

深度学习在眼科疾病辅助诊断中的应用综述

王与点

(北京邮电大学国际学院, 北京 100088)

摘要: 随着人工智能等技术的高速进步, 深度学习等方法在医学领域的研究应用范围和影响力日益扩大, 新型技术与临床实践的结合也成为近年来的研究热点, 在眼科疾病的辅助诊断中展现出了巨大的潜力。基于卷积神经网络的深度学习算法在疾病筛查、病灶检测、组织分割等任务中表现出了巨大的潜力, 并逐渐用于青光眼、老年黄斑变性、糖尿病性视网膜疾病、白内障等多种眼科疾病的诊断和筛查。首先, 就深度学习在眼科疾病辅助诊断中的相关工作和应用做了综述, 重点介绍了各类眼科疾病的数据集、评价指标和当前的研究进展。然后, 得出结论: 深度学习在眼科疾病诊断中取得了显著成果, 但仍面临数据集规模小、类别不平衡、模型可解释性不足等挑战。最后, 对未来前景和发展做了展望。

关键词: 人工智能; 深度学习; 眼科疾病; 辅助诊断

中图分类号: TP18

文献标志码: A

doi: 10.11959/j.issn.2096-3750.2025.00471

Review of the application of deep learning in the auxiliary diagnosis of ophthalmic diseases

WANG Yudian

International School, Beijing University of Posts and Telecommunications, Beijing 100088, China

Abstract: With the rapid advancement of technologies, such as artificial intelligence, the research and application scope of methods such as deep learning in the medical field are expanding increasingly, and the combination of new technologies and clinical practice has become a research hotspot in recent years, demonstrating great potential in the auxiliary diagnosis of ophthalmic diseases. Deep learning algorithms based on convolutional neural networks have shown great potential in tasks, such as disease screening, lesion detection, and tissue segmentation, and have been gradually applied to the diagnosis and screening of various ophthalmic diseases, including glaucoma, age-related macular degeneration, diabetic retinopathy, and cataracts. Firstly, the related works and applications of deep learning in the auxiliary diagnosis of ophthalmic diseases were reviewed, focusing on the datasets, evaluation metrics, and current research progress of various ophthalmic diseases. Secondly, it was found that deep learning achieved remarkable results in the diagnosis of ophthalmic diseases, but it still faced challenges such as small dataset size, class imbalance, and insufficient model interpretability. Finally, the future prospects and development were also discussed.

Key words: artificial intelligence, deep learning, ophthalmic diseases, assisted diagnosis

0 引言

随着人工智能的发展, 深度学习在计算机视觉领域的目标检测、物体分割、物体分类等任务中取

得了巨大的成功, 在医学影像领域的检测、分类、分割和配准等任务中的应用也成为全球的研究热点。与传统图像处理技术相比, 深度学习具有更强的特征提取能力、更高的准确率和更好的鲁棒性,

收稿日期: 2025-02-14; 修回日期: 2025-06-09

通信作者: 王与点, jmmboy20031022@qq.com

同时在部分任务上,深度学习相较于眼科诊断医生取得了更好的成绩。使用深度学习技术开发的辅助诊断工具不仅提高了医疗诊断的效率和准确性,也为医生在临床治疗上提供了强有力的支撑,进而提升了患者的治疗效果和康复质量。

当前,深度学习在计算机断层扫描(CT, computed tomography)、磁共振成像(MRI, magnetic resonance imaging)、X射线(X-ray)、心脏彩超(Echo, echocardiogram)、病理切片图像(WSI, whole slide image)等多种影像数据上得到了广泛的应用,并取得了令人满意的成绩。在光学相干断层扫描(OCT, optical coherence tomography)、眼底彩照(FP, fundus photography)等眼科检查影像数据上的研究和应用正在日益增多,展现出了巨大的潜力。例如,深度学习模型在青光眼、老年黄斑变性、糖尿病性视网膜疾病和白内障等疾病的筛查和检测任务中取得了不错的表现,从而实现对疾病的早期发现或对疾病的进展进行评估,为疾病的预防和治疗提供重要信息。本文就深度学习在上述4类眼科疾病辅助诊断中的相关工作、相关数据集、评价指标和前景展望进行综述。

1 深度学习在眼部疾病辅助诊断中的应用

1.1 深度学习在青光眼辅助诊断中的应用

青光眼是一种由高眼压引起的慢性神经退行性疾病,在眼压逐渐升高的过程中,眼球内玻璃体在眼内压力的作用下逐渐向视网膜方向发生位移,进而对视神经盘、脉管系统和神经纤维层造成损害。据统计,21世纪以来,在全世界范围内的原发性青光眼患者人数接近6 680万,其中约690万人罹患失明^[1]。

眼底彩照和光学相干断层扫描是当前在临床上筛查青光眼的两种有效方式,具体可以通过眼底彩照的垂直杯盘比(vCDR, vertical cup to disc ratio)和OCT的视网膜神经纤维层厚度等显著的生物标志物来判断患者是否患有青光眼。通过深度学习的方法,模型可以自动识别、检测、分割相关生物标志物和关键结构,对青光眼进行分级,并在临床上提供相应的参考数据。

1.1.1 数据集

1) MICCAI 2021 GAMMA 挑战赛数据集
MICCAI 2021 GAMMA 挑战赛数据集^[2]中包含

300个由2D眼底彩照(2 000像素×2 992像素、1 934像素×1 956像素)和3D光学相干断层扫描(992像素×512像素×256像素)两种临床模态影像组成的数据对,其中,无青光眼病症的数据样本占50%(这些数据样本并非健康人群,而是罹患其他眼科疾病患者的数据),青光眼数据样本占50%(早期青光眼占52%,中期青光眼占28.67%,晚期青光眼占19.33%)。该数据集包含样本的青光眼分级、黄斑中央凹的坐标位置和视杯视盘的分割标注。其中,青光眼的分级由临床记录决定,黄斑中央凹的坐标标注和视杯视盘的分割标注均先由4名临床眼科医生完成,之后再由1名具有10年以上诊断经验的眼科医生校验并整合。

2) RIGA 数据集

RIGA (retinal fundus images for glaucoma analysis) 数据集^[3]包含750张用于分析青光眼的视网膜眼底图像,每幅图像均由6名资深眼科医生对视杯和视盘区域进行手动标注、提取并评估了6个参数(视盘区域和质心、视杯和质心、水平和垂直杯盘比)。在标注之后通过比较6位眼科医生的标注值,剔除了可能存在的异常值,保证了数据集的稳定性。

3) Drishti-GS1 数据集

Drishti-GS1 数据集^[4]包含101张2 896像素×1 944像素的眼底彩照图像,其中,训练集包含50张图像,测试集包含51张图像。所有图像均来自马杜赖亚拉文眼科医院,该数据集主要应用于视神经头(ONH, optic nerve head)的分割任务。该数据集包含了视杯和视盘区域的分割标注,以及青光眼的人工分级,所有图像均由4位具有不同临床经验的眼科专家标注。

4) G1020 数据集

G1020 数据集^[5]是一个大规模公开可用的视网膜眼底图像数据集,该数据集包含1 020张高分辨率的彩色视网膜眼底图像,专门用于青光眼的分类和检测任务。该数据集不仅提供了青光眼的诊断标签,还提供了视盘和视杯的详细分割标注、垂直杯盘比、神经视网膜缘大小等信息,以更好地支持多种任务以及复杂算法的开发。G1020数据集遵循了常规眼科检查中的标准实践流程,旨在成为青光眼检测的标准基准数据集,能够支持研究人员开发和评估更鲁棒、泛化能力更强的自动化青光眼诊断算法。

青光眼相关数据集各有千秋,但在不同程度

上存在局限性。青光眼相关数据集对比见表 1，MICCAI 2021 GAMMA 挑战赛数据集融合 2D 眼底彩照与 3D 光学相干断层扫描数据，多模态信息丰富，为青光眼分级提供了全面的依据。然而，在训练复杂模型时有限的规模难以充分覆盖疾病的多样性特征，限制了模型的泛化能力，面对不同临床场景时的适应性可能不足。RIGA 数据集凭借多位资深眼科医生的手动标注，对视杯和视盘区域的分割精度较高，为形态学分析奠定了良好的基础。但仅依赖眼底图像，缺乏其他模态数据，无法从多角度反映青光眼的病理特征，在多模态融合研究方面存在明显短板。Drishti-GS1 数据集虽规模较小，但其标注经多位专家协作完成，具有较高的可靠性，在视神经头分割任务中表现出色。但是，较小的规模导致数据多样性匮乏，难以满足复杂模型对数据量的需求，可能影响模型对不同病情特征的捕捉能力。G1020 数据集以大规模和高分辨率为显著优势，详细的视盘和视杯分割标注有助于开发泛化性强的诊断算法。但潜在的数据不平衡问题不容忽视，不同病情阶段数据样本分布不均，可能使模型训练产生偏差，对少数类别的识别能力下降。未来需要的青光眼数据集应在扩充数据规模的同时，确保多模态数据的均衡采集，涵盖更多样化的病例，包括不同种族、年龄段和病情严重程度，解决数据不平衡问题，统一数据标注标准，以提升模型训练效果和泛化能力。

1.1.2 评价指标

1) 准确率

准确率 (accuracy) 是一种较为常见的分类任务评价指标，准确率表示模型正确预测的样本数量在总样本中的比例，是模型预测正确的概率或者频率，计算式为：

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

其中，TP (true positive) 为真正例、FP (false positive) 为假正例、FN (false negative) 为假反例、TN (true negative) 为真反例 (true negative)。准确率可以用以衡量整个分类器在数据集上的整体性能，较高的准确率表示分类器正确分类了大部分样本。

2) 精确率

精确率 (precision) 也称查准率，指在同一模型中，所有预测结果为正例的样本中，实际标注同为正例的样本。计算式为：

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

如果把正例定义为某种疾病的阳性，反例定义为某种疾病的阴性，那么精确率就可以作为衡量模型对疾病的检测能力，精确率越高，说明在预测阳性样本上的准确率越高。

3) 召回率

召回率 (recall) 指在同一模型中，所有实际结果为正例的样本中，预测结果同为正例的样本，计算式为：

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

召回率关注模型对实际阳性样本的识别能力，重点在于模型是否可以捕捉到所有的实际阳性样本，是衡量漏检率的重要指标。

4) 基于混淆矩阵的 kappa 系数

混淆矩阵是一类用于展示分类器在测试数据集上的预测结果的二维矩阵，其中，一个维度为预测值，另一个维度为实际值，可以用来衡量分类器的预测值和实际值的一致性。例如，在青光眼等疾病的分级任务中，混淆矩阵会显示疾病分级 (如未患病、早期、中晚期) 的预测和实际分布。kappa 系数计算式为：

