

深度学习在糖尿病视网膜病变分级中的应用^①



张志强¹, 赵可辉², 牛惠芳³, 张子宇¹, 周连田⁴

¹(山东中医药大学 智能与信息工程学院, 济南 250355)

²(山东中医药大学第二附属医院 特检科, 济南 250002)

³(山东省药品不良反应监测中心, 济南 250014)

⁴(菏泽市中医医院 碎石科, 菏泽 274035)

通信作者: 赵可辉, E-mail: zhaokh1202@sina.com

摘要: 近年来, 糖尿病视网膜病变 (diabetic retinopathy, DR) 成为全球失明人口增加的主要原因, 早期的 DR 严重程度分级对防止 DR 患者视力丧失尤为重要。由于糖尿病患者数量的逐年上升, DR 分级的需求量也不断增加, 然而传统的人工分级不能满足日益增长的需求, 且人工分级耗时费力。深度学习技术的发展, 为 DR 检测和分级提供了高效率且更可靠的手段。虽然, 目前的 DR 二元检测已经取得十分好的效果, 然而由于糖尿病视网膜病变的复杂性和病变程度之间的差距细微, DR 严重程度分级仍然是一个具有挑战性的问题。本文对近年来涌现的 DR 分级方法进行了研究和总结: 介绍了基于 VGG、InceptionNet、ResNet、EfficientNet、DenseNet、CapsNet 模型的 6 种深度学习分级方法; 并介绍了基于多网络融合的 DR 分级方法; 最后对基于深度学习的 DR 分级方法的研究趋势进行总结和展望。

关键词: 糖尿病视网膜病变; 深度学习; 卷积神经网络; 多网络融合; 多标签分类

引用格式: 张志强, 赵可辉, 牛惠芳, 张子宇, 周连田. 深度学习在糖尿病视网膜病变分级中的应用. 计算机系统应用, 2024, 33(1): 231-244. <http://www.c-s-a.org.cn/1003-3254/9368.html>

Application of Deep Learning in Grading of Diabetic Retinopathy

ZHANG Zhi-Qiang¹, ZHAO Ke-Hui², NIU Hui-Fang³, ZHANG Zi-Yu¹, ZHOU Lian-Tian⁴

¹(College of Intelligence and Information Engineering, Shandong University of Traditional Chinese Medicine, Jinan 250355, China)

²(Department of Special Examination, The Second Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan 250002, China)

³(Shandong Province Adverse Drug Reaction Monitoring Center, Jinan 250014, China)

⁴(Department of Lithotripsy, Heze Traditional Chinese Medicine Hospital, Heze 274035, China)

Abstract: In recent years, diabetic retinopathy (DR) has become the main reason for the global blind population increase. The early DR severity classification is particularly important to prevent vision loss in DR patients. As the number of diabetes patients grows year by year, the demand for DR grading is also rising. However, the traditional manual grading cannot meet the growing demands, and it is time-consuming and laborious. The development of deep learning technology provides a more efficient and reliable means for DR detection and grading. Although the current DR binary detection has yielded good results, DR severity grading is still challenging due to the slight differences between DR complexity and lesion degree. This work studies and summarizes DR grading methods in recent years. It introduces six deep learning classification methods based on VGG, InceptionNet, ResNet, EfficientNet, DenseNet, and CapsNet models. In addition, the study presents DR grading methods based on multi-network fusion. Finally, summary and prospect are provided for the research trends of DR grading methods based on deep learning.

