否有玻璃膜疣、是否有地图样萎缩、是否有新生血管等情况的判别（图3）。在临床应用中，可针对每种病变体征的分类结果进行评估，评估指标可用灵敏度、特异性、Kappa值等。一般认为，与临床诊断报告中病变体征是否出现这一参考标准（参考标准：诊断和治疗过程或基于标注过程建立的基准，可包含疾病、生理状态或生理异常以及位置和程度等信息标签^[1]）比较的Kappa值 ≥ 0.80 时，二者的一致性非常好；Kappa值介于0.61~0.80之间，二者的一致性较好；Kappa值介于0.41~0.60之间，二者的一致性为中等；Kappa值介于0.21~0.40之间，二者的一致性勉强可接受；当Kappa值 ≤ 0.20 时，则认为几乎没有一致性^[31]。AI算法

的性能目标是使模型预测结果达到与专业人员诊断结果非常好的一致率，即Kappa值达到0.8以上。

（二） 病变体征定位

AI算法可对输入眼底彩照中的指定病变体征进行定位。如针对本指南第2部分第2节“黄斑区域病变体征”定义的病变体征，AI模型可从图像中定位出具体的病灶出现区域（图4）。图中交叠的检测框说明交叠区域内存在不止一种病变体征。请注意，按照计算机视觉算法，目标检测框多为矩形框，检测出的病灶在矩形框内，但不一定完全充溢整个矩形框。定位模型在临床中的应用目的通常是提供目标病变或组织的位置信息，提醒医生重点观察该区域。这种应用的效果评估通常与计算机视觉中目标检测^[32]任务的评估指标一致，具体指标将在本指南第4部分第5节第2点“病变体征定位”中介绍。

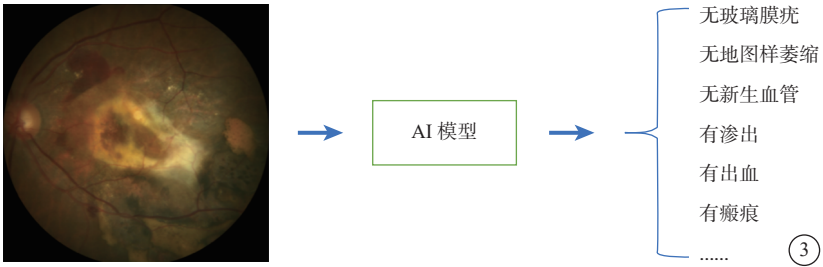


图3 AI模型对眼底彩照进行多标签的分类。AI：人工智能



图4 AI模型对眼底彩照中病变区域进行定位。图中交叠的检测框说明交叠区域内存在不止一种病变体征。粉色线内区域：出血；蓝色线内区域：瘢痕；绿色线内区域：渗出。AI：人工智能

（三） 病变体征分割

AI算法可对输入眼底彩照中的指定病变体征进行分割。如针对本指南第2部分第2节“黄斑区域病变体征”定义的病变体征，AI模型可从图像中分割出具体的病变轮廓^[33]（图5）。分割模型在临床中的应用目的是可视化或量化目标病变或组织，辅助医生观察以及诊断。这种应用的效果通常与计算机视觉中分割任务的评估指标^[34]一致，具体指标将在本指南第4部分第5节第3点“病变体征分割”中介绍。

（四） AI算法的其他应用

AI算法除了用于病变体征的分析，还可以扩展至疾病维度的应用中，本指南给出疾病分类和分级两个应用参考，可作为筛查系统的后续扩展功能。病变体征筛查系统无需包含这两种疾病维度的功能。

1、疾病分类。AI算法可对输入图像进行是否有黄斑区域疾病的判别，该功能的结论是有或无黄斑区域疾病，属于二分类任务。若需细分到具体黄斑区域疾病种类，可扩展成多分类或多类别分类任务。该场景中AI算法和临床中疾病诊断结果参考标准的Kappa值可衡量AI算法与专业医生诊断结果的一致性。

2、疾病分级。AI算法可根据输入图像，对疾病进行识别与分级（分期）（图6）。疾病分级是一个多分类任务。临床应用中的评估方式与疾病分类相似，可用灵敏度、特异性、准确率、Kappa值等评估测度来衡量应用效果的好坏。

四 AI黄斑区域病变体征筛查系统的算法构建和准确性要求

（一） 数据收集与质量控制

数据作为AI“思考”和“决策”的基础，是AI技术得以实现的基石。在AI黄斑区域病变体征筛查系统构建、验证、生产等过程中，均离不开大量数据集的支持。眼底数据采集过程需要拍摄医生/技师严格按照标准图像的拍摄要求进行操作，避免上传质量不合格的图像。本指南将彩色眼底图像质量分为合格、可接受和不合格三类^[35]。合格：图像不存在质量问题，拍摄位置符合本指南的彩色眼底图像采集标准对拍摄位置的要求（图7A）；可接受：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像等（图7B，7C）；不合格：严重曝光异常、严重屈光间质混浊、大片污损、信息缺失及无关图像等（图8）。

每只受检眼拍摄至少1张可供AI系统分析和医生阅读的图像并保存，要求位置正确，对焦清晰，曝光适中，没有间质混浊图像，能够清晰显示视网膜、黄斑区及视盘结构，具体质量要求如下^[36]：（1）除玻璃膜疣、色素异常或新生血管等眼底黄斑疾病相关体征外，图中90%的血管可以辨认。（2）主要眼底结构位置正确，使用单视野图像筛查时，图像视野范围水平与垂直方向均不小于45°，黄斑中心小凹距离图的边缘

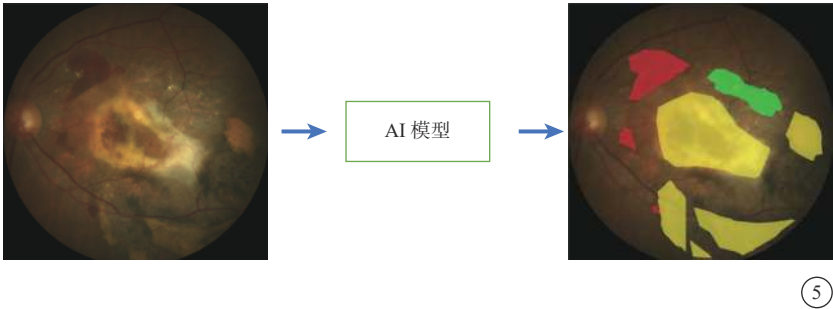


图 5 AI模型可对眼底彩照中病变轮廓进行分割。红色区域：出血；黄色区域：瘢痕；绿色区域：渗出。AI：人工智能

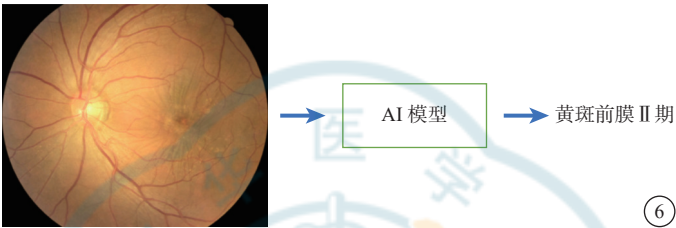


图 6 AI模型可对眼底彩照中的黄斑前膜进行分级识别。AI：人工智能

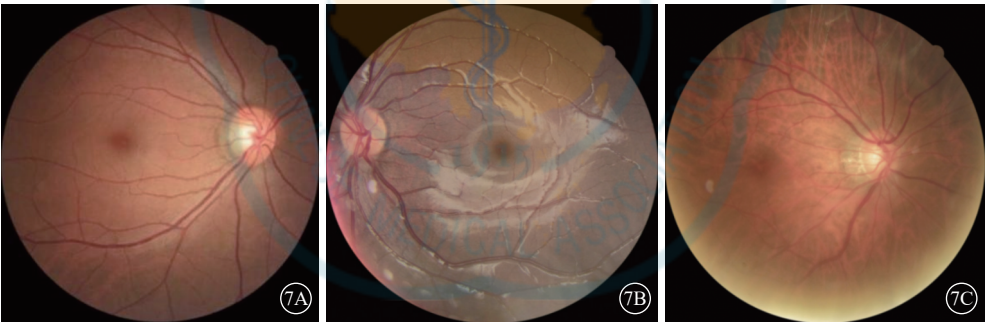


图 7 合格及可接受眼底彩照示意图。7A示合格眼底彩照。7B、7C示可接受眼底彩照，7B示周边可见反光，黄斑区不受影响；7C示边缘轻度漏光，不影响结果判读

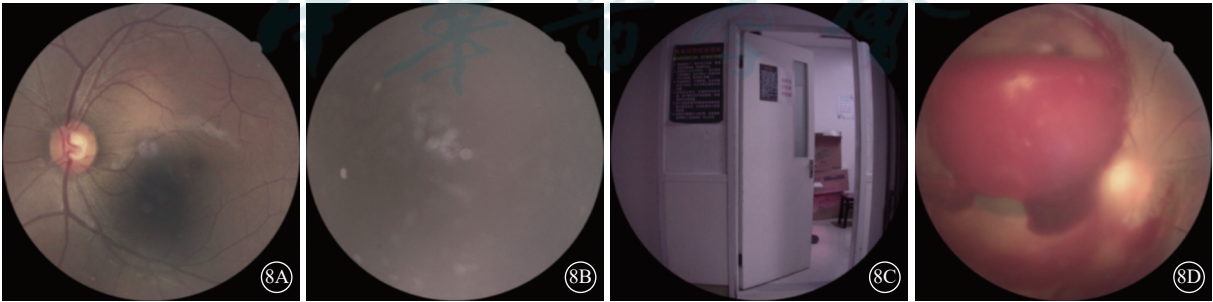


图 8 不合格眼底彩照示意图。8A示曝光不足；8B示屈光间质严重混浊；8C示无关图像；8D示黄斑区大片遮盖

超过2 DD，视盘距离图的边缘超过2 DD；使用双视野图像时，每个视野水平与垂直方向均不小于45°，黄斑区图像要求黄斑中心小凹距离图像中心<1.5 DD，视盘区图像要求视盘中心距离图像中心<1.5 DD；视盘与黄斑中心连线和水平线的夹角不大于24°。（3）成像范围内无影响判读的暗影和（或）高亮反光区域。（4）曝