表 1 青光眼相关数据集对比

数据集	数据规模	标注信息	应用场景
MICCAI 2021 GAMMA 挑战赛数据集	300 个 2D 眼底彩照和 3D 光学相干断层扫描数据对	青光眼分级、黄斑中央凹坐标、视杯视盘分割标注	青光眼分级、生物标志物定位与分割研究
RIGA 数据集	750 张视网膜眼底图像	视杯和视盘区域标注、6 个相关参数	青光眼特征分析、诊断模型训练
Drishti-GS1 数据集	101 张眼底彩照图像	视杯和视盘区域分割标注、青光眼人工分级	视神经头分割任务研究
G1020 数据集	1 020 张高分辨率彩色视网膜眼底图像	青光眼诊断标签、视盘和视杯详细分割标注等多种信息	青光眼分类、检测及复杂算法开发

$$\text{kappa} = \frac{p_o - p_e}{1 - p_e} \quad (4)$$

其中, p_o 表示准确率, 计算方式为对角线元素之和与整个矩阵的元素之和的比值。 p_e 表示实际数量与预测数量的乘积和与样本总数 (整个矩阵的元素之和) 的平方的比值, kappa 系数的取值在 $[-1, 1]$ 内, 但通常不会取到负值。当 kappa 系数为 1 时, 表示分类器的预测结果和实际结果完全一致, 是一个理想值; 当 kappa 系数为 0 时, 表示分类器的预测结果与随机分类产生的结果没有差异; 当 kappa 系数的值为负数时, 表示分类器的预测结果比随机分类的结果更差。因此, 本文认为现实中分类器的 kappa 系数通常分布于 $[0, 1]$ 内。在面对不同类别的疾病分级图像数据不平衡时基于混淆矩阵的 kappa 系数有助于保持数据的一致性和可靠性。

5) F1 分数

F1 分数 (F1-score) 是精确率与召回率的调和平均数, 该指标综合考虑了对于阳性样本的预测工作中的准确性和覆盖率, 在面对数据类别不平衡的数据集时可以作为综合评估模型预测能力的指标。计算式为:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

对于上述眼科疾病的辅助诊断过程中, 预测模型既需要确保高精确率以避免误诊, 又需要保证高召回率以减少漏诊。F1 分数是对两项指标的整体表现的衡量指标。

6) 平均欧氏距离

欧氏距离 (ED, Euclidean distance) 指在任意维度中的两个点之间的真实距离, 对于在任意维度 n 中的两个点, ED 计算式为:

$$\text{ED} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \dots + (x_n - y_n)^2} \quad (6)$$

计算所有预测值和这些预测值所对应的所有实际值之间的平均欧氏距离 (AED, average Euclidean distance) 就可以实现在整组数据集的维度上对分类器的预测结果和实际结果的差异对比, AED 的值越趋近于 0, 表示预测结果越接近实际结果, 分类器的预测效果越好, 反之则表示分类器的效果越差。例如, 在黄斑中央凹的定位任务中, 可通过 AED 来衡量模型预测黄斑中央凹位置的准确度。

7) 垂直杯盘比

垂直杯盘比 (vCDR, vertical cup-to-disc ratio)

是评估青光眼眼底图像中的视神经盘是否健康的重要指标, 其具体为视神经盘杯部的垂直直径与视神经盘垂直直径的比值。通常情况下, 垂直杯盘比应当在 $[0.1, 0.7]$, vCDR 的值越高, 表示在眼底图像中视神经盘的扩张程度越大。在基于眼底图像的青光眼预测任务中, vCDR 可以客观地反映青光眼的进展。

8) Dice 系数

Dice 系数 (Dice coefficient) 是一种用于衡量两个样本集合相似度的统计方法, 通常被用于评估自动图像分割算法的性能, 具体计算式为:

$$\text{Dice} = \frac{2 \times |A \cap B|}{|A| + |B|} \quad (7)$$

其中, A 和 B 分别代表图像的分割结果和图像样本集合, Dice 系数可以用来评估这些自动分割结果与人工标注的准确性, Dice 系数的取值范围为 $[0, 1]$, 较高的 Dice 系数表示分割结果与真实标注的一致性较高。

1.1.3 相关工作

从积极方面来看, 众多学者的研究成果为青光眼诊断技术的革新奠定了坚实的基础。例如, Tampa 等^[6]提出了一种用于自动检测眼部疾病的深度学习算法, 该算法利用 Origa-light 和 Origa 数据库中的 1 300 张图像来进行训练和测试, 该团队首先使用蓝色分量提取、灰度图像转化、椭圆拟合、中值滤波、索贝尔滤波器应用和二值化的方法对图像进行预处理, 之后使用基于迁移学习的 VGGNet19 模型, 最终取得了 0.988 4 的准确率、0.976 9 的特异性、1.000 0 的敏感性、0.988 5 的 F1-Score 和 0.989 0 的 AUC (area under the curve), 证明了深度学习在青光眼自动检测中的巨大潜力, 为后续研究提供了重要的技术思路与实践范例。Sharma 等^[7]部署的卷积神经网络 (CNN, convolutional neural network) 模型包含 3 个卷积层和一个全连接层, 使用更少的可调参数进行学习, 有效地避免了过拟合的问题, 同时, 该团队使用了主成分分析 (PCA, principal component analysis) 和线性判别分析 (LDA, linear discriminant analysis) 的组合特征降维技术, 有效地降低了特征维度, 同时保留了能够区分不同类别的关键信息, 使后续的分类步骤更高效。然后, 经过降维的特征输入极限学习机 (ELM, extreme learning machine) 进行分类, 之后再引入改进的粒

子群优化算法对 ELM 的隐藏节点参数进行优化, 最终取得了 0.984 6 的准确率。该文献的思路为优化模型结构、提升模型性能提供了可行路径。Govindan 等^[9]设计了一种基于眼底彩色图像和 CNN 模型的早期青光眼识别算法, 把 ResNet50 和 InceptionV3 模型组合成一个混合模型, 最终实现了 0.992 0 的 F1-Score。这种模型融合策略在提高青光眼早期识别准确率方面成效显著, 拓宽了模型构建的思路。Aljohani 和 Aburasain^[9]使用 CNN 和随机森林搭建了一个混合青光眼检测框架, 该模型对预处理后的图像分别使用 ResNet50 和 VGG16 进行特征提取, 同时对提取的特征进行组合, 最后使用随机森林将 CNN 提取的特征组合后进行最终分类。该模型达到了 0.954 1 的准确率、0.993 7 的精确率、0.883 7 的召回率和 0.935 2 的 F1-Score。Bander 等^[10]开发了一个基于 CNN 的全自动系统, 用于非青光眼和青光眼分级工作, 该系统通过对原始图像自动提取视盘等特征, 并将这些特征输入支持向量机 (SVM, support vector machine) 分类器, 将图像分类为正常或异常。该方法的准确率、特异性和敏感性分别达到了 0.882 0、0.908 0 和 0.850 0, 实现了高效地检测青光眼病情。Wang 等^[11]设计的深度学习模型首先通过随机水平翻转、缩放、旋转、倾斜等数据增强技术, 对视野光学相干断层扫描 (VF-OCT, visual field optical coherence tomography) 图像进行处理。空间对齐的视野 (VF, visual field) 图像和 OCT 图像被输入 ViT (vision transformer) 模型中提取空间特征, 最终输入全连接的分类器中, 分类器预测指定时间范围内青光眼需要进行手术的概率。Chai 等^[12]设计了一个可以自动分割图像并提取特征的多分支神经网络 (MB-NN, multi-branch neural network) 模型, 该模型在真实数据集上实现了 0.915 1 的准确度、0.923 3 的灵敏度和 0.909 0 的特异性。Chiang 等^[13]的 CNN 模型通过集成卷积注意力模块 (CBAM, convolutional block attention module) 以提升模型性能, 实验数据使用了 5 重交叉验证, 在 3 088 张图片的数据集中集成模型表现性能最佳 AUC 为 0.894 0, 准确率为 0.821 6, 灵敏度为 0.810 4, 特异性为 0.832 7, F1-Score 为 0.819 2。Madadi 等^[14]开发了一种名为青光眼域适应 (GDA, glaucoma domain adaptation) 的深度学习模型, 用于青光眼的诊断。GDA 模型的核心目标是

通过域适应技术从不同数据源中提取青光眼的特征, 从而提升模型的泛化能力。为了实现有效的域适应, 该文献还为 GDA 引入了一种渐进式加权机制, 该机制能够在知识转移过程中逐步加权, 以确保有价值的源域知识转移到目标域, 同时减少负面迁移的影响 (即不适合的源域知识进入目标域), 该团队根据 3 种不同情况来训练 GDA, 最终达到了 0.910 0 的准确率和 0.980 0 的 AUC, 具有较强的泛化能力。Panahi 等^[15]提出了一种基于简化 U-Net 架构的快速分割算法, 应用于视网膜血管和视盘的分割任务, 可以将准确性保持在其他先进分割方法的水平的基础上, 在 0.008 s 内完成视盘分割任务, 在 0.03 s 内完成血管分割任务。Aurangzeb 等^[16]提出了一种基于 Anam-Net DNN 的轻量级模型, 用于计算 CDR 并对视杯和视盘进行分割, 该轻量化 DNN 模型通过减少计算复杂度, 高效地完成了分割任务。Kashyap 等^[17]的研究通过集成预训练的迁移学习模型和 U-Net 架构, 并使用 DenseNet-201 深度卷积神经网络 (DCNN, deep convolutional neural networks) 作为特征提取器完成青光眼的识别任务, 这种方法在处理视网膜眼底图像时能够有效地判断一个人是否患有青光眼。Sudhan 等^[18]使用基于深度学习算法的 U-Net 架构实现图像分割和预训练迁移学习模型, 在特征提取过程中同样使用 DenseNet-201 深度卷积神经网络。Abbas^[19]提出了一种基于卷积神经网络的无监督架构, 即 Glaucoma-Deep 系统, 从原始像素强度中提取多层特征, 并使用深度信念网络 (DBN, deep belief network) 模型选择出最具辨识度的特征, 最后通过 Softmax 线性分类器区分青光眼和非青光眼视网膜图像, 并在 1 200 张来自公开和私有数据集的视网膜图像上进行了测试。该系统达到了 0.845 0 的敏感度、0.980 1 的特异性、0.990 0 的准确率和 0.840 0 的精确率。Elakkiya 和 Saraniya^[20]提出了一种迁移学习模型, 用于预测视神经中的眼内压, 进而预测青光眼的发生概率。Wang 等^[21]采用了 4 种类型的输入数据 (视网膜眼底图像、感兴趣区域 (ROI, region of interest)、血管感兴趣区域 (VROI, vascular region of interest) 和颜色增强图像) 来反映结构问题, 同时通过改进损失函数解决了数据不平衡问题, 并提出了一种基于视觉大语言模型 (VLLM, vision-language large model) 的集成模型, 从而提高了青光眼分类的准确性。