① 基金项目: 中国药品监管科学研究行动计划第二批重点项目 (2022SDADRKY06)

收稿时间: 2023-07-23; 修改时间: 2023-08-21; 采用时间: 2023-08-24; csa 在线出版时间: 2023-11-17

CNKI 网络首发时间: 2023-11-20

Key words: diabetic retinopathy; deep learning; convolutional neural network (CNN); multi-network fusion; multi-label classification

近年来, 全球范围内糖尿病患者的数量逐年增加, 经查阅 IDF 糖尿病地图谱^[1], 得知在全球范围已经有 5 亿人口被诊断出患有此病, 且预计到 2045 年, 这个数字将达到 7 亿之多. 由此看来, 糖尿病已经成为一个全球性的健康问题. 并且 IDF 糖尿病地图谱^[1]中还指出, 2040 年约有 1/3 的糖尿病患者将会有患有糖尿病视网膜病变 (diabetic retinopathy, DR). 在国内, Yin 等^[2]的有关研究, 指出了糖尿病患者患糖尿病视网膜病变的患病率, 并用统计学方法分析了糖尿病和 DR 之间关联性, 且指出 DR 是导致我国人口失明的主要原因.

DR 是一种视网膜血管受损导致的眼部病变, 视网膜血管会表现为出血、渗出物、微动脉瘤、畸形和血管扭曲等症状, 甚至可能导致视网膜血管的异常生长,

图 1 显示了有无 DR 的不同临床表现. 依据 DR 病变严重程度可以将 DR 分为 5 个阶段, 分别为无病变 (NO-DR); 轻度非增殖性病变 (Mild NPDR); 中度非增殖性病变 (Moderate NPDR); 重度非增殖性病变 (Severe NPDR); 增殖性病变 (PDR)^[3], 这 5 个不同病变等级有着不同的临床表现 (如图 2 所示), 临床医生可以通过检查患者的眼底图像并结合 DR 分级标准 (见表 1) 来评估 DR 的严重程度, 然而这种传统的人工分级将耗费大量时间, 且就 DR 患者的数量而言, 目前临床经验十足的医生还是相对缺乏的^[4]. 然而 DR 的早期检测和分级对于防止 DR 患者视力失明尤为重要^[5], 随着深度学习在临床诊断中的广泛应用, 深度学习为 DR 的检测和分级提供了更为高效且可靠的诊断手段.

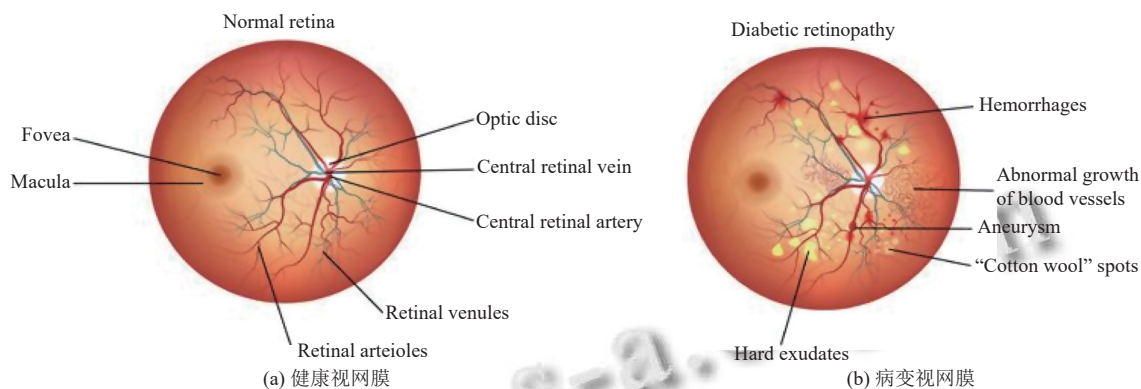


图 1 健康视网膜和病变视网膜对比图^[4]

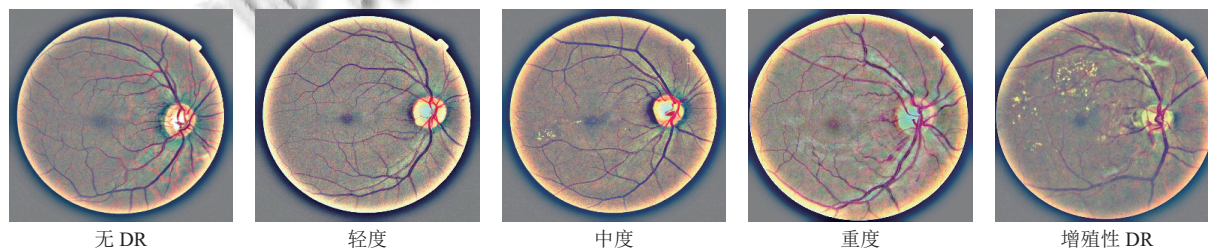


图 2 糖尿病病变的 5 个阶段

目前基于深度学习的 DR 分级综述研究已经取得了丰富的成果, Ismail 等^[6]总结归纳了几种传统机器学习方法在 DR 分级中的应用, 并通过对比各类方法的