光适度，无过曝光或欠曝光。（5）无镜头污渍、眼睑和（或）睫毛等遮挡影，无运动伪影。（6）无图像中未显示患眼、拍摄非眼底范围的眼底彩照等图像错误。

图像未达到以上质量要求时，需进行如下调整：

（1）主要眼底结构位置不正确：调整受试者坐姿，调

节固视点，确认患者有无斜视或其他眼部异常情况，重新拍摄图像。（2）过曝光、欠曝光及对焦错误：调节眼底相机曝光与对焦设定，重新拍摄图像；如观察到图像过暗，应确认患者瞳孔大小，并相应缩短患眼暴露在明亮灯光下的时间，降低检查室亮度。（3）瞳孔过小，无法获得满意的小瞳孔眼底图像结果：临床医生排除散瞳的禁忌症后考虑散瞳后采集眼底图像。（4）眼睑、睫毛遮挡：提示患者在拍照过程中睁大眼睑，必要时协助患者提高眼睑，并重新拍摄图像。（5）虹膜反光：提示患者紧盯固视点，不要移开视线，重新拍摄图像。（6）镜头污渍：检查并清理镜头。（7）信息缺失：确认图像信息存储完整，重新拍摄并储存图像。（8）外眼、非眼底图像：防止上传镜头盖未打开图像、误触拍摄的环境图像以及外眼检查图像等无关图像。

本指南建议数据集中图像的收集需考虑以下纳入与排除标准：纳入标准：（1）单视野/双视野拍摄的眼底彩照；（2）图像质量为合格或可接受。排除标准：（1）图像质量为不合格；（2）图像中存在治疗痕迹。

（二）数据集构建

本指南按照AI模型构建、模型效果检验以及临床试验三个环节将筛查系统设计和应用过程中涉及到的数据集分为三大类（图9）。其中，模型构建数据集中调优集可称验证集，测试集可称内部测试集，模型效果检验以及临床试验数据集可称外部测试集。各类数据集间不可出现重复的样本以及重复的被检者。另外，医疗数据的收录需通过伦理审批并注意提前将数据做脱敏处理。指南参考人员可根据实际应用场景和条件，选择性参考本指南相关内容。

1、模型构建数据集来源

模型构建数据集的作用是为模型训练、调优和自测提供带标签的数据，使模型学习到最优参数，实现好的筛查效果。由于不同眼底彩色相机存在拍摄角度、成像色彩、清晰度等方面的差异，为了提升AI模型的泛化能力，需使收集的图像数据采集自多种机型、多种拍摄条件。泛化能力指AI算法对新鲜样本的适应能力。AI技术中机器学习的目的是学到隐含在数据背后的规律，对具有同一规律的数据集以外的数据，经过训练的模型也能给出合适的输出。综合参考糖尿病视网膜病变^[10]和青光眼^[9]两种眼疾筛查的AI系统应用指南，本指南建议收集的数据集来源不少于3家医疗机构（满足采集环境、采集人员操作多样性），涉及机型不少于3种（满足采集设备多样性）。由于本指南针对的是中国人群眼疾筛查应用场景，所以模型构建的数据来源种族为单一的中国人群。同时，数据集

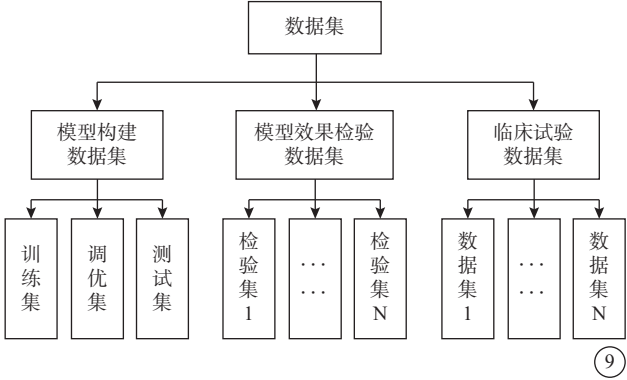


图9 筛查系统设计与应用过程中涉及到的数据集分类示意图

中样本需覆盖多民族、多地区、多年龄层人群。AI模型构建过程中，数据集的来源可包含各类公开数据集（如各级别医院、眼科中心、流行病学调查数据和互联网公开数据等），但必须符合本指南第4部分第3节介绍的标注办法，不符合的必须重新标注。所有数据的获取须经相关单位的伦理审查委员会批准。

值得注意的是，为了符合筛查系统应用时输入的图像可能存在质量高低不一致的实际情况，在收集数据时，不可人为挑取质量好的眼底图像。在各家医疗机构收集图像时，应当从以往队列中随机抽取样本构建数据集。

2、模型构建数据集划分

本指南建议将模型构建数据集采用留出法^[37]进行划分，留出法是直接将数据集划分为三个互斥的集合，分别为训练集、调优集和测试集。参考Google公司在JAMA发表的糖尿病视网膜病变识别论文^[38]中使用的的数据划分比例：训练集和调优集总比例为91.56%，测试集比例为8.35%。本指南建议在模型构建数据集中，训练集和调优集占比在90.00%左右，测试集占比为10.00%左右。对于训练集和调优集的划分，可根据具体情况，按照AI技术应用中常用的8：1或7：2比例划分。请注意，此处的测试集仅用于模型构建环节中的自测。另外，除了训练集、调优集和测试集三个集合的划分外，筛查系统也可根据模型特性，将数据集仅划分为训练集（包括调优集）和测试集。

数据集划分建议注意的一点是保证训练集、验证集和测试集数据分布的一致性^[39]，即所包含的各类别图像数占比保持一致，这里的类别需考虑图像所属病变体征类别、受试者年龄与性别类别、图像采集机构与机型类别等。同时，建议测试集符合预期应用场景的数据分布，并且数据划分过程中需注意同一受试者的多张眼底彩照应划分至训练集、验证集或测试集中的某一个集合，不可出现同一受试者的多张图像划分至

不同数据集的情况。

3、模型构建数据集数量

由上一节模型构建数据集划分可知,本指南建议测试集占比达到总数据集数量的10.00%。下面本指南将根据AI模型的预期效果,推导测试集中各病变体征类别图像需要的数量,以此提供模型构建数据集总数量的建议。

根据评估指标的置信区间样本量计算公式^[40]计算所需各类别样本数量:

$$N = \frac{[Z_{1-\alpha}]^2 P(1-P)}{\Delta^2} \quad (1)$$

公式(1)中: Z 为可信水平的 Z 统计量, Δ 为允许误差, P 为预期评估指标(灵敏度或特异性), N 为所需样本量。设定显著性水平 $\alpha=0.05$ (双侧),则预期评估指标的允许误差 Δ 为5%:

(1) 当AI模型预期的灵敏度为90%、特异性为90%时,则各病变体征所属图像对应的阳性或阴性样本量最小分别为:

$$N = \frac{1.96^2 \times 0.9 \times (1-0.9)}{0.05^2} \approx 138 \quad (2)$$

即,测试集中包含本指南第2部分第2节“黄斑区域病变体征”中定义各病变体征和无目标体征样本分别至少需要138张图像。(2) 当AI模型预期的灵敏度为85%、特异性为85%时,单类最少需196张图像。

参考上述计算结果,本指南建议模型构建数据集中测试集里涉及第2部分第2节“黄斑区域病变体征”定义的每类病变体征的样本数量至少为200张。参考青光眼辅助筛查系统规范化设计及应用指南(筛查阳性样本:正常眼底样本:其他眼疾样本=2:7:1),建议^[4]测试集中包含的临床正常眼底图像至少700张,包含有其他眼底疾病的眼底图像至少100张(若设计的筛查系统可筛查多种病变体征,则不同体征样本可互为“有其他眼底疾病的”样本)。根据测试集在模型构建总数据集中的占比,推导可得模型构建环节总数据集中符合本指南第2部分第2节“黄斑区域病变体征”定义各类病变体征的样本至少为2 000例。建议模型构建总数据集中包含的临床正常眼底图像至少7 000张,包含有其他眼底疾病的眼底图像至少1 000张。在保证每类阳性样本数量至少2 000的基础上,收集到与之满足流行病学分布的正常眼底样本和其他眼疾样本数量十分困难,因此本指南对正常眼底样本和其他眼疾样本数量进行了建议。在此前提下,本指南也建议收集的数据集中各类型样本量尽可能地满足流行病学分布。请注意,同一张图像上可出现多种病变体征,即一张图像上可存在多个标签。

此外,本指南建议收集的数据集中每类样本的样本数量按照采集机构和机型均衡分配;每类样本中样本性别和所属年龄段分布建议参考病变体征对应的黄斑区域疾病在不同性别和不同年龄段的发病率。

4、模型效果检验数据集来源及数量

为了验证AI模型的效果,除了在模型构建时使用测试集进行验证外,在模型构建完成后,还可收集额外的数据集对效果进行检验。此环节收集的数据要求与模型构建数据集一致,需满足采集环境、采集人员操作多样性,满足采集设备多样性,满足民族、地区、年龄多样性等。本数据仍可来源于公开数据集,但必须符合本指南第4部分第3节介绍的标注办法,不符合的需重新标注。所有数据的获取须经相关单位的伦理审查委员会批准。为了满足对筛查系统泛化能力的验证,此数据集涉及的采集环境、采集人员、采集设备应尽可能不同于上述模型构建数据集。如图9所示,模型效果检验数据集可由多个构成,这些数据集可用于多个厂商或第三方机构对筛查系统的验证。