尽管青光眼相关的深度学习研究取得了显著的进展,但以上工作在模型性能评估、可解释性、数据利用等方面仍然存在明显的局限性。在模型性能评估上,不少研究存在不足。Tampa等^[6]利用Origa-light和Origa数据库训练的深度学习算法,虽然在该特定数据库上取得了高准确率等良好指标,但缺乏在其他多样化数据集以及复杂临床场景中的测试。实际临床数据的多样性远超特定数据库,包含不同设备采集的图像、不同病情复杂程度的样本等,该模型在这样的环境中,泛化能力未经充分验证,可能无法保持高性能。在模型可解释性方面,多数研究存在局限性。例如, Govindan等^[8]将ResNet50和InceptionV3组合的混合模型和Aljohani等^[9]基于CNN和随机森林搭建的混合框架,虽然在青光眼诊断中具有较好的性能,但这些深度学习模型的决策过程难以解释。临床医生在使用模型辅助诊断时,需要清晰地了解模型依据何种具体特征做出了诊断决策,而此类模型复杂的网络结构和运算过程形成了一个“黑箱过程”,限制了其在临床实践中的广泛应用。数据利用问题在现有研究中较为普遍。Shoukat等^[22]开发的全自动系统使用的数据集, Hu和Wang^[23]研究中涉及的数据,都存在数据规模有限和数据不平衡的情况。Bander等^[10]的系统在利用卷积神经网络提取特征时,由于数据规模的限制,模型可能未能充分学习到青光眼的所有特征,对一些特殊的青光眼病例难以准确诊断。同时,数据集中不同青光眼阶段的数据样本数量差异可能导致模型训练偏向多数类别,对早期或轻度青光眼的识别能力不足。Wang等^[21]的研究中,数据不平衡可能使模型在预测青光眼手术概率时,对某些病情阶段的判断出现偏差,影响临床决策的准确性。

1.2 深度学习在年龄相关性黄斑变性辅助诊断中的应用

年龄相关性黄斑变性(AMD, age-related macular degeneration)是一种常见的、进行性加重的眼部致盲性疾病,该疾病通常会严重影响老年人的视力。它的主要影响范围为人眼位于视网膜后部的黄斑区域,进而影响中央视觉和细节视觉的功能。在医学界通常把AMD分为干性和湿性两种类型。年龄相关性黄斑变性是老年群体视力下降乃至失明的主要原因,至2040年,预计全世界范围内的AMD患者数量将达到2.88亿例^[24-25],随着我国人口老龄化现象加剧,在接

下来10年内,我国的AMD患者人数将持续上升。

AMD的早期病理表现为基板沉积物和基线沉积物^[26],在临床上通常可以使用眼底照相、OCT、相干光层析血管成像术(OCTA, optical coherence tomography angiography)来进行检查。通过深度学习的方式可以对眼底彩照、OCT和OCTA等模态的图像进行分析,通过对视盘区域进行分割、黄斑中央凹定位、标志物或病变区域分割,进而判断是否患有AMD,并对AMD进行分型。

1.2.1 年龄相关性黄斑变性数据集

1) ADAM挑战赛数据集

ISBI 2020曾发布一个关于年龄相关性黄斑变性的自动检测挑战赛(ADAM, automatic detection challenge on age-related macular degeneration),并提供了一个数据集^[27]。在该挑战赛的AMD检测部分,赛方发布了一个由中山大学中山眼科中心提供的包含1 200张眼底图像的综合数据集,其中包括824张分辨率为2 124像素×2 056像素的眼底图像和376张分辨率为1 444像素×1 444像素的眼底图像。该数据集样例的53%为男性、47%为女性,年龄区间在53.19岁±15.59岁。其标注包含AMD诊断标签(AMD和非AMD)、视盘区域和AMD相关病变(包括玻璃膜疣、出血、渗出物等)标签等信息,其中,276例为AMD,924例为非AMD。该数据集由7名眼科医生手动完成掩模分割和黄斑中央凹的坐标定位,是一组信息较为完善全面的AMD数据集。

2) AREDS数据集

年龄相关性眼病研究(AREDS, age-related eye disease study)^[28]旨在研究年龄相关性黄斑变性的临床过程。在AREDS发布的数据集中,被试者加入AREDS时的年龄区间在55到80岁之间,并且在临床结束后,AREDS机构额外对参与者进行了5年随访,并收集了其他基本身体检测数据。该数据集包含来自595名参与者的超过72 000张高质量眼底图像。数据集中包括患者基本信息(如年龄、性别、种族等)、眼科检查结果(如视力、眼压、视网膜图像等)、AMD分级结果、玻璃样点统计、视网膜色素是否异常、病程记录等,其中,AMD分级被分为4个类别(无病症、早期AMD、中期AMD和晚期AMD)。

3) RFMiD数据集

RFMiD数据集^[29]共收录3 200张眼底图像,该

数据集由两位专家共同标识病症类型，为眼科日常临床中的疾病检测和分类提供了重要的资源。整个数据集被分为两个部分，第一部分为RFMiD_All_Classes_Dataset，包含3 200张原始的彩色眼底png图像，其中，1 920张为训练集、640张为验证集、640张为测试集，并提供正常和患病两种标注。第二部分为RFMiD_Challenge_Dataset，包含3 200张彩色眼底图像和相同的训练集、验证集、测试集的划分，同时还提供了28类疾病的标注数据，其中，包含超过10张图像的的疾病被单独归类为独立类别，其余稀有疾病则被合并标记为“OTHER”。

AMD数据集同样优劣并存，AMD相关数据集对比见表2，ADAM挑战赛数据集包含丰富的眼底图像及详细病变标注，涵盖多种病变类型，在多标签分类任务中优势明显。然而，数据来源的局限性使其特征代表性受限，可能无法准确反映全球不同人群的AMD特征，降低了在广泛应用中模型的准确性。AREDS数据集凭借长期随访数据和疾病分级信息，为病程预测和风险评估提供了宝贵资源。但数据获取难度较大，且图像采集过程中的设备差异和时间跨度可能导致数据质量参差不齐，干扰研究结果的可靠性。RFMiD数据集经多位专家标注，提供了多种眼科疾病的分类信息，利于多疾病联合诊断。但对AMD疾病状态的标注相对简略，缺乏细致的亚型和病变程度细分，难以满足精准诊断和病情深入分析的需求。未来AMD数据集的构建应着重拓展全球多中心、多族群的数据来源，确保样本的代表性和普适性。同时，还需建立标准化的多层次标注框架，涵盖疾病亚型、病变程度、病程特征等关键维度，并采用统一的采集与质控标准，为AMD机制的研究、早期诊断及个体化治疗提供高质量的数据支撑。

1.2.2 相关工作

Peyman等^[30]提出了一种新型的多尺度卷积神经网络架构，用于AMD的准确诊断，该多尺度CNN架构具有7个卷积层，用于对是否患有AMD进行分类，同时使用sigmoid函数作为分类器，该网

络在Mendeley数据集上进行训练，并在Mendeley、OCTID、Duke、SD-OCT Noor这4个数据集上进行了测试，准确率分别达到99.73%、98.08%、96.66%和97.95%。这一成果不仅验证了多尺度架构在AMD诊断中的有效性，更为后续研究在模型构建方面提供了重要的参考范式，有力地推动了基于深度学习的AMD诊断技术向高精度方向发展。Sendeck等^[31]首先对数据集进行图像去噪并筛选低质量图像，该文献使用了经过微调的DenseNet121模型，对AMD患者进行多基因风险评分（PRS, polygenic risk score），模型通过学习眼底图像中与AMD相关的特征来估计每个受试者的PRS。在训练完成后，特征会被提取到回归模型中增强PRS的准确度。除此之外，该文献还采用了Grad-CAM（gradient-weighted class activation mapping）技术生成热力图，可以直观地展示模型在做出预测时所关注的图像区域。模型通过五重交叉验证进行全面评估，模型的平均绝对误差（MAE, mean absolute error）为0.74，均方误差（MSE, mean squared error）为0.85，均方根误差（RMSE, root mean squared error）为0.92，决定系数（R², r-squared）为0.18，平均绝对百分比误差（MAPE, mean absolute percentage error）为2.41。表明该模型在预测PRS方面具有一定的准确性和稳定性，但是尚未捕捉到所有影响PRS的因素。Sendeck^[32]在另外一项研究中采用了CNN作为核心模型，用于从OCT图像中提取特征并估计AMD的PRS，同时通过将CNN提取的特征输入到Extra Tree回归模型中，进一步增强PRS估计的准确性。在引入AMD诊断信息、年龄、吸烟史等额外特征后，模型的预测性能达到了0.54的MAE、0.44的MSE、0.66的RMSE、0.42的R²，能够较为有效地预测AMD的遗传风险。Lee等^[33]利用OCT图像开发了一种基于深度神经网络的自动化诊断模型。研究对象包含两组患者，一组为正常人群，另一组为AMD患者，每次扫描提取中心区域的11张图像进行分析，最终在不同分析层面上（独立图像层面、黄斑OCT层面和患者层面）取得了92%的

表2 AMD相关数据集对比

数据集	数据规模	标注信息	应用场景
ADAM挑战赛数据集	1 200张眼底图像	AMD诊断标签、视盘区域和相关病变标签等	AMD检测、病变区域识别研究
AREDS数据集	超72 000张高质量眼底图像	患者基本信息、眼科检查结果、AMD分级等丰富信息	AMD临床过程研究、疾病进展分析
RFMiD数据集	3 200张眼底图像	疾病标注(正常/患病及28类疾病细分标注)	眼科疾病检测和分类