性能指标, 最终得出贝叶斯模型更符合 DR 分级的性能要求. 虽然贝叶斯模型适合多分类实时预测, 然而贝叶斯模型所要求的样本分布独立在现实应用中是很难

实现的,且该文章的总结研究局限于传统机器学习算法.范家伟等^[7]从视网膜成像技术、图像标注方法、DR病灶区域分割、DR二分类检测、DR多分类分级等方面依次阐述,其中着重介绍了DR病灶区域分割方法,且该文的出发点是对DR智能诊断的概括性综述,缺乏对DR分级的详细化总结.Tsiknakis等^[5]依照DR分级数据集展开阐述,主要对比不同深度学习模型在同一DR数据集上的性能表现,且主要评价标准局限于高精度指标,模型的可解释性和轻量化问题并没有过多提及.由此可见,DR分级研究需要更系统化和更详细化的综述研究.因此,本文的主要贡献集中在以下3点.

(1) 详细且系统地总结深度学习在DR分级中的应用,分别以数据集质量、数据预处理方式、模型基础架构及改进方法等作为关键变量进行探讨研究.

(2) 按照深度学习分类网络展开讨论,分析同一网络在不同数据集上的性能表现.

(3) 模型评价标准既涉及高精度指标,又兼顾了轻量化和模型可解释性.

表1 糖尿病视网膜病变分级标准

类别	DR病变等级	临床表现
0	NO-DR	无病变特征
1	Mild NPDR	仅出现微动脉瘤
2	Moderate NPDR	除了微动脉瘤外,出现血点絮状静脉串珠,病变程度介于轻度和重度之间
3	Severe NPDR	大于两个象限出现静脉串珠,四象限内,每个象限出现20个以上的出血点,至少一个象限出现微血管异常
4	PDR	玻璃体/视网膜出血,增生新血管

DR诊断手段主要包括两种技术:检测和分级.DR检测技术采用二元分类(DR或无DR),DR分级即多元分类,运用深度学习技术将DR严重程度划分为5个等级.目前,DR的二元检测技术已经相当成熟,且取得了不错的成果,已经能够成功应用于临床实践中.然而,由于DR病变的复杂性,DR的多元分级仍然存在着各种困难与挑战.因此,本文主要总结介绍DR分级的相关研究,首先整理了DR分级研究的常用数据集;其次介绍了基于深度学习的DR分级诊断方法;将其按照经典分类网络模型的改进工作依次阐述;然后介绍基于多网络融合的DR诊断方法;并且总结比较不同改进方法的优缺点;最后结合深度学习在DR分

级诊断领域所面临的挑战,对未来的研究方向进行展望.

1 DR分级的数据集

深度学习在DR分级上的成功应用在很大程度上取决于所使用的数据集,所用数据集的质量和其注释的精度将直接影响DR分级结果的好坏.通过DR分级的相关期刊会议报告和相关开源数据网站对近年来有关DR分级的公开数据集进行总结和介绍,表2总结了DR分级的常用数据集.

表2 糖尿病病变(DR)分级常用数据集

数据集	格式	图像数量	病变等级
APTOS 2019	PNG	3 662	5
EyePACS	JPEG	88 702	5
IDRiD	JPEG	597	5
MESSIDOR	TIFF	1 200	4
MESSIDOR-2	TIFF	1 200	4
DDR	JPEG	13 673	5

1.1 APTOS 2019

APTOS 2019^[8]包含从印度农村许多参与者收集的3 662个样本,由印度Aravind眼科医院组织建立了此数据集.后来,让一组临床经验丰富的医生根据国际临床糖尿病视网膜病变严重程度表对样本进行了审查和标记,并标记成5个等级.此数据集只有训练集是公开的,且数据集高度不平衡,其中1 805图像是无病变特征的图像,只有183张图像具有严重非增殖性病变特征^[9].