模型效果检验数据集数量通常由厂商和第三方机构根据检测要求制定。此外,也可参考本指南第4部分第2节第3点“模型构建数据集数量”中模型构建过程中测试集数量,准备的数据集中涉及本指南第2部分第2节“黄斑区域病变体征”定义的每类病变体征的样本数量至少为200张,包含的临床正常眼底图像至少700张,包含有其他眼底疾病的眼底图像至少100张。同时,为了满足对不同质量图像筛查效果的验证,检验数据集中需要有图像质量合格和可接受两种情况的测试样本,其中图像质量的判断标准请参考本指南“数据收集与质量控制”中数据质量控制标准。按照评估指标的可信区间样本量计算公式(1),检验数据集中每种质量的图像应至少有200张。每个检验图像上的图像质量标签和包含病变体征标签同时存在。

5、临床试验数据集来源及数量

设计临床试验的目的是验证AI模型的性能,本指南建议临床试验采用前瞻性、配对、多中心、预期目标法进行试验设计。设计的主要依据为《医疗器械临床试验质量管理规范》^[41]、《人工智能医疗器械注册审查指导原则》^[39]和《深度学习辅助决策医疗器械软件审批要点》^[42],具体包括:(1)多中心:医疗器械临床试验应在3个及以上临床试验机构中进行,即多中心临床试验。各中心与模型构建数据集来源非同源,且每个中心包含每类病变体征的眼底图像数量建议均衡。(2)配对:试验中同一受试者的眼底彩照将同时采用AI模型和对照方法(中心阅片)来进行阅片,即配对设计。(3)预期目标法:试验所收集数据量需满足可

验证AI模型在临床场景应用的灵敏度和特异性是否达到预期目标的条件。(4) 用于检验AI模型临床效果的临床试验数据集需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,应当保证样本分布符合临床实际情况,且与模型构建过程使用的数据集无交集。

在临床试验中,为验证AI模型是否达到预期目标,需要的样本数量可与本指南“数据集构建”中计算的测试集数据量一致,即本指南建议临床试验提供的数据集中涉及第2部分第2节“黄斑区域病变体征”定义的每种病变体征图像至少为200张,包含的临床正常眼底图像至少700张,包含有其他眼底疾病的眼底图像至少100张。与第4部分第2节第4点“模型效果检验数据集来源及数量”中介绍的模型验证数据集模型验证数据集类似,为了满足对不同质量图像筛查效果的检测,临床试验数据集中需要有图像质量合格和可接受两种情况的测试样本,按照评估指标的可信区间样本量计算公式(1),临床试验数据集中每种质量的图像应至少有200张。

值得注意的是,本指南建议用来进行模型测试和检验的数据集符合预期应用场景的数据分布,根据黄斑区域病变相关的流行病学调查^[43-45],本指南第2部分第2节“黄斑区域病变体征”中涉及的病变体征中患病率最低的为色素增生(0.10%),患病率最高的为玻璃膜疣和地图样萎缩(7.21%)。若用来进行模型测试或检验的数据集中色素增生体征所属图像数为200张,则玻璃膜疣和地图样萎缩两种体征对应的图像数需达到14 420张,无黄斑区域病变相关体征的图像数至少需达到16万张,在实践中难以实现。因此,本指南建议在保证用来进行模型测试或检验的数据集中涉及第2部分第2节“黄斑区域病变体征”定义的每种病变体征图像至少为200张,包含的临床正常眼底图像至少700张,包含有其他眼底疾病的眼底图像至少100张的前提下,尽可能地增加患病率大的体征和无病变体征等类别对应的样本数。

(三) 数据集标注

上述数据集的标注和临床试验数据集的中心阅片,均可采用本节描述的标注方式,标注后的结果作为各个任务的“参考标准”。

1、标注模式

按照《眼底彩照标注与质量控制规范》^[14],本指南建议标注采用3+2+1或3+2+2的方式,即3名标注医生+2名高级医生+1或2名仲裁专家。标注医生为有执业医师资格证或是经过相关专业培训合格的医学专业人员;高级医生为在眼底病领域工作5年及以上的专业人员;仲裁专家为职称为副主任医师及以上的更高级别

专家。若三名标注医生中存在两位标注结果不完全一致,则引入两名高级医生。若两位高级医生标注一致,则以两位高级医生的标注结果为准;若不一致,则再引入一至两名仲裁专家。仲裁专家评审后,提供最终结论。

2、标注人员

标注人员应接受数据标注培训,培训内容应包含标注软件操作流程,标注规程等。人员培训及考核具体内容如下:(1) 培训内容:1) 为了规范标注过程,需由标注系统研发工程师对标注医生、评估医生和仲裁专家就如何使用标注系统进行培训。2) 为了统一对眼底图像标注规则的认识,需由眼底病专家(三甲医院眼底病专业主任医师)依据病变体征的定义对标注医生进行培训。(2) 考核机制:1) 采用实操的方法进行考核。从每个病变体征随机选取一定张数的图像作为考核图像,由眼底病专家和参加考核的标注医生同时标注所有选取图像指定形式的标签(详见第4部分第3节“数据集标注”),计算标注医生与眼底病专家标注结果的一致率。2) 考核指标:标注一致率要求不低于一定比例。根据眼底病专家评估,本指南建议考核合格的标注医生和眼底病专家标注一致的样本数占比不低于80%。

3、标注过程

本指南涉及的标注对象为彩色眼底图像,标注形式需按照筛查系统功能进行设定(表2)。筛查系统可同时包含多项功能,因此标注形式可有多种。

标注过程包含以下三个重要内容:图像中治疗痕迹的判定、图像可读性的判定以及图像具体标注流程(图10)。医生标注前须由标注系统管理员分配用户名和密码,随后标注医生登录进入标注系统,开始标注工作。(1) 治疗痕迹判定:如果发现治疗痕迹,如激光斑、硅油填充、气体填充、外加压/环扎等,则该图像不需要标注,直接剔除;(2) 图像可读性判定:图像拍摄质量会影响可读性,若图像因欠曝、过曝或其他拍摄原因导致黄斑区域1/3以上不可读,则建议该图像不标注,直接剔除。

4、标注质量评估

在标注过程中,本指南建议对标注质量进行评估。评估可从正确性和重复性两个角度进行。评估过程与标注人员考核过程相似,从已标注图像中随机抽取一批图像(建议包含200个样本)作为评估图像集,请眼底病专家对评估图像集进行标注。正确性评估:对比标注医生和眼底病专家对评估图像集的标注结果,若标注结果一致率达到预设比例,则认为标注医生标注正确性合格。重复性评估:将评估图像集重复

表 2 筛查系统功能（拟提供的信息）与标注形式对应表

筛查系统拟提供的信息	标注形式
转诊建议	标出无需转诊、建议转诊、确定转诊中的一个标签
病变体征名称	标出病变体征名称（0个，1个或多个）
病变体征区域	标出病变体征的区域框（0个，1个或多个）
病变体征坐标	标出病变体征区域的质心坐标（0个，1个或多个）
病变体征轮廓	标出病变体征的轮廓线（0个，1个或多个）
相关疾病	标出黄斑区域疾病名称（0个，1个或多个）
疾病严重程度	标出疾病严重程度，如分级/分期（0个，1个或多个）
图像质量	给出合格、可接受、不合格中的一个标签

投入标注数据池，由标注医生进行二次标注，随后对比评估图像集中各图像两次标注结果，若标注结果一致率达到预设比例，则认为标注医生标注重复性合格。标注质量不合格的医生将接受二次培训，标注不合格的图像将抹去已有标签后放回标注数据池重新进行标注。

（四） 算法模型构建

本指南中涉及的第2部分第2节“黄斑区域病变体征”筛查AI模型负责提供图像中有或无黄斑区域病变体征存在的参考意见、具体病变体征的分割或定位可视化展示等。

在模型构建过程中，训练集图像和相应标注会输入模型，供其学习图像与目标任务的映射关系、更新模型参数^[37]。在训练过程中，算法可对图像进行初步的预处理，如图像前景提取、图像缩放到固定尺寸、图像色调增强等，其目的是通过对输入图像的尺寸和色调标准化来减小机型之间的样本差异。同时算法可以对原始图像进行随机左右翻转、旋转、平移和剪裁，调节图像色彩和明暗程度等图像增强操作，目的是增强训练样本的多样性以及数量。此外，算法还可以在训练过程中对图像引入随机噪声，增强模型对于噪声和对抗样本攻击的健壮性^[46]。训练结束后，在调优集上对模型进行优化。调优结束后，模型可以在测试集上进行测试及性能评估。

（五） 算法模型评估指标

本指南第3部分“AI算法在黄斑区域病变体征筛查汇总的应用场景”中介绍了AI算法在黄斑区域病变体征筛查中的应用场景有病变体征检测、病变体征定位、病变体征分割以及扩展的疾病分级和疾病诊断。本指南将按照不同的应用场景介绍算法模型的评估指标。