准确率 (ACC, accuracy), 92% 的接收者操作特征曲线 (ROC, receiver operating characteristic), 该模型在 AMD 的诊断中具有良好的准确性和泛化性, 有望提高临床诊断效率, 减少人为诊断误差。Saha 等^[34]开发了一种用于早期 AMD OCT 生物标志物自动检测和分类的深度学习方法, 训练了一个深度卷积神经网络, 具备自动检测和分类高反射病灶、玻璃膜疣内的低反射病灶和 B 超扫描的视网膜下玻璃膜疣样沉积物。该工作经过大量实验, 最终在来自 153 名患者的约 20 000 张 OCT 图像数据集上达到了 87% 的 ACC, 该研究有望提高临床诊断效率并减少人为诊断的误差。Mathieu 等^[35]开发了一个用于自动检测眼底彩色图像中年龄相关性黄斑变性临床特征的模型 DeepAlienorNet, 共计对 7 个标签进行分类 (大软玻璃膜疣、中等软玻璃膜疣、大面积软玻璃膜疣、中央软玻璃膜疣、色素沉着过度、色素减退、晚期 AMD), 通过交叉实验验证, 最终灵敏度为 0.77, 特异性为 0.83, 该模型在预测晚期 AMD 和大面积软玻璃膜疣上表现较好。Peng 等^[37]开发了 DeepSeeNet 深度学习模型, 使用双侧彩色眼底摄影 (CFP, color fundus photography) 对 AMD 患者进行严重程度自动评分。DeepSeeNet 首先对每只眼睛检测 AMD 的风险因素 (是否存在色素异常、玻璃膜疣大小), 之后参考 AREDS 标准进行风险评分计算, 最终获得了高于人类眼科专家的诊断准确度和 kappa 系数指标。为临床医生判断疾病严重程度和制定治疗方案提供了有价值的参考信息。在糖尿病的自动筛查工作中, Ali 等^[38]首先在图像的预处理任务中使用了伽马矫正, 以此增强了相关特征的强度, 之后利用卷积神经网络对预处理后的图像进行特征提取, 同时使用长短期记忆 (LSTM, long short-term memory) 网络识别特征, 将该模型命名为 AMDNet23, 该模型最终取得了 96.5% 的准确率。Jung 等^[39]开发了一个基于 DenseNet201 模型的系统, 该系统可以预测患者在接受治疗后 AMD 的复发风险。研究选取了 100 个随机样本, 将该模型的预测性能与 6 名眼科医生的预测表现进行了比较, 该系统在对于单次注射前图像和 3 次注射前后的 4 张图像的表现均优于眼科医生的预测。与之类似, Kim 等^[40]为评估新生血管性年龄相关性黄斑变性, 分别使用了 BaseM、SegM、RawM 这 3 种模型进行训练并引入多模态数据来预

测一年后视力恶化超过最小分辨率角 1.0 的可能性, 经过实验最终得出 RawM 模型 ROC 达到 0.95, 具有很强的泛化性, 性能最佳。Chen 等^[41]通过生成对抗网络 (GAN, generative adversarial network) 模型, 从彩色眼底摄影 (CF, color fundus photography) 图像生成逼真的吲哚菁绿血管造影 (ICGA, indocyanine green angiography) 图像, 创新性地开发了一种 CF-to-ICGA 的跨模态转换算法, 并验证得出该方法在 AMD 分类中可以提升 AMD 筛查的准确率, 为 AMD 的筛查提供了一种新的解决方案, 同时为深度学习模型训练提供了数据增强的手段, 丰富了研究方法, 为解决 AMD 研究中的数据不足问题提供了创新思路。根据文献[42], 视网膜层厚度也能反映出 AMD 的进展 (随着 AMD 病情的发展, 视网膜层厚度增加), Gao 等^[43]开发了一种根据 OCT 图像显示的视网膜层厚度预测 AMD 进展的方法。该方法首先由眼科专家手动分割并通过计算水平轴上平均两层之间的像素数量来测量视网膜层厚度, 训练卷积神经网络从 OCT 图像中捕获厚度, 形成了针对特定患者的厚度时间序列, 最后使用 LSTM 来处理这些时间序列, 预测未来的厚度变化, 最终推断出患者的 AMD 未来发展情况。Ricardi 等^[44]基于深度学习模型开发了一种分割模型, 用于对新生血管性年龄相关性黄斑变性中与疾病相关的 5 种视网膜生物标志物 (视网膜内积液、视网膜下积液、视网膜下高反射材料、玻璃膜疣、新生血管性) 的自动分割。模型在检测新生血管性年龄相关性黄斑变性 (NAMD, neovascular age-related macular degeneration) 中 5 种视网膜生物标志物的任务上表现出较高的准确性: 切片水平上的平均 AUC 为 0.93 (± 0.04), 体积水平上的 AUC 为 0.88 (± 0.07), 模型自动分割与手动分割的相关性评分 (R^2) 平均为 0.89 (± 0.05), 2D 相关评分平均为 0.69 (± 0.04), Dice 系数平均得分为 0.61 (± 0.10), 模型能够实现相对精确的分割, 与手动分割结果具有较高的重叠程度。

当前 AMD 的相关研究工作存在诸多问题。在模型的普适性方面, 多数研究使用的数据集具有局限性, 导致模型在不同人群和临床环境中的泛化能力不足。例如, Thomas 等^[45]的多尺度 CNN 架构虽然在特定数据集上表现优异, 但这些数据集的样本可能存在地域、种族等局限性, 无法涵盖全球不同

人群的AMD特征,使模型在应用于其他人群时,诊断准确率可能受到影响。在研究的全面性上,部分研究仅关注单一模态的数据,忽视了多模态数据融合带来的优势。如Saha等^[35]的研究主要基于OCT图像检测生物标志物,未充分考虑将眼底彩照、OCTA等其他模态数据与OCT图像结合,可能无法全面捕捉AMD的病变信息,影响对疾病的准确诊断和评估。从临床实用性角度看,一些研究成果在实际临床应用中面临挑战。Sendeki等^[32]在估计AMD多基因风险评估的研究中,虽然采用了多种技术,但模型的预测性能仍有待提高,且计算过程相对复杂,难以在临床快速诊断中广泛应用。在AMD诊断中同样存在深度学习模型的可解释性较差的局限性。以Lee等^[34]的自动化诊断模型为例,虽然该模型诊断准确率较高,但医生难以理解模型的决策过程和依据,使临床医生在使用这些模型时对模型本身的置信度较低,极大地限制了模型在临床实践中的推广与应用。

1.3 深度学习在糖尿病视网膜病变辅助诊断中的应用

糖尿病视网膜病变(DR, diabetic retinopathy)是一种由慢性进行性的糖尿病所导致、由视网膜微血管的渗漏和阻塞而引起的眼底病变,临床表现有微血管瘤、硬性渗出、棉絮斑、新生血管、玻璃体增殖、黄斑水肿,甚至视网膜脱离等。具体可以以DR患者眼底图中是否存在从视网膜发出的异常新生血管作为判断标准,分为增殖性糖尿病视网膜病变和非增殖性糖尿病视网膜病变。据统计,全球范围内的糖尿病患者群体中约有22.7%的糖尿病视网膜病变患者。根据预测,到2030年,全球DR患者将增加25.9%,达到1.298 4亿,并且会在2045年达到1.605 0亿^[46]。

1.3.1 糖尿病视网膜病变数据集

1) OIA-DDR数据集

OIA-DDR^[47]数据集由南开大学和北京上工医信科技有限公司联合推出,包含了我国23个省份,从9 598名患者中收集到的13 673张眼底图像。该数据集包含757张标注的眼底图像,其中涵盖4种与DR相关的病变点:微动脉瘤(MA, microaneurysm)、软渗出物(SE, soft exudate)、硬渗出物(HE, hard exudate)等。数据由眼科医生手动进行像素级和bounding-box级的病变点标注,之后再根

据DR的严重程度分为5类:无DR病症、轻度DR、中度DR、重度DR和增殖性DR。

2) Maple-DR数据集

Maple-DR数据集源于MESSIDOR数据集^[48]。MESSIDOR数据集中包括1 200张糖尿病患者的彩色眼底图像(分辨率为1 440像素×960像素或2 240像素×1 488像素),并附有AMD的病变标注(视网膜出血、微动脉瘤、软性渗出、硬性渗出等)。2019年2月至2020年2月,加拿大蒙特卡罗和多伦多的7位资深眼科医生对MESSIDOR公共数据中提取的198张具有典型性的眼底图像进行了与DR相关的解剖和病理结构分割(视盘、视杯、黄斑、血管、渗出物和结节等)标注,产生了Maple-DR数据集。

3) IDRiD数据集

IDRiD数据集(Indian diabetic retinopathy image dataset)^[49]中包含516张高清视网膜图像,图像标签有两类,一类为带有DR迹象的视网膜图像,另一类为正常的视网膜图像。IDRiD数据集在像素级别上对典型病灶进行了详细标注,包括微动脉瘤、软性渗出物、硬性渗出物和出血点的精确标注。同时提供了关于糖尿病视网膜病变和糖尿病黄斑水肿严重程度的分级信息,除此之外还提供了视盘和黄斑中心的坐标。该数据集中的81张图像附有DR病灶的二值掩码和视盘掩码,适用于病灶检测和分割算法的性能评估,使其成为早期检测糖尿病视网膜病变的图像分析算法开发和评估的理想资源。

4) DRD数据集

DRD数据集(Kaggle diabetic retinopathy detection training dataset)^[50]是一个公开的大规模视网膜图像数据集,包含了来自印度和美国眼科诊所的35 126张高分辨率视网膜图像,展示了糖尿病视网膜病变不同程度的特征。每张图像由专业眼科医生根据DR严重程度分为5个等级:0级表示无病变(正常视网膜),1级表示轻度病变,2级表示中度病变,3级表示重度病变,4级表示增生性糖尿病视网膜病变(最严重)。但是该数据集的0级和1级数据数量远多于3级和4级数据量,存在较为严重的数据不平衡的问题。虽然如此,该数据集整体质量较高,样本具有典型性,具有重要的应用潜力。

5) AGAR300数据集

AGAR300数据集^[51]是专为DR中MA检测而设计的基准数据集。第一批发布的AGAR300数据集包含28张.jpg格式的彩色眼底图像,这些图像展示了典型的微动脉瘤特征。每张图像的视野角度为45°,覆盖了视网膜的主要部分,包括黄斑区和视盘区,病灶清晰可见,分辨率为2 448像素×3 264像素,清晰地呈现了微动脉瘤等微细结构,有助于开发与评估自动检测算法。数据集图像均由眼底摄影机拍摄,能够捕捉视网膜的高分辨率彩色图像,清晰地展现了眼底中的血管、视盘和黄斑结构,便于检测微动脉瘤和出血等糖尿病视网膜病变的早期病灶,适用于微动脉瘤检测算法的开发与测试。