1.2 EyePACS

EyePACS^[10]是由加利福尼亚州和其他地方的多个初级保健点收集的,由于各个站点的采集设备和采集环境不同,造成此数据集的分辨率大小不一,且部分数据存在瑕疵、失焦、曝光不足、曝光过度等问题.该数据集包括35 126张训练样本和53 576张测试样本,共计88 702张图像.该数据集来源亚洲、欧洲、非洲、美洲土著等多个地区,覆盖范围广,囊括的种族多样,这也是该数据集被广泛使用的重要原因.

1.3 IDRiD

IDRiD^[11]来源于IEEE国际生物医学影像研讨会(ISBI-2018)举办的“糖尿病性视网膜病变:分割和分级挑战”竞赛,包括597张彩色眼底图像组成,像素值为4288×2848,存储格式为JPEG.此数据集由3部分组成:语义标注的81张眼底图像;第2部分是516张通过CSV文件记录的病变等级;第3部分通过CSV文

件记录了每张图像中视盘和中间凹的位置信息。

1.4 MESSIDOR

MESSIDOR^[12]是由 MESSIDOR 项目创建,用于评估视网膜病变的分割和分级,它是一个完全标记的数据集,共包含 1200 张 TIFF 格式的眼底图像,具有多规格的分辨率,包括 1400×960、2240×1488、2304×1536,并且每张图像都由专业医生进行了标记。但是,不同数据在图像质量、图像尺寸和视觉效果等方面存在着显著性差异,因此,选用该数据集作评价数据集时,数据预处理的方法可能对实验的性能指标影响较大。

1.5 MESSIDOR-2

MESSIDOR-2^[13]是 MESSIDOR 的继承与发展,包含 1200 张高质量图像,其中一部分数据继承了 MESSIDOR^[10],另一部分由法国布雷斯特大学医院提供,并且由 3 名临床经验丰富的专家根据国际临床糖尿病视网

膜病变严重程度表进行标定。

1.6 DDR

DDR 数据集^[14]包含来自中国 13 个省份的 673 家医院的 14723 张眼底图像。根据 DR 严重程度,图像分为 5 类:无、轻度、中度、重度和增殖性 DR。第 6 类表示图像质量差。对外提供的数据集(共 13673)不包括质量差的图像,并且所有图像都做了预处理来删除黑色背景。

2 深度学习在 DR 分级诊断中的应用

深度学习作为机器学习的重要分支,通过端到端的学习方式从大量的训练样本中自动提取视网膜图像的抽象特征信息,从而获得高精度的 DR 分级诊断模型。本节将深度学习在 DR 分级诊断中的应用方法按照经典分类模型分为 6 类,具体方法见图 3。