1、病变体征检测

病变体征检测属于多标签分类任务，该任务的模型评估指标主要有灵敏度、特异性、漏检率、精确

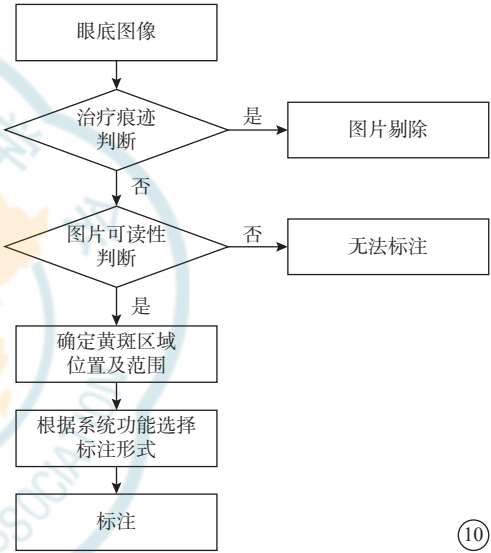


图 10 黄斑区域病变体征筛查系统中AI任务的标注流程。AI：人工智能

率、分数、AUC分数及Kappa值等，具体评估指标定义及计算公式如下所示^[11-12]。灵敏度（sensitive，Sen）：又可称召回率（recall）、查全率，是真阳性样本占全体阳性样本的比例。

$$Sen = \frac{TP}{TP + FN}$$

(3)

特异性（special，Spe）：真阴性病例占全体阴性病例的比例。

$$Spe = \frac{TN}{TN + FP}$$

(4)

漏检率（miss rate，MR）：即检测中未发现的阳性样本占全体阳性样本的比例。

$$MR = 1 - \frac{TP}{TP + FN}$$

(5)

精确率（precision，Pre）：又称查准率、阳性预测值，是真阳性样本占算法判为阳性样本的比例。

$$Pre = \frac{TP}{TP + FP}$$

(6)

阴性预测值 (negative prediction value, NPV) : 真阴性样本占被算法判为阴性样本的比例。

$$NPV = \frac{TN}{TN + FN} \quad (7)$$

准确率 (accuracy, Acc) : 算法诊断正确的样本占全体样本的比例。

$$Acc = \frac{TP + TN}{N} \quad (8)$$

F₁分数: 召回率和精确率的调和平均数。

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (9)$$

约登指数: 也称正确指数, 假设假阴性 (漏诊率) 和假阳性 (误诊率) 危害同等意义, 约登指数为灵敏度与特异度之和减去1, 指数越大说明筛查效果越好。

$$Jorden\ index = Sen + Spe - 1 \quad (10)$$

受试者操作特征 (receiver operating characteristic) 曲线下面积 (area under curve)。受试者操作特征是通过在一组预设阈值下计算筛查系统在测试集上的灵敏度和特异性, 从而产生一组 (1-特异度, 灵敏度) 操作点, 将这些操作点依次连接形成的曲线。

Kappa系数: 用于评价筛查系统与参考标注诊断结果一致性的指标。

$$Kappa = \frac{N(TP + TN) - (R_1C_1 + R_2C_2)}{N^2 - (R_1C_1 + R_2C_2)} \quad (11)$$

公式 (11) 中R₁是真阳性和假阳性例数的总和, R₂是假阴性和真阴性例数的总和, C₁是真阳性和假阴性例数的总和, C₂是假阳性和真阴性例数的总和, N是样本例数总和。

2、病变体征定位

定位任务的结果可以有两种表现形式: 一是输出目标中心点的坐标信息, 二是输出目标所在方框的信息。对于第一种只输出一个点坐标的模型, 评估指标通常选用平均欧式距离; 对于第二种输出一个矩形框的模型, 处理的是目标探测任务, 评估指标通常为交并比和平均的精确率均值。

平均欧式距离 (mean euclidean distance, MED) :

$$MED = \frac{1}{N} \sum_{i=1}^N \left(\sqrt{(x_i^1 - y_i^1)^2 + (x_i^2 - y_i^2)^2} \right) \quad (12)$$

交并比 (intersection over Union, IoU), 衡量预测框和医生标注的目标框参考标准的重叠程度:

$$IoU(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} = \frac{TP}{TP + FN + FP} \quad (13)$$

平均精确度均值 (mean average precision, mAP) : 在多目标定位任务中用于计算各类目标定位

的平均精确度的平均值^[47], 计算过程如下: 在目标定位任务中, 模型会提供目标框内对应某类别的概率值, 这个概率值被称为置信度。在置信度高于设定阈值时, 保留当前预测结果。计算预测框与目标框参考标准的IoU, 在IoU结果大于设定阈值且预测框内预测类别与目标框参考标准内类别一致时, 认为预测结果为真阳性。其余被认为是假阳性结果。通过由大至小改变置信度阈值, 计算模型输出结果的精确率和召回率, 可以画出Precision-Recall曲线, 曲线下面积记为精确率均值 (average precision, AP)^[48]。在得到每个类别的目标检测AP后, 可求目标检测任务的mAP测定:

$$mAP = \frac{\sum_{i=1}^K AP_i}{K} \quad (14)$$

公式 (14) 中K为目标检测任务中包含类别总数, AP_i为第i个类别目标预测结果的AP测定值。

3、病变体征分割

分割任务的模型评估指标主要有DICE系数 (dice coefficient)、Jaccard系数 (交并比)、灵敏度、特异性等。其中, 由于分割任务可视为将图像像素点分为前景与背景的二分类任务, 因此灵敏度和特异性两个指标可用于评估分割模型对每个像素点判别的效果。

DICE系数, 分割轮廓与参考标准轮廓的交集占分割轮廓与参考标准轮廓平均值的比例:

$$DICE(X, Y) = \frac{|X \cap Y|}{\frac{1}{2}(|X| + |Y|)} = \frac{2 \times TP}{(TP + FN) + (TP + FP)} \quad (15)$$

其中|X∩Y|是X和Y之间的交集, |X|和|Y|分表表示X和Y的元素个数。

Jaccard系数 (jaccard coefficient) : 分割轮廓与参考标准轮廓的交集占分割轮廓与目标轮廓并集的比例, 又称交并比 (IoU)。

$$Jaccard(X, Y) = IoU(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} = \frac{TP}{TP + FN + FP} \quad (16)$$

4、其他应用模型评估

本指南第3部分第4节“AI算法的其他应用”介绍了两类在疾病维度分析的AI模型, 分别用于疾病的分类 (诊断) 和分级。由于这两类应用均属于分类研究, 因此评估指标可参考本指南第4部分第5节第1条。

5、模型综合评估指标

(1) 筛查符合率: 模型提供的阳性病例结果占中心阅片判定阳性病例结果的百分比。(2) 重复性需求: 同一操作医生/技师在满足配置环境要求的不同服

务器上部署运行AI辅助筛查系统,对于同一张图像,应能再现其预测结果。(3)再现性需求:不同操作医生/技师在同一配置环境下于不同时间段分别独立部署运行AI辅助筛查系统,对于同一张图像,应能再现其预测结果。(4)健壮性需求:AI筛查系统对于多中心、多机型临床数据应有稳定的指标表现,主要指标的统计性能变化建议控制在一定范围内,具体范围数值需视筛查系统功能和相应医疗器械标准、法规而定。

五 AI黄斑区域病变体征筛查系统的应用标准

(一) AI辅助筛查系统的类型

离线版AI系统可安装在计算机或移动电子设备上,无需网络即可对输入的眼底彩照进行黄斑区域病变体征的智能筛查辅助诊断,并生成筛查报告。离线版AI系统反馈速度快,参考《基于眼底照相的糖尿病视网膜病变人工智能筛查系统应用指南》^[10]建议,离线版AI系统输入彩色眼底照片至输出筛查结果时间应控制在1 min以内。本指南建议应用机构保留筛查系统的离线版本。

在线版AI系统适用于有一定网络使用条件的单位,需要将脱敏的眼底彩照通过网络传输上传到云端,随后在云端进行黄斑区域病变体征的智能筛查,并传回结果。之后结合受检者信息,生成报告供使用者下载、确认。在线版系统对硬件设备、网络传输速度和传输安全要求高。参考《基于眼底照相的糖尿病视网膜病变人工智能筛查系统应用指南》^[10],在线版AI系统输入彩色眼底照片至输出筛查结果时间应控制在5 min内。

(二) 筛查系统的架构和功能

黄斑区域病变体征辅助筛查系统应当具备的架构包括受检者管理、采集图像显示、采集图像质量评估、AI筛查、专科医生拟诊意见输入和报告生成。筛查系统的功能视实际应用需求而定,可包括判别病变体征是否存在、病变等级、提供病变体征所处区域、具体轮廓等功能。系统的部署应当适应相应机构的环境和网络情况。

1、受检者管理

在系统中可进行受检者信息的录入、修改和查看。受检者信息包括个人基本信息[如姓名、年龄、性别、医疗身份标识(ID)等]和相关检查信息(如视力、眼压、血压等),可为医生提供详细的患者信息,以此核实采集图像与患者匹配无误。

2、图像显示

展示采集的右眼、左眼的眼底图像。该图不可编辑,但可被缩放、平移,方便医生分析。

3、图像质量评估

本指南建议采用的质量评分方案应覆盖拍摄位置、亮度、清晰度、以及信息是否缺失四部分(详见第4部分第1节“数据收集与质量控制”),以判断图像质量是否能够用来观察或检测黄斑区域病变体征。筛查系统可提供当前采集图像的质量评估结果(合格、可接受、不合格)。

4、AI筛查

筛查系统最主要的模块即为AI筛查,系统开发者需根据设计的筛查功能,训练相应的AI模型,使其完成AI筛查环节。AI模型输出的结果将展示在筛查系统提供的最终报告中,且需注明“该结果来自AI筛查系统”。报告中展示的筛查结果形式取决于系统预设的筛查功能,即预测受检者是否属于转诊病例、判断图像中具体出现了何种病变体征、提供病变体征所处区域、病变体征轮廓信息等。

5、专科医生拟诊意见

专科医生复核AI筛查结果,并提供拟诊意见。参考国内外现有的AI辅助筛查系统应用指南^[9],拟诊意见建议在24 h内提供。本指南建议设计的筛查系统中预留专科医生拟诊意见输入模块。