DR数据集在推动研究进展的同时,也暴露出一些问题。DR相关数据集对比见表3,OIA-DDR数据集覆盖范围广,病变点标注和疾病严重程度分级详细,对区域性流行病学研究意义重大。但标注样本占比相对较低,大量未标注数据增加了数据处理的难度,在模型训练过程中可能无法充分利用全部数据信息,影响模型性能的提升。Maple-DR数据集基于MESSIDOR数据集进行再标注,对视盘、视杯等解剖结构的分割精确,为解剖结构自动分割研究提供了有力支持。然而,其规模相对较小,难以支撑复杂深度学习模型的充分训练,在模型的泛化和准确性方面存在提升瓶颈。IDRiD数据集以高清图像和像素级标注为特色,适合病灶检测和分割算法的性能评估。但有限的规模难以涵盖DR所有的病变特征,导致模型在面对复杂临床病例时的检测和分割能力不足,无法满足实际临床应用的多样化需求。DRD数据集规模庞大且分辨率高,是开发自动诊断算法的重要基础。但严重的数据不平衡问题使模型训练易偏向多数类别,对少数类别病变的识别能力较弱,降低了模型在实际诊断中的可靠性。未来的DR数据集的构建应着重扩大样本规模并优化数据分布,确保各类别及不同严重程度病变

数据的均衡性。提升标注效率与质量,完善多层次标注体系,并整合患者病史、治疗方案及随访结果等临床信息,从而为深度学习模型的训练提供更丰富、精准的数据支持,提升诊断准确性及临床实用价值。

1.3.2 相关工作

众多学者的努力极大地推动了DR研究的发展。Sivapriya等^[52]提出了名为ResEAD2Net的DR检测和分级方法,ResEAD2Net的设计基于U-Net架构,减少了每个阶段的下采样和上采样层数量,从原来的4个减少到2个,便于更好地识别细小特征,并且引入残差块代替普通卷积块,加快了计算速度并使用了正则化随机游走算法使分割出的血管更加完整。之后使用MLRF分类器对不同DR样本进行5个分级。模型取得了98.07%的ACC、90.24%的灵敏度、99.01%的特异性、97.51%的AUC,模型表现出色,尤其在细微血管分割和不同DR分级的准确性上具有明显优势。Dai等^[53]设计了DeepDR Plus系统,用于从糖尿病患者的眼底图像中预测在未来5年内DR的进展情况。该系统突破了传统的仅仅预测是否会发生DR的局限,创新点在于它并非只预测患者在未来5年之内是否会发生DR,而是预测DR进展的具体时间窗口。模型的一致性指数在0.754至0.846之间,综合Brier评分的范围在0.153至0.241之间,表明该模型具有较高的准确性和较小的时间误差。为临床提前干预、制定个性化治疗方案提供了关键依据。Guefrachi等^[54]提出了一个DR自动诊断的综合系统,首先对来自眼科医院的3 662张视网膜图像进行了标注,分为5个严重程度(0到4),之后团队对这些进行了图像去噪和增强对比度,以突出视网膜中的病灶区域。接着使用ResNet152-V2和InceptionResNet-V2对这些经过预处理的图像进行分级,实验结果表明,ResNet152-V2性能优异,ACC、召回率、F1-Score均接近100%,有效地解决了实际临床图像分析中

表3

DR相关数据集对比

数据集	数据规模	标注信息	应用场景
OIA-DDR数据集	13 673张眼底图像(757张标注)	4种DR相关病变点标注、DR严重程度分5类	DR病变检测、分级研究
Maple-DR数据集	198张眼底图像	与DR相关的解剖和病理结构分割标注	DR解剖和病理结构研究
IDRiD数据集	1 516张高清视网膜图像	DR迹象标注、典型病灶像素级标注、疾病严重程度分级等	DR病灶检测、分割算法评估
DRD数据集	35 126张高分辨率视网膜图像	DR严重程度5个等级标注	DR分级研究
AGAR300数据集	28张彩色眼底图像	微动脉瘤特征	微动脉瘤检测算法开发测试

的关键问题,提高了DR诊断的可靠性。DR患者的眼底彩照中视网膜血管与背景对比度低且视网膜结构复杂,使对这些彩照的检查变得困难。Soomro等^[55]与Jena等^[56]在研究中提出了一种图像增强技术,应对在部分眼底图像中对比度低难以识别的问题,在RGB颜色模型中,G通道(绿色通道)对于高亮特征呈现出较为理想的对比度,易于识别。为了进一步提升病变特征的可辨识度,在G通道使用了用于血管增强的对比限制自适应直方图均衡(CLAHE, contrast limited adaptive histogram equalisation)以达到增强图像对比度的效果,进而提升病变检测的准确性。除此之外,Jena等^[61]还进一步使用U-Net提取非对称深度特征,用来完成视盘和血管的分割任务。该文献使用了CNN提取图像的空间特征,接着使用SVM进行DR病变分类,最终取得了91.9%以上的准确率。Prabhakar等^[57]利用了U-Net、塘鹅鸚鵡优化算法、Deep Q网络和指数塘鹅鸚鵡优化算法开展糖尿病视网膜病变检测工作,在数据增强和特征提取后使用塘鹅优化算法和鸚鵡优化算法进行组合,开发出塘鹅鸚鵡优化算法以更好地识别病灶特点,进而用于训练U-Net模型。接着,使用训练好的U-Net对眼底图像中的病灶区域进行分割,为后续的分类和检测提供有效的区域。之后,使用指数塘鹅鸚鵡优化算法对Deep Q网络进行训练,使用被分割好的病灶作为输入,并使用训练好的Deep Q网络进行分类,最终实现了91.60%的准确度、92.22%的敏感性和91.91%的特异性,综合表现良好,展现出了在糖尿病视网膜病变检测任务中的优越性能。Saranya和Umamaheswari^[58]通过背景去除、视盘去除、血管提取、病灶分割来对视网膜图像进行处理,之后通过CNN进行DR不同阶段的分类任务。最终在MESSIDOR数据集上取得了97.54%的ACC、90.34%的灵敏度、98.24%的特异性和93.28%的F1-Score。Qummar等^[59]训练了5种深度卷积神经网络模型(Resnet50、Inceptionv3、Xception、Dense121和Dense169)的集合,以改进DR不同阶段的分类。Lalithadevi和Krishnaveni^[60]提出了一种名为OptiDex的创新深度学习模型,用于DR的检测和分类任务。OptiDex使用了一种结合Nas和MobileNet的架构对预处理后的图片进行特征提取,同时该模型结合了DCNN和增强型猫群优化算法,通过优化DCNN的网络参

数(如权重和偏差)来优化模型性能。经过实验,模型取得了97.65%的ACC、96.46%的灵敏度和93.45%的特异性。Mansour^[61]应用基于卷积神经网络的AlexNet DNN来实现DR CAD解决方案,该模型采用多级优化度量,结合预处理,基于自适应学习的高斯混合模型(GMM, Gaussian mixture model)区域分割、感兴趣区域(ROI, regions of interest)定位、AlexNet DNN的高维特征提取、基于主成分分析(PCA, principal component analysis)和线性判别分析(LDA, linear discriminant analysis)的特征选择以及支持向量机的分类,以确保最佳的5类DR分类。Wang等^[62]提出了一种基于梯度的元学习算法,用于完成DR样本数据有限的情况下的小样本分类任务。该文献使用FIAML-LR元学习方法,将逻辑回归纳入了元学习的内循环,帮助模型在每个任务中快速调整参数,以增强模型的可解释性和快速适应性。Singh等^[63]使用了微调的Efficient-NetB0 CNN模型和迁移学习检测糖尿病视网膜病变的不同阶段。经过灰度图像转换、单独提取RGB的3个通道、绿色通道的直方图均衡化、在红色通道上使用Sobel滤波器等多种实验,绿色通道图像的直方图均衡化处理在所有指标上表现出最优性能,ACC达到了83%,证明了该方法可以帮助模型更好地捕捉DR特征。任志远^[64]提出了3种DR的分类方法:基于DenseNet与SVM、多注意力与长短期记忆网络和双线性多注意力网络。Lim和Chen^[65]提出了一种SLIC-G混合图像处理方法,用于DR的分类任务。该方法分为两个阶段,第一阶段是通过简单线性迭代聚类将图像中的像素聚合成一组超像素,去除多余部分以减小计算复杂度。第二阶段是进行高斯平滑操作,通过模糊图像来去除噪声和不规则变化,使分割的超像素边界更平滑,避免出现杂乱无章的剪切。由于SLIC-G去除了局部冗余,深度学习模型能够专注于更重要的区域,避免受到无关细节的干扰。SLIC-G的预处理使模型在数据不平衡的DR数据集上能获得优秀的表现。

在DR的深度学习研究领域,尽管目前的研究已取得显著进展,但在模型通用性、数据处理、临床转化、研究完整性等方面存在不足。在模型通用性上,现有研究成果存在明显短板。众多研究成果通常依赖特定的数据集训练模型,导致模型通用性欠佳。如Sivapriya等^[52]的ResEAD2Net,虽在其

使用的数据集上表现卓越,可一旦应用于其他的图像数据,由于不同数据集在图像采集设备、拍摄环境、患者群体特征等方面存在差异,模型的性能便可能大打折扣。这意味着该模型难以在广泛的临床场景中稳定发挥作用,限制了其大规模推广与应用,无法有效满足不同地区、不同医疗机构对DR诊断的需求。数据处理环节同样存在问题。数据标注的准确性和一致性难以保障,成为制约研究发展的重要因素。临床转化方面,目前的研究成果与实际临床应用之间存在较大差距。许多研究专注于模型在实验室环境中的性能提升,忽视了临床应用中的实际需求。Dai等^[52]的DeepDR Plus系统在预测DR进展时间窗口上虽有创新,但在实际临床使用时,系统的操作流程可能过于复杂,对硬件设备要求较高,且缺乏与现有临床诊疗流程的有效整合。这使临床医生在繁忙的工作中难以快速上手和应用,无法充分发挥该系统的优势,阻碍了研究成果从实验室到临床实践的转化进程。目前研究的完整性也有待提高。多数研究仅聚焦于DR的诊断,忽略了与其他相关领域的关联研究。DR作为糖尿病的并发症,与患者的全身健康状况密切相关,而目前很少有研究将患者的糖尿病病史、血糖控制情况、其他并发症等多模态临床信息纳入模型考虑范围。这种片面的研究方式限制了对DR发病机制的深入理解,也影响了模型对DR病情的综合评估能力,难以满足临床全面诊疗的需求。

1.4 深度学习在白内障辅助诊断中的应用

白内障(cataract)是一种由眼睛晶状体混浊、透明度下降导致视觉障碍的疾病,是导致患者视力丧失、甚至失明的常见原因之一。截至目前,全世界估计有超过1 600万人受到白内障的困扰。年龄增长、基因遗传、长时间暴露于紫外线下、罹患糖尿病等都是白内障的诱导因素^[66-67]。白内障病例约占中低收入国家中致盲病例的50%,随着我国老龄化进程的发展,预计我国白内障致盲病例将于2050年达到2 000万^[68]。