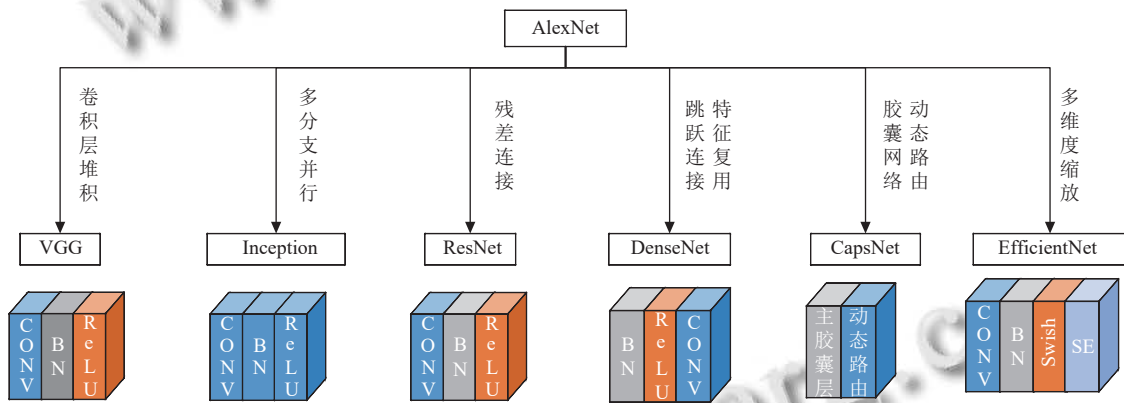


图 3 基于深度学习分级方法的结构图

卷积神经网络 (convolutional neural network, CNN) 是一类包含卷积计算且具有深度结构的前馈神经网络,由输入层、卷积层、池化层、全连接层、输出层组成^[15], DR 分级诊断的基本网络架构如图 4 所示。CNN 是深度学习中的一种监督网络,通常用以处理图像数据,由于 DR 数据集的不断更新完善和图像处理器的发展, CNN 已经成为眼科医学图像分析的主导网络架构^[16]。CNN 及其衍生模型在 DR 分级领域具有极大的应用场景,为 DR 的分级任务提供了可靠且稳定的技术手段。

2.1 基于 VGG 网络的 DR 分级模型

VGGNet^[17]是牛津大学视觉几何小组 (visual geometry group, VGG) 所提出来的,包含 VGG-16 和 VGG-19 两种结构,两者并没有本质区别,只是网络深度不

同, VGG 的提出证明了增加网络深度能够影响网络的性能。VGGNet 通常采用 3×3 卷积核、2×2 池化核,增加网络的深度来提高模型的特征提取能力和模型表达能力,能够在保证相同感受野的同时,减少网络参数且捕获更多细节特征。因此,一些研究者将 VGGNet 作为 DR 分级的骨干分类网络。

Chandrasekaran 等^[18]首先针对 EyePACS 数据集^[10]进行非局部均值去噪,然后使用 VGG-16 和 VGG-19 并结合迁移学习^[19]和超参数调优的方法对 DR 严重程度进行自动分级,其最佳准确率达 95.68%,最优 AUC 为 0.978 6。但其特异性表现一般,并且该网络使用了大量的池化层从而丢失了识别对象的空间信息,为了避免上述问题可以引入胶囊网络和动态路由算法。

Bhardwaj 等^[20]采用 MESSIDOR^[12]作为训练集,并用 IDRiD^[11]中的部分眼底图像来进行性能验证,其提出的 PFTL 方法^[20]应用于 VGG-19 来进行特征提取,并用统计特征选择块与支持向量机分类器相结合的方法来进行分类,PFTL 搭配 VGG-19 在 IDRiD^[11]上性能表现出色,准确率达 88.9%、AUC 曲线下面积为 0.89、特异性为 89.86%。但是,其交叉损失熵为 0.647,相较于 PFTL 搭配 InceptionNet 下的 0.295,PFTL 搭配 VGG-19 的交叉损失熵表现过高。Geetha 等^[21]针对

EyePACS^[10]的类别样本数量不均衡问题,提出通过图像反射从特定级别的现有眼底图像样本生成额外样本来解决,并使用对比度有限自适应直方图均衡 (CLAHE)^[22]和高斯模糊来改善图像的质量。此外,Geetha 等^[21]提出了基于 VGG-16 和空间注意力增强^[23]的 DR 自动分级模型,准确率为 89.20%。该网络通过 VGG-16 来提取眼底图像中的 ROI (region of interest)^[24],而注意力模块负责捕获,而且 VGG-16 的第 4 个池化层有助于构建模型管道,该模型管道不仅性能表现良好而且效率高。