6、报告生成

筛查报告以电子版或纸质版形式提供给受检者,内容应覆盖本指南的筛查范围,且符合健康体检主检报告撰写标准和规范^[49]。

7、筛查系统部署方式

离线版AI黄斑区域病变体征筛查系统部署在本地设备中,在线版AI黄斑区域病变体征筛查系统使用云端远程部署。

(三) 彩色眼底图像采集硬件要求

黄斑区域病变体征筛查系统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,各项指标参数应参考如下要求:(1)对焦方式:机器可支持自动对焦。(2)曝光方式:机器支持自动曝光。(3)视场角度:使用单视野时水平方向不小于45°,垂直方向不小于45°^[35];或使用双视野时每个视野水平方向不小于45°,垂直方向不小于45°。(4)可拍摄的最小瞳孔直径:标准模式瞳孔最小约4 mm,小瞳孔模式瞳孔最小约3.3 mm。(5)分辨率:视野范围内最小分辨率不小于眼底照相机行业标准要求^[50]。(6)屈光补偿范围:不小于±15 D。(7)内固视点:眼底照相机的内置固视点至少可调节至3个位置,分别对应以视盘中心为视野中心、以黄斑中心为视野中心、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。(8)图像存储格式: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。(9)图像拍摄像素:不低于1800万像素。

(四) 图像采集视野标准

1、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^[35](图11A)。

2、双视野拍摄法:视野1以黄斑中心小凹为拍摄视野的中心,成像至少涵盖45°视网膜区域;视野2以视盘为拍摄视野的中心,成像至少涵盖45°视网膜区域,且能涵盖视盘周边2DD的区域,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(图11B, 11C)。

拍摄的图像质量需按照本指南第4部分第1节“数据收集与质量控制”中的标准进行判别,对于不合格图像需重新拍摄。

(五) 黄斑区域病变的AI筛查方案

1、彩色眼底图像的质量评估

AI系统分析前可对图像进行质量评估,此模块将根据图像质量控制标准,从拍摄位置、清晰度和可读范围三个角度考察采集图像的质量:(1)拍摄位置:检查图像各区域是否有偏转、移位。(2)清晰度:检查对焦是否准确,视盘表面、视网膜主干血管、视网膜分支血管、黄斑、视网膜病变等结构是否清晰可辨。(3)可读范围:检查视野范围内是否完整包含视盘、黄斑区、上下血管弓区域,是否有睫毛或其他异物遮挡。

图像质量分为合格、可接受、不合格三种。若检测中采集图像质量合格,则正常进行后续AI筛查分析;若图像质量为可接受,操作者可选择继新进行AI筛查分析,但在“筛查结果”处应注明“图像质量偏差,结果仅供参考”,另外操作者也可选择重新拍摄受检者眼底图像;若图像质量不合格,操作者必须重新采集受检者眼底图像,若3次拍摄均得到质量不达标的图像,则拍摄失败,无法进行后续AI筛查分析。

2、筛查方案指导标准和原则

(1) 黄斑区域病变体征分割/定位指导标准

针对本指南第2部分第2节“黄斑区域病变体征”中列出的黄斑病变体征(玻璃膜疣、地图样萎缩、新生血管、渗出、出血、瘢痕、色素增生等),建议AI筛查系统提供体征分割结果的可视化展示。为了方便医生观察病变,本指南建议筛查系统提供的可视化方式为病变体征分割结果的边缘显示或定位结果的检测框显示。

(2) 黄斑区域病变体征检出指导标准

按照本指南第2部分第2节“黄斑区域病变体征”列出的病变体征定义,建议AI筛查系统提供输入眼底

彩照中是否有以上体征存在的判别结果。

(3) 黄斑区域病变体征筛查的指导原则

医生先参考AI结果,纠正不认可的部分,再以此向患者详细解释说明AI筛查结果的内容和意义,同时需说明AI辅助诊断不等同于专业医生诊断。医生可根据AI筛查结果,并结合受检者的视力^[51]及临床表现,为其提供下一步专业的诊疗建议:1)未见黄斑区域病变相关体征(低风险、无需转诊)的受检者,建议每年进行1次眼底健康检查。2)眼底疑似出现1项相关体征(中风险、建议转诊)的受检者,或仅通过眼底图无法明确判定属于“黄斑区域病变体征”黄斑区域病变项目中体征表现,按照临床经验,倾向于进行进一步的检查;3)眼底确定出现至少1项相关体征(高风险、确定转诊)的受检者,应尽快到医院进行详尽的检查。

(六) 筛查报告要求

(1)受检者的基本信息:病历号/医疗ID、姓名、年龄、性别、受检眼别。(2)眼科相关信息:视力、眼压、眼科疾病史、眼科治疗史(手术、用药)。(3)图像采集设备信息:采集日期、设备型号、成像范围、图像存储格式。(4)筛查系统相关信息:算法类型及版本、算法适用机型、图像分析日期。(5)图像评估:图像质量。(6)AI筛查结果:受检眼的诊断、患病或健康的概率、黄斑区域病变体征性眼底改变方面的描述。具体地,建议包括本指南第2部分“黄斑区域病变体征”中列出的病变体征是否存在结果,病变体征分割/定位显示。(7)专科医生拟诊意见。(8)权责说明。

(七) 数据存储

眼底彩照的存储需要满足数据的存储安全、管理的高效和方便医院信息系统(HIS)^[52]、图像存档通信系统(PACS)^[53]、AI辅助筛查系统在使用上的便捷。具体要求如下:(1)模型构建数据集可由开发机构存储,模型验证数据集可由第三方检测机构存储,临床试验数据集可由各试验中心存储;模型接受第三方检测机构和临床试验中心测试时,建议采用盲测方式,即数据集不流入模型开发机构。(2)PNG格式、JPG格式、TIFF格式或DICOM格式存储。(3)AI模型预测结果需保存,并能关联到受检者、检查信息和眼底彩照。(4)本地和云端部署的AI筛查系统都需要对数据进行备份,在存储容量达到一定的阈值时需要告警,系统管理员要进行存储更新。(5)AI筛查系统和医院PACS、HIS对接需要满足医疗相关协议规范。

(八) 数据安全

参照我国《中华人民共和国数据安全法》、《个

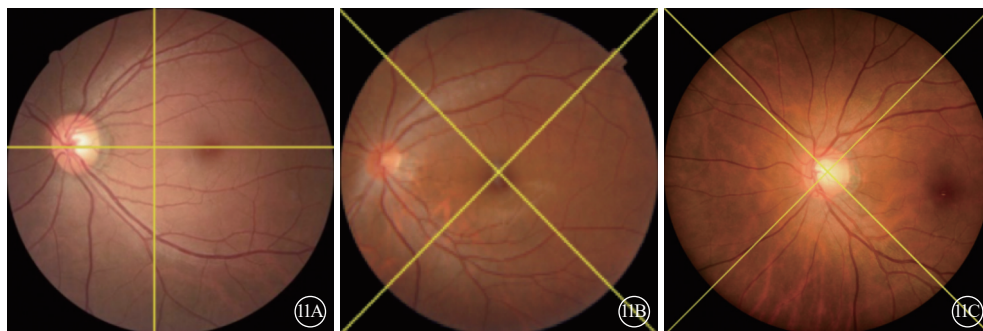


图 11 单视野及双视野拍摄法示意图。11A 示单视野拍摄法；11B 示双视野拍摄法，以黄斑中心小凹为拍摄中心；11C 示双视野拍摄法，以视盘为拍摄中心

人信息保护法》、《网络安全法》、《人口健康信息管理办法》，应做到：

1、数据脱敏：对于不影响疾病诊断的受试者敏感信息，如姓名、联系电话、身份证号、家庭住址等，应进行脱敏处理。数据脱敏时需注意在执行数据漂白，抹去敏感内容的同时保持原有的数据特征，保证开发、测试等不受到脱敏的影响，达成脱敏前后的数据一致性和有效性。

2、数据使用用户管理：包括用户身份管理、认证管理和授权管理。为每个用户创建唯一的用户账号，并对用户身份进行鉴别，确保数据访问控制和安全审计可以追溯到个人账号。数据访问需有统一的身份认证机制，对于敏感数据，可采用多步验证技术，防止误操作。需要根据数据访问主体身份，以及被访问数据的密级，实现对各类数据的访问授权。

3、数据使用日志管理：捕获系统内完整活动的记录且不可更改，确保每一项用户破坏性操作被记录审计，保证操作可回溯，进而可辅助快速定位系统是否遭受恶意的操作和攻击。

4、传输加密：需通过安全传输方式和标准加密协议，满足平台以及传输敏感数据的需求，避免数据被非法访问、窃听或旁路嗅探。设置敏感数据流转监控、数据流转关联分析等环节，保证流转操作安全进行。

5、存储安全：数据存储过程需使用加密措施进行防护，降低数据泄露风险。加密算法需使用安全强度符合国家安全要求的商用密码算法。存储阶段除了保障数据访问权限的控制，还需考虑备份和容灾问题，可通过本地存储、网络存储等多种方式实现数据多副本多数据中心的同步和备份，实现数据的异地容灾，从容应对极端灾难。