1.4.1 白内障数据集

1) CATARACTS数据集

CATARACTS (cataract annotation and recognition of actions in surgical tools)数据集^[69]是来自康奈尔大学的公开白内障数据集,该数据集根据白内障挑战发布的训练视频生成。白内障挑战训练集包

括25个手术视频,平均时长为10 min 56 s(每秒30帧),视频内容是手术中的超声乳化过程,手术由3名不同专业水平的外科医生(1名专家、1名中级外科医生和1名实习外科医生)进行,患者平均年龄为61岁。手术过程被分为14个阶段,每个阶段随机提取10~20个具有典型性的帧,并且相邻两帧之间至少相邻3 s,最终提取4 670张分辨率为960×540的图片,之后由专业人员对数据进行手动数据标注。

2) Cataract Surgery数据集

Cataract Surgery数据集^[70]是由山西省眼科医院采集的共计20例白内障患者接受常规手术的视频资料,视频分辨率为1 920像素×1 080像素,帧率为25 FPS。这些视频中,1~7为测试集,8~10为验证集,11~20为训练集。该工作通过从训练集对应的视频中每50帧提取1帧来构建训练集,共得到3 105张图像。测试集与验证集被分为多个视频片段(测试集53个,验证集16个),同时数据标记了角膜缘后选择的4个分叉点,在每一帧中记录了4个分叉点的坐标。

3) ODIR-5K数据集

ODIR-5K (ocular disease intelligent recognition)数据集^[71]是一个用于眼科疾病智能诊断的高质量眼底图像数据集,由中国的首都医科大学和北京大学共同开发,包含5 000多名患者的眼底图像,总计约10 000张彩色图像,包含了白内障和其他眼病的标注。同时,数据集提供了患者的年龄、性别等基础信息,使其适用于单任务、多任务和多标签分类模型的研究。ODIR-5K数据集中每张图像均标注了病种标签,且由于部分患者同时存在多种眼病,适合多标签分类和共病分析。该数据集中的图像分辨率较高,图像质量较好,能够清晰地呈现细微的病灶特征,为病变的自动检测和分类提供了坚实的数据基础。

白内障数据集在各自应用场景中发挥了重要的作用,同时仍然存在改进空间。白内障相关数据集对比见表4,CATARACTS数据集通过手术视频帧提取和手动标注,为手术辅助系统开发提供了详细的手术过程信息。然而,其数据处理过程复杂且耗时,且只聚焦于手术场景,对术前诊断和病情评估的研究支持不足,极大地限制了其在白内障全病程研究中的应用。Cataract Surgery数据集利用高分辨

表 4 白内障相关数据集对比

数据集名称	数据规模	标注信息	应用场景
CATARACTS 数据集	4 670 张从手术视频提取的图片	手术过程各阶段标注	白内障手术过程分析、手术阶段识别
Cataract Surger 数据集	训练集 3 105 张图像、测试集和验证集多个视频片段	角膜缘后 4 个分叉点坐标	白内障手术中眼部位置和结构研究
ODIR-5K 数据集	约 10 000 张彩色眼底图像	白内障及其他眼病标注、患者基础信息	眼科疾病多任务、多标签分类研究

率视频和多帧提取，提供了手术中眼位和角膜缘分叉点的精确标注，对手术导航系统研究具有重要价值。但数据规模小且来源单一，难以代表不同地区和手术习惯的情况，限制了研究成果的广泛应用。ODIR-5K 数据集的多标签分类和高分辨率图像适用于白内障与其他眼科疾病的联合诊断。但该数据集对白内障专项研究的针对性不足，在白内障特征的深度挖掘和精准诊断方面存在欠缺。未来的白内障数据集应扩大数据采集范围，涵盖不同地区、不同医院和不同手术方式的数据，增加数据规模。同时，丰富数据类型，不仅包含手术过程数据，还应增加术前、术后的多模态数据，如眼底图像、超声图像等，并优化标注内容，突出某一专项特征，满足全病程研究和精准诊断的需求。

1.4.2 相关工作

众多学者的研究成果为白内障的诊疗提供了多样化的思路和方法。Kumari 和 Saxena^[72]采用一个基于 U-Net 的 SMi-UNet 架构进行特征提取，并与一个专门为眼底图像设计的卷积神经网络 RINet 相结合，用于白内障检测。该研究把 SMi-UNet 和 RINet 相结合，其中，RINet 的最后一层使用彩色热图帮助更好地展示白内障的受影响区域，除此之外，该文献引用 CLR 自动优化调整学习率，最终分类准确率达到 96%，表现出了优秀的白内障识别诊断能力，为后续研究在模型构建和训练优化上提供了借鉴。Zia 等^[73]对 SqueezeNet 的网络结构进行改进，引入了瓶颈注意力模块和附加层，优化其特征提取能力，瓶颈注意力模块可以有效地提取与白内障相关的关键特征。该文献还通过预训练模型、引入正则化和压缩连接策略有效地减少了过拟合现象，对模型性能进行了交叉验证，最终模型达到了 98.10% 的 ACC，泛化性能良好，为提升白内障诊断模型的可靠性提供了新的途径。Dong 等^[74]首先采用最大熵法，对数据集中的图像数据进行数据预处理，之后使用了深度学习网络对数据中的典型特征进行分割，最后使用了 SVM 和 Softmax 算法对图

片进行分类，实现了更好的预测准确率。该研究为白内障诊断提供了一种系统的图像处理和分类流程，对后续研究在数据处理和模型选择组合上具有参考价值。Fahmi 和 Estiko^[75]开发了一个基于卷积神经网络的分类系统，利用深度残差网络模型来对白内障图像进行分类，首先通过对图像裁剪调整大小、旋转、翻转等数据增强手段进行数据的预处理，然后通过不同层数的 ResNet 残差网络进行实验，最终得到 ResNet50 的效果最佳，达到了 92.71% 的 ACC、93.25% 的灵敏度、96.16% 的特异性。其研究过程为选择合适的网络结构和数据处理方式提供了实证依据。Jiang 等^[76]提出了一种新的深度学习模型 DAMC_Net 用于白内障分级任务。DAMC_Net 模型结合了注意力机制和多级分类器的特点，其双重注意机制通过引入空间注意机制和通道注意机制分别对病灶位置和特征通道进行聚焦，优化了模型识别能力。DAMC_Net 还通过对不同类别赋予不同的权重，惩罚错误分类较高风险的类别，例如，给病变样本更高的权重，从而减少假阴性。该模型对轻度、中度、重度白内障的检测灵敏度均高于 Densnet121 和 Resnet50，性能优异，为提高白内障分级的准确性提供了更有效的模型。Ran 等^[77]将深度卷积神经网络和随机森林 (RF, random forest) 相结合，将白内障分为 6 个等级，更加准确地对患者病情进行定位，同时该模型将 ACC 提升至 90.69%。这种模型融合的方式为更细致地评估白内障病情提供了新的思路。Yang 等^[78]构建了一个基于视网膜图像分类神经网络的白内障分类器，该方法首先使用了一种形态学图像处理技术，增强物体与前景、背景的对比度，同时使用滤波器实现精准地去除图像噪点的任务，在对于图像的分析过程中通过识别图像的亮度和纹理信息作为特征，同时选择两层 BP 神经网络构成分类器，最终实现了在短时间内保证较高的阳性识别准确率。Zhang 等^[79]提出的新型 6 级白内障分级方法结合了残差网络 (ResNet18) 中的高级特征和灰度共生矩阵 (GLCM, gray level co-

occurrence matrix) 扩展的纹理特征, 并通过多特征融合进行堆叠, 最后使用两个支持向量机(SVM, support vector machine) 分类器来获取每个眼底图像的概率输出, 并使用全连接神经网络(FCNN, fully connected neural network) 输出最终分类结果, 实现对白内障的自动分级。Gu等^[80]提出了一种名为Ranking-MFCNet的方法, 用于对白内障进行严重程度分级, Ranking-MFCNet引入了一个基于排名的分类框架, 将白内障的严重程度分类任务视为多个二元分类任务的组合, 模型的MFCNet可以捕捉到较小尺度上的纹理细节, 除此之外, Ranking-MFCNet还引入了eaMFC模块, 通过外部注意力机制强化多尺度特征的捕捉能力, 可以进一步加强对于微小细节的捕捉能力。Yadav和Yadav^[81]通过在深度学习模型中引入二维离散傅里叶变换(DFT, discrete Fourier transform) 光谱, 提出了一种白内障早期的分类方法。该文献先利用2D-DFT计算眼底图像的频谱图, 然后在该频谱图上进行特征提取, 最后由Softmax完成分类任务, 该方法最终取得了93.10%的准确率。

尽管深度学习在白内障研究领域已取得了一定的成果, 为该疾病的诊疗带来了新的技术思路, 但仍面临诸多挑战, 在模型性能、临床转化及研究的全面性等维度存在显著短板, 极大地阻碍了该领域的深入发展和临床应用的广泛普及。当前白内障相关研究中的模型在泛化能力上表现欠佳。多数模型基于特定数据集进行训练, 这些数据集通常具有局限性, 无法涵盖所有可能的白内障病例特征。不同地区人群的白内障可能因遗传、环境等因素呈现出不同特点, 不同设备采集的图像在分辨率、成像质量等方面也存在差异。Kumari等^[72]的基于U-Net的SMi-UNet和RINet结合的模型, 虽在训练数据集上分类准确率达到96%, 但面对其他来源的数据集或实际临床中复杂多样的图像时, 性能可能会受到显著影响。模型在实际应用中可能无法准确诊断各种类型的白内障, 降低诊断的可靠性, 从研究成果到临床应用的转化过程面临诸多障碍。在操作复杂性上, 许多模型的使用需要专业的技术知识和复杂的操作流程, 这对于临床一线医生来说难以快速掌握和应用。Jiang等^[76]的DAMC_Net模型虽然在白内障分级上表现出色, 但可能因复杂的网络结构和操作要求, 在临床实际使用中存在困难。在与现有

临床工作流程的融合方面, 多数研究并未充分考虑如何与医院现有的诊疗流程相结合, 导致模型难以融入日常临床工作, 无法有效地提高诊疗效率。此外, 目前的研究仍存在研究全面性的欠缺。目前的研究主要聚焦于白内障的诊断和分级, 对疾病的其他重要方面关注不足。在发病机制研究上, 虽然深度学习模型能够识别白内障的特征, 但对于白内障发生发展的内在机制研究较少, 无法为预防和治疗提供深入的理论依据。在与其他眼部疾病的关联性研究方面, 白内障通常与其他眼部疾病相互影响, 然而现有研究对此涉及不多, 难以满足临床对复杂眼部疾病综合诊断和治疗的需求。