6、明确责任人：AI 黄斑区域病变体征筛查系统落地应用的机构需明确数据安全责任人，并制定数据安全的管理制度和应急机制。

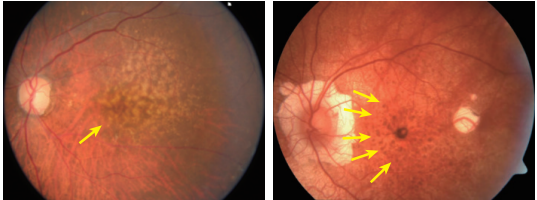
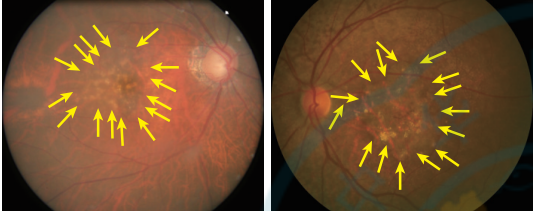
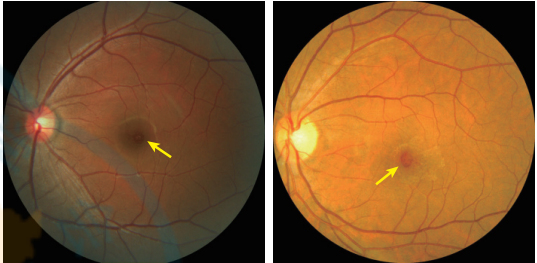
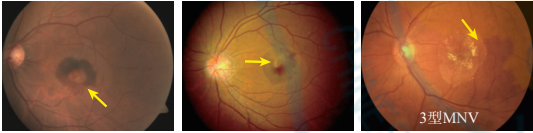
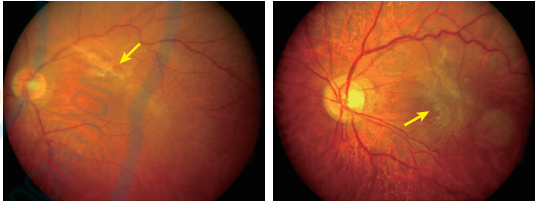
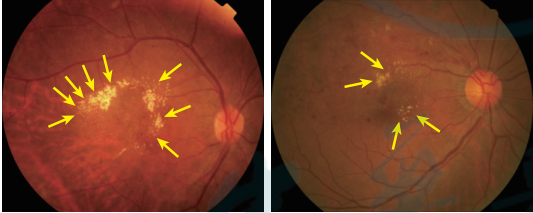
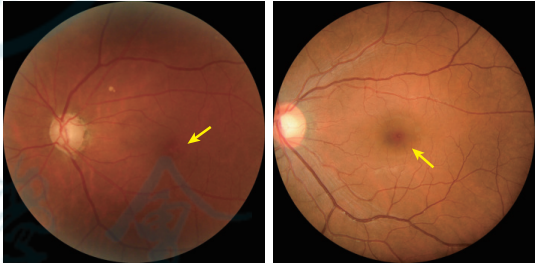
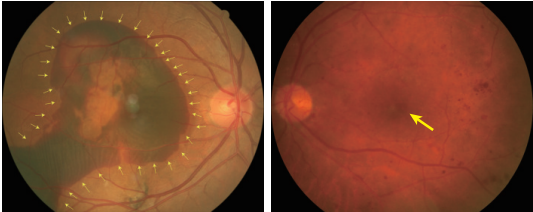
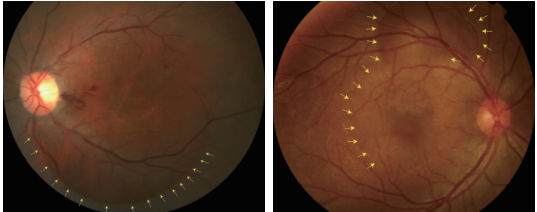
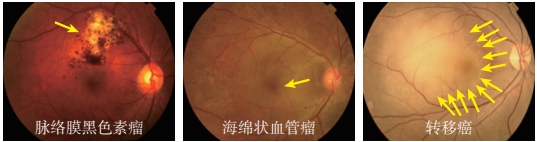
六 AI 黄斑区域病变体征筛查系统的局限与未来发展方向

目前 AI 系统对不同机型和不同拍摄质量图像的筛查效果稳定性有待提高，因此，现存的 AI 筛查系统需对拍摄设备和拍摄方式进行严格的约束。值得期待的是，AI 领域出现了域自适应学习（domain adaptation learning，AI 技术，一种迁移学习方式，能够有效地解决训练样本和测试样本概率分布不一致的学习问题，是当前机器学习的热点研究领域，在自然语言处理、文本分析、影像分析、生物信息学、跨语言分析、视频分析、情感分析和手写体识别等领域有广泛应用）相关研究^[54-55]，该研究可缓解 AI 模型在不同机型采集的不同质量图像上筛查效果健壮性差的问题。此外，由于 AI 中的深度学习技术通常是个“黑盒”，因此基于此项技术实现疾病判别的 AI 系统存在难以解释的问题。本指南以病变体征为检测目标，在第 3 部分“AI 算法在黄斑区域病变体征筛查中的应用场景”中提出了病变体征检测、定位、分割等任务。其中，检测出病变体征位置可作为受试者复查时的关注区域提示，提示医师注意一段时间内该区域内病变体征增多或减少。因此，建议在医疗场景应用 AI 系统时要覆盖疾病诊断过程中的若干子任务。

本指南针对的是输入眼底彩照单一模态数据的病变体征筛查场景，由于黄斑区域疾病的复杂性，单一模态数据无法提供全面的眼结构和病变信息，因此未来可发展多种眼底图像模态输入的 AI 辅助黄斑区域疾病的筛查系统。值得注意的是，这种辅助筛查系统迫切需要眼底检查硬件的发展，如出现方便拍摄、价格低的光相断层扫描采集设备等。目前已达成共识的是 AI 系统在医疗中的应用需要有专业医生的复核，因此，AI 可视为辅助医生的工具，AI 筛查不等同于专业医生诊断。目前 AI 筛查系统在某些疾病/体征的早期筛查中有很好的应用，可缓解筛查人口基数大、专业医生人数不足等问题。相信在未来，会有更多基于智能技术的疾病/体征筛查系统出现。

七 附录:黄斑区域病变体征标注参考示例 (表3)

表 3 黄斑区域病变体征标注参考示例

标注描述	示例图像	标注描述	示例图像
玻璃膜疣		色素增生	
地图样萎缩		黄斑裂孔	
脉络膜新生血管		黄斑前膜	
渗出		黄斑水肿	
出血		视网膜脱离	
瘢痕		眼内占位性病变	

形成指南意见的专家组成员

医学意见指导专家组

中华医学会眼科学分会眼底病学组

张 明 四川大学华西医院 (本文主要负责人, 执笔)

许 迅 上海交通大学附属第一人民医院 (组长)

王雨生 第四军医大学西京医院 (副组长)

魏文斌 首都医科大学附属北京同仁医院 (副组长)

赵明威 北京大学人民医院 (副组长)

陈有信 北京协和医院

李筱荣 天津医科大学眼科中心

陆 方 四川大学华西医院

刘庆淮 江苏省人民医院

雷 博 河南省眼科研究所

曲进锋 北京大学人民医院

文 峰 中山大学中山眼科中心

俞素勤 上海交通大学附属第一人民医院

杨卫华 南京医科大学附属眼科医院

张少冲 深圳市眼科医院

杜 亮 四川大学华西医院 中国循证医学中心

王冬玥 四川大学华西医院 (非学组委员, 秘书)

人工智能技术意见指导专家组

许言午 北京百度网讯科技有限公司 (执笔)

刘 江 南方科技大学计算机科学与工程系

徐向民 华南理工大学未来技术学院

陈 杰 鹏城实验室

谭明奎 华南理工大学软件学院

夏 勇 西北工业大学计算机学院

雷柏英 深圳大学生物医学工程学院

赵一天 中国科学院宁波材料所慈溪生物医学工程研究所

朱鹏志 广东省医疗器械质量监督检验所

方慧卉 北京康夫子健康技术有限公司 (秘书)

杨大陆 北京百度网讯科技有限公司

孙钦佩 北京康夫子健康技术有限公司

张思琦 北京百度网讯科技有限公司

指南外审专家组成员

医学专家组

姚克 浙江大学医学院附属第二医院

黎晓新 厦门大学附属厦门眼科中心

王宁利 北京同仁医院

孙兴怀 复旦大学附属眼耳鼻喉科医院

惠延年 空军军医大学西京医院

陈长征 武汉大学人民医院

李世迎 厦门大学附属翔安医院

刘晓玲 温州医科大学附属眼视光医院

卢 海 北京同仁医院

吕 林 中山大学中山眼科中心

彭晓燕 北京同仁医院

沈 吟 武汉大学人民医院

孙晓东 上海交通大学附属第一人民医院

王 敏 复旦大学附属眼耳鼻喉科医院

徐格致 复旦大学附属眼耳鼻喉科医院

严 宏 西安市第四医院

颜 华 天津医科大学总医院

于伟泓 北京协和医院

张 含 中国医科大学附属第一医院

赵培泉 上海交通大学医学院附属新华医院

人工智能专家组

闵 栋 中国信息通信研究院云计算与大数据研究所

王 浩 中国食品药品检定研究院

何晖光 中国科学院自动化研究所

段立新 电子科技大学计算机科学与工程学院

吴庆耀 华南理工大学软件学院

声明 本指南制定严格按照世界卫生组织、中华医学会的指南制定规范, 并参考指南研究与评价工具和国际实践指南报告标准, 旨在为基层眼底筛查服务提供指导, 不是在各种情况下都必须遵循的医疗标准, 也不是为特殊个体提供的保健措施; 本指南所包括的信息或资料的准确性、完整性、合法性、可靠性、可操作性不承担任何法律责任; 本指南制定过程中严格遵循世界卫生组织指南制定利益冲突条例和指南伦理道德标准, 全体参与成员均填写利益冲突声明表, 经过评价与指南制定不存在直接利益冲突