2 结束语

近年来, 随着人工智能和计算机视觉等技术的快速发展, 使深度学习应用在医学领域辅助诊断治疗成为可能, 并且目前已有大量成熟的应用在临床上投入使用, 尤其是对于基眼底彩照和OCT等图像数据进行检查的眼科疾病筛查和诊断。但在现阶段的实际应用中还存在一些问题。

1) 数据集规模小、数据量少。目前应用于医学领域, 特别是眼科相关疾病的训练数据还是较少, 一方面是由于数据标注成本高, 以及数据隐私安全等原因, 各个数据集的规模都比较小, 另一方面各个医院使用的数据采集设备不一致, 导致数据格式和质量多种多样, 这对于训练一个高准确率、高鲁棒性的模型具有一定的挑战性。未来可以通过一系列的图像处理和生成训练图像等方式来进行数据增广, 以尽可能地提高训练集的数据量。

2) 类别不平衡问题。在医学领域的数据集和实际临床应用中, 类别不平衡问题对于分类任务是一个普遍存在的挑战, 由于某些疾病的分类或亚型分布不均, 部分类别的数据量远远少于主要类别, 这极大地影响了深度学习模型的训练效果和泛化能力。未来一方面可以通过重采样或不同的数据增广权重来调整数据集各类别的比例, 另一方面在模型层面可以设计用于不平衡数据的算法, 或在特征的选择和提取以及损失函数上做相应的改进, 从而在一定程度上缓解医学领域数据集类别不平衡的问题。

3) 可解释性问题。现阶段深度学习模型的可解释性仍然是一大难题, 对于在医学领域的应用来

说, 医生在临床决策中通常需要相应的诊断依据, 但目前大多数基于深度学习的模型算法在决策过程中仍缺乏可解释性。未来可以通过使用多模态数据训练、优化算法等方式来提高模型可信度以及泛化性, 同时可以通过引入注意力机制、可视化技术手段, 帮助医生理解模型的决策过程, 增强模型的可信度和临床实用性。

4) 研究全面性问题。当前的研究主要聚焦于疾病的诊断和分级, 而对疾病的发病机制、预防措施、与其他疾病的关联性等方面的研究相对不足。未来的研究应更加注重疾病的全面性研究, 探索深度学习在疾病预防、机制研究、多病联合诊断等方面的应用, 以推动眼科疾病的综合诊疗水平提升。

5) 普适性和临床转化能力不足。许多模型在特定数据集上表现优异, 但在面对不同来源的数据或复杂临床场景时, 性能大幅下降, 泛化能力不足。且对于医务工作者具有较大的操作复杂性, 与现有临床流程的整合难度较大。未来的研究应更注重模型的跨数据集、跨设备、跨族群的泛化能力, 探索多模态数据融合、迁移学习等技术, 提升模型在不同临床场景中的适应性和稳定性。注重与临床实际需求的结合, 开发易于操作、与现有诊疗流程无缝衔接的辅助诊断工具, 以加速研究成果的临床应用。

通过不断地技术创新和跨学科合作, 深度学习将在医学领域发挥更重要的作用, 通过进一步的优化, 未来在眼科疾病的辅助诊断治疗上一定会有更广泛的应用, 为提高医生诊断的准确率和效率, 以及为患者提供更高效便捷的医疗服务做出更大的贡献。

参考文献:

- [1] QUIGLEY H A. Number of people with glaucoma worldwide[J]. *British Journal of Ophthalmology*, 1996, 80(5): 389-393.
- [2] WU J D, FANG H H, LI F, et al. GAMMA challenge: glaucoma grading from multi-modality images[J]. *Medical Image Analysis*, 2023, 90: 102938.
- [3] ALMAZROA A A, ALODHAYB S, OSMAN E, et al. Retinal fundus images for glaucoma analysis: the RIGA dataset[C]//*Proceedings of the Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*. SPIE, 2018: 8.
- [4] SIVASWAMY J, KRISHNADAS S R, DATT JOSHI G, et al. Drishti-GS: retinal image dataset for optic nerve head(ONH) segmentation[C]//*Proceedings of the 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*. Piscataway: IEEE Press, 2014: 53-56.
- [5] BAJWA M N, SINGH G A P, NEUMEIER W, et al. G1020 dataset [DB]. 2023.
- [6] TAMPA H, MEKONGO M, TIEDEU A. Deep learning-based algorithm for automated detection of glaucoma on eye fundus images[J]. *Multimedia Tools and Applications*, 2024: 1-18.
- [7] SHARMA S K, MUDULI D, PRIYADARSHINI R, et al. An evolutionary supply chain management service model based on deep learning features for automated glaucoma detection using fundus images[J]. *Engineering Applications of Artificial Intelligence*, 2024, 128: 107449.
- [8] GOVINDAN M, DHAKSHNAMURTHY V K, SREERANGAN K, et al. A framework for early detection of glaucoma in retinal fundus images using deep learning[J]. *Engineering Proceedings*, 2024, 62(1): 3.
- [9] ALJOHANI A, ABURASAIN R Y. A hybrid framework for glaucoma detection through federated machine learning and deep learning models[J]. *BMC Medical Informatics and Decision Making*, 2024, 24(1): 115.
- [10] AL-BANDER B, AL-NUAIMY W, AL-TAEE M A, et al. Automated glaucoma diagnosis using deep learning approach[C]//*Proceedings of the 2017 14th International Multi-Conference on Systems, Signals & Devices (SSD)*. Piscataway: IEEE Press, 2017: 207-210.
- [11] WANG R L, BRADLEY C, HERBERT P, et al. Deep learning-based identification of eyes at risk for glaucoma surgery[J]. *Scientific Reports*, 2024, 14: 599.
- [12] CHAI Y D, LIU H Y, XU J. Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models[J]. *Knowledge-Based Systems*, 2018, 161: 147-156.
- [13] CHIANG Y Y, CHEN C L, CHEN Y H. Deep learning evaluation of glaucoma detection using fundus photographs in highly myopic populations[J]. *Biomedicine*, 2024, 12(7): 1394.
- [14] MADADI Y, ABU-SERHAN H, YOUSEFI S. Domain adaptation-based deep learning model for forecasting and diagnosis of glaucoma disease[J]. *Biomedical Signal Processing and Control*, 2024, 92: 106061.
- [15] PANAHI A, ASKARI MOGHADAM R, TARVIRDIZADEH B, et al. Simplified U-Net as a deep learning intelligent medical assistive tool in glaucoma detection[J]. *Evolutionary Intelligence*, 2024, 17(2): 1023-1034.
- [16] AURANGZEB K, HAIDER S I, ALHUSSEIN M. Refined AnamNet: lightweight deep learning model for improved segmentation performance of optic cup and disc for glaucoma diagnosis[J]. *Computers, Materials & Continua*, 2024, 80(1): 1381-1405.
- [17] KASHYAP R, NAIR R, GANGADHARAN S M P, et al. Glaucoma detection and classification using improved U-Net deep learning model[J]. *Healthcare*, 2022, 10(12): 2497.
- [18] SUDHAN M B, SINTHUJA M, PRAVINTH RAJA S, et al. Segmentation and classification of glaucoma using U-Net with deep

- learning model[J]. *Journal of Healthcare Engineering*, 2022, 2022(1): 1601354.
- [19] ABBAS Q. Glaucoma-deep: detection of glaucoma eye disease on retinal fundus images using deep learning [J]. *International Journal of Advanced Computer Science and Applications*, 2017, 8(6): 79-85.
- [20] ELAKKIYA B, SARANIYA O. A comparative analysis of pre-trained and transfer-learning model for automatic diagnosis of glaucoma[C]//*Proceedings of the 2019 11th International Conference on Advanced Computing (ICoAC)*. Piscataway: IEEE Press, 2019: 167-172.
- [21] WANG S, KIM B, KANG J, et al. Precision diagnosis of glaucoma with VLLM ensemble deep learning[J]. *Applied Sciences*, 2024, 14(11): 4588.
- [22] SHOUKAT A, AKBAR S, HASSAN S A, et al. Automatic diagnosis of glaucoma from retinal images using deep learning approach[J]. *Diagnostics*, 2023, 13(10): 1738.
- [23] HU W, WANG S Y. Predicting glaucoma progression requiring surgery using clinical free-text notes and transfer learning with transformers[J]. *Translational Vision Science & Technology*, 2022, 11(3): 37.
- [24] WONG W L, SU X Y, LI X, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis[J]. *The Lancet Global Health*, 2014, 2(2): e106-e116.
- [25] YE H H, ZHANG Q, LIU X H, et al. Prevalence of age-related macular degeneration in an elderly urban Chinese population in China: the Jiangning Eye Study[J]. *Investigative Ophthalmology & Visual Science*, 2014, 55(10): 6374-6380.
- [26] 许迅, 刘堃, 苏莉. 《中国年龄相关性黄斑变性临床诊疗指南(2023年)》更新点[J]. *中华眼底病杂志*, 2023(11): 879-882.
- XU X, LIU K, SU L. Update points in the Evidence-based guidelines for diagnosis and treatment of age-related macular degeneration in China(2023) [J]. *Chinese Journal of Ocular Fundus Diseases*, 2023(11): 879-882.
- [27] FANG H H, LI F, FU H Z, et al. ADAM challenge: detecting age-related macular degeneration from fundus images[J]. *IEEE Transactions on Medical Imaging*, 2022, 41(10): 2828-2847.
- [28] National Eye Institute of the United State. NEI age-related eye disease study (AREDS) [DB]. 2023.
- [29] PACHADE S, PORWAL P, THULKAR D, et al. Retinal fundus multi-disease image dataset (RFMiD) [DB]. *IEEE Dataport*, 2020. doi:10.21227/s3g7-st65.
- [30] GHOLAMI P, ROY P, LAKSHMINARAYANAN V. Age-related macular degeneration retinal OCT images [DB]. (2018)
- [31] SENDECKI A, LEDWOŃ D, TUSZY A, et al. Fundus image deep learning study to explore the association of retinal morphology with age-related macular degeneration polygenic risk score[J]. *Biomedicines*, 2024, 12(9): 2092.
- [32] SENDECKI A, LEDWOŃ D, NYCZ J, et al. A deep learning approach to explore the association of age-related macular degeneration polygenic risk score with retinal optical coherence tomography: a preliminary study[J]. *Acta Ophthalmologica*, 2024, 102(7): e1029-e1039.
- [33] LEE C S, BAUGHMAN D M, LEE A Y. Deep learning is effective for classifying normal versus age-related macular degeneration OCT images[J]. *Ophthalmology Retina*, 2017, 1(4): 322-327.
- [34] SAHA S, NASSISI M, WANG M, et al. Automated detection and classification of early AMD biomarkers using deep learning[J]. *Scientific Reports*, 2019, 9: 10990.
- [35] MATHIEU A, AJANA S, KOROBELNIK J F, et al. DeepAlienNet: a deep learning model to extract clinical features from colour fundus photography in age-related macular degeneration[J]. *Acta Ophthalmologica*, 2024, 102(5): e823-e830.
- [36] PENG Y F, DHARSSI S, CHEN Q Y, et al. DeepSeeNet: a deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs[J]. *Ophthalmology*, 2019, 126(4): 565-575.
- [37] ALI M A, HOSSAIN M S, HOSSAIN M K, et al. AMDNet23: Hybrid CNN-LSTM deep learning approach with enhanced pre-processing for age-related macular degeneration (AMD) detection[J]. *Intelligent Systems with Applications*, 2024, 21: 200334.
- [38] JUNG J, HAN J, HAN J M, et al. Prediction of neovascular age-related macular degeneration recurrence using optical coherence tomography images with a deep neural network[J]. *Scientific Reports*, 2024, 14: 5854.
- [39] KIM N, LEE M, CHUNG H, et al. Prediction of post-treatment visual acuity in age-related macular degeneration patients with an interpretable machine learning method[J]. *Translational Vision Science & Technology*, 2024, 13(9): 3.
- [40] CHEN R Y, ZHANG W Y, SONG F, et al. Translating color fundus photography to indocyanine green angiography using deep learning for age-related macular degeneration screening[J]. *NPJ Digital Medicine*, 2024, 7: 34.
- [41] SAßMANNSHAUSEN M, BEHNING C, WEINZ J, et al. Characteristics and spatial distribution of structural features in age-related macular degeneration: a MACUSTAR study report[J]. *Ophthalmology Retina*, 2023, 7(5): 420-430.
- [42] GAO Q T, AMASON J, RATHINAVELU J, et al. Deep learning for longitudinal OCT retinal layer thickness and AMD progression prediction[J]. *Investigative Ophthalmology & Visual Science*, 2023, 64(8).
- [43] RICARDI F, OAKLEY J, RUSSAKOFF D, et al. Validation of a deep learning model for automatic detection and quantification of five OCT critical retinal features associated with neovascular age-related macular degeneration[J]. *British Journal of Ophthalmology*, 2024, 108(10): 1436-1442.
- [44] THOMAS A, HARIKRISHNAN P M, KRISHNA A K, et al. A novel multiscale convolutional neural network based age-related macular degeneration detection using OCT images[J]. *Biomedical Signal Processing and Control*, 2021, 67: 102538.