利益冲突 所有作者均声明不存在利益冲突

8 参考文献

- [1] 陈浩宇. 糖尿病视网膜病变筛查手册诊断和处理(第2版)[M]. 北京: 人民卫生出版社, 2016: 10-12.
Chen HY. Handbook of retinal screening in diabetes diagnosis & management (2 nd ed)[M]. Beijing: People's Medical Publishing House, 2016: 10-12.
- [2] 沙文茹. 《中国眼健康白皮书》发布: 致盲性眼病有效遏制[J]. 中国医药科学, 2020, 10(13): 3. DOI: [10.3969/j.issn.2095-0616.2020.13.002](https://doi.org/10.3969/j.issn.2095-0616.2020.13.002).
Sha WR. "China Eye Health White Paper" released: effectively inhibition of blinding eye diseases[J]. China Medicine and Pharmacy, 2020, 10(13): 3. DOI: [10.3969/j.issn.2095-0616.2020.13.002](https://doi.org/10.3969/j.issn.2095-0616.2020.13.002).
- [3] Balyen L, Peto T. Promising artificial intelligence-machine learning-deep learning algorithms in ophthalmology[J]. Asia Pac J Ophthalmol (Phila), 2019, 8(3): 264-272. DOI: [10.22608/APO.2018479](https://doi.org/10.22608/APO.2018479).
- [4] Orlando JJ, Fu H, Breda JB, et al. Refuge challenge: a unified framework for evaluating automated methods for glaucoma assessment from fundus photographs[J/OL]. Med Image Anal, 2020,

- 59: 101570[2019-10-08]. <https://pubmed.ncbi.nlm.nih.gov/31630011/>. DOI: 10.1016/j.media.2019.101570.
- [5] 吴乐正, 林顺潮. 临床眼黄斑病学[M]. 北京: 北京科学技术出版社, 2007: 29-104.
- Wu LZ, Lin SC. Clinical maculopathy[M]. Beijing: Beijing Science and Technology Publishing, 2007: 29-104.
- [6] Esteva A, Kuprel B, Novoa RA, et al. Correction: corrigendum: dermatologist-level classification of skin cancer with deep neural networks[J]. *Nature*, 2017, 546(7660): 686. DOI: 10.1038/nature22985.
- [7] 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.
- [8] Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning[J]. *Ophthalmology*, 2017, 124(7): 962-969. DOI: 10.1016/j.ophtha.2017.02.008.
- [9] 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.
- Chinese Medical Association Ophthalmology Branch Glaucoma Group, China Medical Equipment Association Ophthalmology Artificial Intelligence Group. Guidelines for standardized design and application of artificial intelligence glaucoma auxiliary screening system based on fundus photography in China (2020)[J]. *Chin J Ophthalmol*, 2020, 56(6): 423-432. DOI: 10.3760/cma.j.cn112142-20200102-00003.
- [10] 袁进, 雷博, 张明, 等. 基于眼底照相的糖尿病视网膜病变人工智能筛查系统应用指南[J]. *中华实验眼科杂志*, 2019, 37(8): 593-598. DOI: 10.3760/cma.j.issn.2095-0160.2019.08.001.
- Yuan J, Lei B, Zhang M, et al. Guidelines for artificial intelligent diabetic retinopathy screening system based on fundus photography[J]. *Chin J Exp Ophthalmol*, 2019, 37(8): 593-598. DOI: 10.3760/cma.j.issn.2095-0160.2019.08.001.
- [11] 中国食品药品检定研究院. 人工智能医疗器械质量要求和评价第1部分: 术语. YY/T 1833.1-2022[S/OL]. (2020-07-03)[2022-06-16]. <http://www.cnpharm.com/upload/resources/file/2020/07/03/54879.pdf>.
- National Institutes for Food and Drug Control. Artificial intelligence medical device—Quality requirements and evaluation—Part 1: Terminology. YY/T 1833.1-2022[S/OL]. (2020-07-03)[2022-06-16]. <http://www.cnpharm.com/upload/resources/file/2020/07/03/54879.pdf>.
- [12] 中国食品药品检定研究院. 人工智能医疗器械质量要求和评价第2部分: 数据集通用要求. YY/T 1833.2-2022[S/OL]. (2020-07-03)[2022-06-16]. <https://www.nifdc.org.cn/directory/web/nifdc/infoAttach/38b8b027-8b43-43c1-b744-c0ac891b5ec8.pdf>.
- National Institutes for Food and Drug Control. Artificial intelligence medical device—Quality requirements and evaluation—Part 2: General requirements for datasets. YY/T 1833.2-2022[S/OL]. (2020-07-03)[2022-06-16]. <https://www.nifdc.org.cn/directory/web/nifdc/infoAttach/38b8b027-8b43-43c1-b744-c0ac891b5ec8.pdf>.
- [13] 中国食品药品检定研究院. 人工智能医疗器械质量要求和评价第3部分: 数据标注通用要求(征求意见稿)[S/OL](2021-06-04)[2022-06-16]. <https://www.nifdc.org.cn/directory/web/nifdc/images/2021080213295600571.pdf>.
- National Institutes for Food and Drug Control. Artificial intelligence medical device—Quality requirements and evaluation—Part 2: General requirements for data annotation (draft for comments)[S/OL]. (2021-06-04)[2022-06-16]. <https://www.nifdc.org.cn/directory/web/nifdc/images/2021080213295600571.pdf>.
- [14] 中国质量检验协会. 眼底彩照标注与质量控制规范(T/CAQI 166-2020)[J]. *中华实验眼科杂志*, 2021, 39(9): 761-768. DOI: 10.3760/cma.j.cn115989-20210106-00013.
- China Association for Quality Inspection. Annotation and quality control specifications for fundus color photographs (T/CAQI 166-2020)[J]. *Chin J Exp Ophthalmol*, 2021, 39(9): 761-768. DOI: 10.3760/cma.j.cn115989-20210106-00013.
- [15] Davis MD, Gangnon RE, Lee LY, et al. The age-related eye disease study severity scale for age-related macular degeneration: AREDS report no. 17[J]. *Arch Ophthalmol*, 2005, 123(11): 1484-1498. DOI: 10.1001/archophth.123.11.1484.
- [16] Britton G, Liaaen Jensen S, Hanspeter P. Carotenoids Volume 5: nutrition and health[J]. *Carotenoids*, 2009, 24(10): 497-522.
- [17] 美国眼科学会, 赵家良. 眼科临床指南[M]. 北京: 人民卫生出版社, 2013: 57-59.
- American Academy of Ophthalmology, Zhao JL. Preferred practice pattern[M]. Beijing: People's Medical Publishing House, 2013: 57-59.
- [18] 张承芬. 眼底病学(第2版)[M]. 北京: 人民卫生出版社, 2010: 397-398.
- Zhang CF. Diseases of ocular fundus (2nd ed)[M]. Beijing: People's Medical Publishing House, 2010: 397-398.
- [19] 杨依柳, 杨婷婷, 陆方, 等. 新生血管性老年性黄斑变性亚型报告的国际新命名专家共识解读[J]. *中华眼底病杂志*, 2022, 38(2): 99-107. DOI: 10.3760/cma.j.cn11434-20220128-00055.
- Yang YL, Yang TT, Lu F, et al. Brief interpretation of the consensus nomenclature for reporting neovascular age-related macular degeneration data[J]. *Chin J Ocul Fundus Dis*, 2022, 38(2): 99-107. DOI: 10.3760/cma.j.cn11434-20220128-00055.
- [20] 中华医学会眼科学分会眼底病学组. 中国老年性黄斑变性临床诊断治疗路径[J]. *中华眼底病杂志*, 2013, 29(4): 343-355. DOI: 10.3760/cma.j.issn.1005-1015.2013.04.002.
- Fundus Disease Group of Chinese Medical Association Ophthalmology Branch. Clinical pathway of age-related macular degeneration in China[J]. *Chin J Ocul Fundus Dis*, 2013, 29(4): 343-355. DOI: 10.3760/cma.j.issn.1005-1015.2013.04.002.
- [21] 文峰, 张雄泽. 提高对视网膜出血的分类及临床意义的认识[J]. *眼科*, 2009, 18(4): 221-224.
- Wen F, Zhang XZ. Classification and clinical significance of retinal hemorrhage[J]. *Ophthalmol CHN*, 2009, 18(4): 221-224.
- [22] Sarks J, Tang K, Killingsworth M, et al. Development of atrophy of the retinal pigment epithelium around disciform scars[J]. *Br J Ophthalmol*, 2006, 90(4): 442-446. DOI: 10.1136/bjo.2005.083022.
- [23] Gass JD. Reappraisal of biomicroscopic classification of stages of development of a macular hole[J]. *Am J Ophthalmol*, 1995, 119(6): 752-759. DOI: 10.1016/s0002-9394(14)72781-3.
- [24] 张惠蓉, 王薇. 特发性黄斑视网膜前膜[J]. *中国实用眼科杂志*, 1997, 15(10): 4.
- Zhang HR, Wang W. Idiopathic preretinal macular membrane[J]. *Chin J Pract Ophthalmol*, 1997, 15(10): 4.
- [25] Early treatment diabetic retinopathy study design and baseline patient characteristics: ETDRS report number 7[J]. *Ophthalmology*, 1991, 98(5): 741-756. DOI: 10.1016/s0161-6420(13)38009-9.
- [26] 严密. 黄斑囊样水肿[J]. *中华眼底病杂志*, 2002, 18(3): 234-235. DOI: 10.3760/j.issn.1005-1015.2002.03.035.
- Yan M. Cystoid macular edema[J]. *Chin J Ocul Fundus Dis*, 2002, 18(3): 234-235. DOI: 10.3760/j.issn.1005-1015.2002.03.035.
- [27] Sigler EJ, Randolph JC, Kiernan DF. Longitudinal analysis of the structural pattern of pseudophakic cystoid macular edema using multimodal imaging[J]. *Graefes Arch Clin Exp Ophthalmol*, 2016, 254(1): 43-51. DOI: 10.1007/s00417-015-3000-8.
- [28] Brinton DA, Wilkinson CP. 视网膜脱离: 原理与实践(第3版)[M]. 马凯, 杨庆松, 徐军, 译. 2版. 北京: 人民卫生出版社, 2011: 65-67.
- Brinton DA, Wilkinson CP. Retinal detachment: principles and practice (3rd ed)[M]. Ma K, Yang QS, Xu J, translation. 2nd ed. Beijing: People's Medical Publishing House, 2011: 65-67.
- [29] Zayit-Soudry S, Moroz I, Loewenstein A. Retinal pigment epithelial detachment[J]. *Surv Ophthalmol*, 2007, 52(3): 227-243. DOI: 10.1016/j.survophthal.2007.02.008.
- [30] Madjarov G, Koccev D, Gjorgjevikj D, et al. An extensive experimental comparison of methods for multi-label learning[J]. *Pattern Recogn*, 2012, 45(9): 3084-3104. DOI: 10.1016/j.patcog.2012.03.004.
- [31] Landis JR, Koch GG. The measurement of observer agreement for categorical data[J]. *Biometrics*, 1977, 33(1): 159-174. DOI: 10.2307/

2529310.

- [32] Liu L, Ouyang W, Wang X, et al. Deep learning for generic object detection: a survey[J]. *Int J Comput Vision*, 2020, 128(2): 261-318. DOI: [10.1007/s11263-019-01247-4](#).
- [33] Li T, Bo W, Hu C, et al. Applications of deep learning in fundus images: a review[J/OL]. *Med Image Anal*, 2021, 69: 101971[2021-01-20]. <https://pubmed.ncbi.nlm.nih.gov/33524824/>. DOI: [10.1016/j.media.2021.101971](#).
- [34] Zhang YJ. A survey on evaluation methods for image segmentation[J]. *Pattern Recogn*, 1996, 29(8): 1335-1346. DOI: [10.1016/0031-3203\(95\)00169-7](#).
- [35] 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(1): 11-12. DOI: [10.13281/j.cnki.issn.1004-4469.2015.01.005](#).
Li JJ, Xu L, Peng XY, et al. Quality standard for single fundus images in teleophthalmology (draft for comments)[J]. *Ophthalmol CHN*, 2015(1): 11-12. DOI: [10.13281/j.cnki.issn.1004-4469.2015.01.005](#).
- [36] 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](#).
Fundus Disease Group of Ophthalmology Branch of Chinese Medical Association, Fundus Disease Professional Committee of Ophthalmologist Branch of Chinese Medical Doctor Association. Guidelines for image acquisition and reading of diabetic retinopathy screening in my country (2017)[J]. *Chin J Ophthalmol*, 2017, 53(12): 890-896. DOI: [10.3760/cma.j.issn.0412-4081.2017.12.003](#).
- [37] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 25.
Zhou ZH. *Machine learning*[M]. Beijing: Tsinghua University Press, 2016: 25.
- [38] Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs[J]. *JAMA*, 2016, 316(22): 2402-2410. DOI: [10.1001/jama.2016.17216](#).
- [39] 国家药品监督管理局医疗器械技术审评中心. 人工智能医疗器械注册审查指导原则 [EB/OL]. (2022-03-09)[2022-06-16]. <https://www.cmde.org.cn/flfg/zdyz/zdyzwbk/20220309091014461.html>.
Center for Medical Device Evaluation. NMPA. Guidelines for the registration review of artificial intelligence medical devices[EB/OL]. (2022-03-09)[2022-06-16]. <https://www.cmde.org.cn/flfg/zdyz/zdyzwbk/20220309091014461.html>.
- [40] 邵志强. 抽样调查中样本容量的确定方法[J]. *统计与决策*, 2012, 22: 12-14. DOI: [10.13546/j.cnki.tjyjc.2012.22.002](#).
Shao ZQ. The method of determining sample size in a sampling survey [J]. *Statistics & Decision*. 2012, 22: 12-14. DOI: [10.13546/j.cnki.tjyjc.2012.22.002](#).
- [41] 国家食品药品监督管理总局. 医疗器械临床试验质量管理规范 [EB/OL]. (2021-09-28)[2022-06-16]. <https://www.nmpa.gov.cn/xxgk/gfwj/bmgzh/20160323141701747.html?type=pc&m=>.
National Medical Products Administration. Practice for quality management of clinical trials of medical devices[EB/OL]. (2021-09-28)[2022-06-16]. <https://www.nmpa.gov.cn/xxgk/gfwj/bmgzh/20160323141701747.html?type=pc&m=>.
- [42] 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]. (2021-09-28)[2022-06-16]. <https://www.cmde.org.cn/xwdt/zxyw/20190628151300923.html>.
Center for Medical Device Evaluation. NMPA. Key points of deep learning assisted decision making medical device software review and related description. [EB/OL]. (2021-09-28)[2022-06-16]. <https://www.cmde.org.cn/xwdt/zxyw/20190628151300923.html>.
- [43] 曹霞, 姚勇, 傅东红, 等. 无锡市50岁及以上人群年龄相关性黄斑变性流行病学调查[J]. *中国实用眼科杂志*, 2013, 31(4): 494-498. DOI: [3760/cma.j.issn.1006-4443.2013.04.032](#).
Cao J, Yao Y, Fu DH, et al. Prevalence study of age-related macular degeneration over the age of 50's in Wuxi[J]. *Chin J Pract Ophthalmol*, 2013, 31(4): 494-498. DOI: [10.3760/cma.j.issn.1006-4443.2013.04.032](#).
- [44] 杨倩, 韩磊, 刘宁, 等. 免散瞳数码眼底照相在眼底黄斑部疾病筛查中的应用[J]. *眼科新进展*, 2012, 32(5): 473-475. DOI: [10.13389/j.cnki.rao.2012.05.010](#).
Yang Q, Han C, Liu N, et al. Application of non-mydratic digital fundus photography in fundus macular disease screening[J]. *Rec Adv Ophthalmol*, 2012, 32(5): 473-475. DOI: [10.13389/j.cnki.rao.2012.05.010](#).
- [45] Jonas JB, Xu L, Wang YX. The Beijing eye study[J]. *Acta ophthalmologica*, 2009, 87(3): 247-261. DOI: [10.1111/j.1755-3768.2008.01385.x](#).
- [46] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning[J]. *Big Data*, 2019, 6(1): 1-48. DOI: [10.1186/s40537-019-0197-0](#).
- [47] Padilla R, Netto SL, Da Silva EA. A survey on performance metrics for object-detection algorithms[C]. 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), Wuhan, 2020: 237-242.
- [48] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Trans Pattern Anal Mach Intell*, 2017, 39(6): 1137-1149. DOI: [10.1109/TPAMI.2016.2577031](#).
- [49] 中华医学会健康管理学分会, 《中华健康管理学杂志》编辑委员会. 健康体检主检报告撰写专家共识[J]. *中华健康管理学杂志*, 2020, 14(1): 8-11. DOI: [10.3760/cma.j.issn.1674-0815.2020.01.003](#).
Chinese Society of Health Management, The Editorial Board of Chinese Journal of Health Management. Expert consensus on the chief physician report for health checkup[J]. *Chin J Health Manage*, 2020, 14(1): 8-11. DOI: [10.3760/cma.j.issn.1674-0815.2020.01.003](#).
- [50] 国家食品药品监督管理总局. 中华人民共和国医药行业标准: 眼科仪器眼底照相机[S]. 北京: 中国标准出版社, 2008.
National Medical Products Administration. People's Republic of China Pharmaceutical Industry Standard: ophthalmic instruments fundus camera[S]. Beijing: China Standards Press, 2008.
- [51] 宋琳琳, 李志清, 马文江, 等. 运用云医疗远程平台构建糖尿病视网膜病变分级诊疗模式[J]. *现代医院管理*, 2020, 18(3): 17-20. DOI: [10.3969/j.issn.1672-4232.2020.03.005](#).
Song LL, Li ZQ, Ma WJ, et al. Establishment of a hierarchical diagnosis and treatment model for diabetic retinopathy by cloud medical remote platform[J]. *Modern Hospital Management*, 2020, 18(3): 17-20. DOI: [10.3969/j.issn.1672-4232.2020.03.005](#).
- [52] 谭丹华, 万里飞, 邹溢辉. 医院信息系统HIS和其应用[J]. *中国医疗器械信息*, 2007, 13(9): 39-40. DOI: [10.3969/j.issn.1006-6586.2007.09.011](#).
Tan DH, Wan LF, Zou YH. Hospital information system and its application[J]. *China Medical Device Information*, 2007, 13(9): 39-40. DOI: [10.3969/j.issn.1006-6586.2007.09.011](#).
- [53] 李贵祥, 刘国祥, 李刚荣, 等. 大型综合性医院 PACS 系统设计与实施[J]. *中国医院管理*, 2005, 25(1): 46-48. DOI: [10.3969/j.issn.1001-5329.2005.01.019](#).
Li GX, Liu GX, Li GR, et al. The design and implementation of PACS system in large-scale comprehensive hospitals[J]. *Chinese Hospital Management*, 2005, 25(1): 46-48. DOI: [10.3969/j.issn.1001-5329.2005.01.019](#).
- [54] Long M, Zhu H, Wang J, et al. Unsupervised domain adaptation with residual transfer networks[C]. 30th International Conference on Neural Information Processing Systems, New York: Curran Associates Inc., 2016: 136-144.
- [55] Ganin Y, Lempitsky V. Unsupervised domain adaptation by backpropagation[C]. The 32 nd International Conference on Machine Learning, Lille, 2015: 1180-1189.

(收稿日期: 2022-06-16)
(本文编辑: 杨婷婷)