- [45] 刘雪立. 全球糖尿病视网膜病变患病率和2045年负担预测: 美国眼科学会报告要点摘编[J]. 眼科新进展, 2024, 44(1): 80-81.
LIU X L. Global prevalence of diabetic retinopathy and its burden forecast in 2045: excerpts from the report of American Ophthalmology Association[J]. *Recent Advances in Ophthalmology*, 2024, 44(1): 80-81.
- [46] LI T, GAO Y Q, WANG K, et al. Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening[J]. *Information Sciences*, 2019, 501: 511-522.
- [47] DECENCIÈRE E, ZHANG X W, CAZUGUEL G, et al. Feedback on a publicly distributed image database: the messidor database[J]. *Image Analysis & Stereology*, 2014, 33(3): 231.
- [48] PORWAL P, PACHADE S, KAMBLE R, et al. Indian diabetic retinopathy image dataset (IDRiD): a database for diabetic retinopathy screening research[J]. *Data*, 2018, 3(3): 25.
- [49] Kaggle diabetic retinopathy detection training dataset (DRD) [DB]. 2019.
- [50] JEBA D D, PRIESTLY S B. Diabetic retinopathy-fundus image dataset (AGAR300) [DB]. 2020. DOI:10.21227/fsnq-tn19.
- [51] SIVAPRIYA G, MANJULA DEVI R, KEERTHIKA P, et al. Automated diagnostic classification of diabetic retinopathy with microvascular structure of fundus images using deep learning method[J]. *Biomedical Signal Processing and Control*, 2024, 88: 105616.
- [52] DAI L, SHENG B, CHEN T L, et al. A deep learning system for predicting time to progression of diabetic retinopathy[J]. *Nature Medicine*, 2024, 30(2): 584-594.
- [53] GUEFRACHI S, ECHTIOUI A, HAMAM H. Automated diabetic retinopathy screening using deep learning[J]. *Multimedia Tools and Applications*, 2024, 83(24): 65249-65266.
- [54] SOOMRO T A, GAO J B, KHAN M A U, et al. Role of image contrast enhancement technique for ophthalmologist as diagnostic tool for diabetic retinopathy[C]//*Proceedings of the 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. Piscataway: IEEE Press, 2016: 1-8.
- [55] JENA P K, KHUNTIA B, PALAI C, et al. A novel approach for diabetic retinopathy screening using asymmetric deep learning features[J]. *Big Data and Cognitive Computing*, 2023, 7(1): 25.
- [56] PRABHAKAR T, MADHUSUDHANA RAO T V, MARAM B, et al. Exponential gannet firefly optimization algorithm enabled deep learning for diabetic retinopathy detection[J]. *Biomedical Signal Processing and Control*, 2024, 87: 105376.
- [57] SARANYA P, UMAMAHESWARI K M. Detection of exudates from retinal images for non-proliferative diabetic retinopathy detection using deep learning model[J]. *Multimedia Tools and Applications*, 2024, 83(17): 52253-52273.
- [58] QUMMAR S, KHAN F G, SHAH S, et al. A deep learning ensemble approach for diabetic retinopathy detection[J]. *IEEE Access*, 2019, 7: 150530-150539.
- [59] LALITHADEVI B, KRISHNAVENI S. Diabetic retinopathy detection and severity classification using optimized deep learning with explainable AI technique[J]. *Multimedia Tools and Applications*, 2024, 83(42): 89949-90013.
- [60] MANSOUR R F. Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy[J]. *Biomedical Engineering Letters*, 2018, 8(1): 41-57.
- [61] WANG M F, GONG Q Z, WAN Q, et al. A fast interpretable adaptive meta-learning enhanced deep learning framework for diagnosis of diabetic retinopathy[J]. *Expert Systems with Applications*, 2024, 244: 123074.
- [62] SINGH S P, GUPTA P, DUNG R. Diabetic retinopathy detection by fundus images using fine tuned deep learning model[J]. *Multimedia Tools and Applications*, 2024, 83(39): 86657-86679.
- [63] 任志远, 基于深度学习的糖尿病视网膜病变分类方法研究[D]. 齐鲁工业大学, 2024.
REN Z Y. Research on classification methods of diabetic retinopathy based on deep learning [D]. Qilu University of Technology, 2024.
- [64] LIM W X, CHEN Z Y. Enhancing deep learning pre-trained networks on diabetic retinopathy fundus photographs with SLIC-G[J]. *Medical & Biological Engineering & Computing*, 2024, 62(8): 2571-2583.
- [65] ASBELL P, DUALAN I, MINDEL J, et al. Age-related cataract[J]. *The Lancet*, 2005, 365(9459): 599-609.
- [66] GUPTA V B, RAJAGOPALA M, RAVISHANKAR B. Etiopathogenesis of cataract: an appraisal[J]. *Indian Journal of Ophthalmology*, 2014, 62(2): 103-110.
- [67] 颜钰玲, 薛春燕. 人工智能在白内障诊断的应用进展[J]. *眼科学报*, 2024, 39(3): 160-168.
YAN Y L, XUE C Y. Advances in artificial intelligence for cataract diagnosis[J]. *Eye Science*, 2024, 39(3): 160-168.
- [68] AL HAJJ H, LAMARD M, CONZE P H, et al. CATARACTS: Challenge on automatic tool annotation for cataract surgery[J]. *Medical Image Analysis*, 2019, 52: 24-41.
- [69] ZHAI Y X, ZHANG G H, ZHENG L S, et al. Cataract surgery dataset for eye positioning and alignment[DB]. *IEEE Dataport*, 2021. doi:10.21227/1v4j-9f35.
- [70] SCHWARTZ K. Ocular disease intelligent recognition[DB]. 2023.
- [71] KUMARI P, SAXENA P. Cataract detection and visualization based on multi-scale deep features by RINet tuned with cyclic learning rate hyperparameter[J]. *Biomedical Signal Processing and Control*, 2024, 87: 105452.
- [72] ZIA A, MAHUM R, AHMAD N, et al. Eye diseases detection using deep learning with BAM attention module[J]. *Multimedia Tools and Applications*, 2024, 83(20): 59061-59084.
- [73] DONG Y Y, ZHANG Q Y, QIAO Z Q, et al. Classification of cataract fundus image based on deep learning[C]//*Proceedings of the 2017 IEEE International Conference on Imaging Systems and Techniques (IST)*. Piscataway: IEEE Press, 2017: 1-5.
- [74] FAHMI F, ESTIKO H. Cataract detection using deep learning in teleophthalmology systems[C]//*Proceedings of the 2024 14th In-*

- ternational Conference on Electrical Engineering (ICEENG). Piscataway: IEEE Press, 2024: 260-264.
- [75] JIANG J W, ZHANG Y, XIE H, et al. A deep learning based fine-grained classification algorithm for grading of visual impairment in cataract patients[J]. Optoelectronics Letters, 2024, 20(1): 48-57.
- [76] RAN J, NIU K, HE Z Q, et al. Cataract detection and grading based on combination of deep convolutional neural network and random forests[C]//Proceedings of the 2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC). Piscataway: IEEE Press, 2018: 155-159.
- [77] YANG M M, YANG J J, ZHANG Q Y, et al. Classification of retinal image for automatic cataract detection[C]//Proceedings of the 2013 IEEE 15th International Conference on e-Health Networking, Applications and Services (Healthcom 2013). Piscataway: IEEE Press, 2013: 674-679.
- [78] ZHANG H Y, NIU K, XIONG Y M, et al. Automatic cataract grading methods based on deep learning[J]. Computer Methods and

Programs in Biomedicine, 2019, 182: 104978.

- [79] GU Y Y, FANG L X, MOU L, et al. A ranking-based multi-scale feature calibration network for nuclear cataract grading in AS-OCT images[J]. Biomedical Signal Processing and Control, 2024, 90: 105836.
- [80] YADAV S, YADAV J K P S. Automatic cataract severity detection and grading using deep learning[J]. Journal of Sensors, 2023(1): 2973836.

[作者简介]



王与点(2003-)，男，北京邮电大学国际学院在读，主要研究方向为物联网工程、人工智能、深度学习。