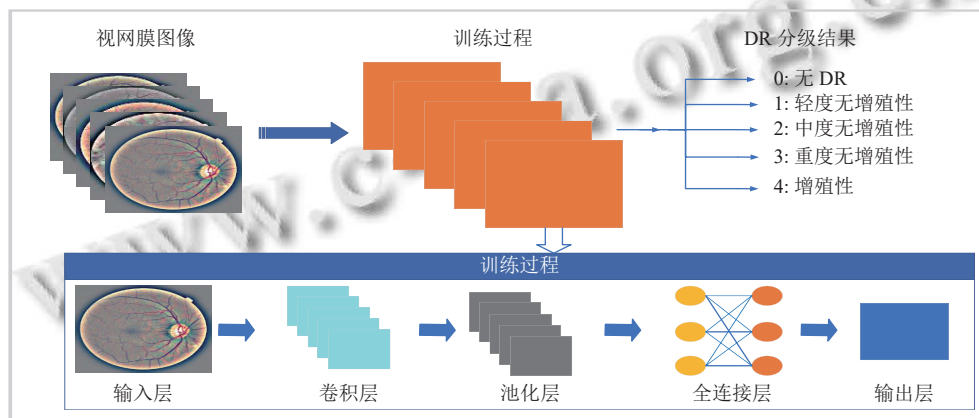


图4 DR 分级诊断过程及其网络架构图

Khan 等^[25]设计了用于 DR 分级的 VGG-NIN 网络,同时引入空间金字塔池化层 (SPP)^[26],形成一个高度非线性且尺度不变的深度模型,该模型在识别每个 DR 阶段时模型参数量小,从而加速了模型的训练和收敛,且 AUC 曲线下面积为 0.95,但是其数据预处理方法单一。针对此问题,Beevi 等^[27]在 DR 分级之前先对图像进行显著特征挖掘,使用 U-Net 对眼底图像进行病变分割,然后 VGG-16 结合迁移学习进行 DR 严重程度分级,DCNN^[27]相较于 VGG-16-NIN^[25],在缩短计算时间的同时也降低了内存的开销。

总体来看,VGG 网络作为骨干网络的模型大部分都是针对大数据集,进行模型微调,或是使用迁移学习,然而 VGG 网络对于小样本数据集 IDRiD^[11]和 MESSIDOR^[12]的性能表现并不是十分优秀,对比结果可以发现,其中 EyePACS 数据集^[10]更适用于 VGG 网络架构,同时也证明了 VGG 这种深层网络结构在平衡的大数据集上表现更为优秀。

2.2 基于 InceptionNet 的 DR 分级模型

相对于 VGGNet 的多层卷积堆积,InceptionNet 更

注重提取多尺度特征,通过采用不同尺度的多分支结构来聚集相关性强的特征,从而保证了网络结构的稀疏性,在模型轻量化的角度,采用更小尺度的卷积核,还利用 1×1 卷积核减少参数量并加深网络,从而提高了网络的计算效率和性能。此外,模型使用全局平均池化层代替全连接层,进一步提高了运算速度。因此,InceptionNet 也常常被用于 DR 分级任务。

Li 等^[28]提出以 Inception-V3 模型作为基础架构来进行迁移学习^[19],他们用在 ImageNet 数据集上训练的模型的权重来初始化模型的权重,数据集选用上海中山医院收集的私人数据集和公开数据集 MESSIDOR-2^[13],按照 9:1 的比例划分成训练集和验证集,并进行了 10 倍的交叉验证,最后实现 93.49% 的准确率、93.45% 的特异性、96.93% 的召回率。但是该模型仍存有改进的空间,由于 DR 的各种特征在提取和分类方面也很复杂,单一模型在单独使用时无法发挥其最佳性能,尤其针对大数据集,可以将其划分成多个小数据集,同时结合集成学习算法,分类精度会进一步提高。Devi 等^[29]使用 MESSIDOR^[12]和 DDR^[14]对 Eye-PACS^[10]进行扩

充,使用 InceptionNet 模型作为特征提取器,同时集成了 SVM、KNN 和 Logistic 回归等经典机器学习算法对图像进行 DR 分级,在测试集上的准确率为 95.4%。该模型用 InceptionNet 对 DR 进行特征提取,再把这些特征传递给 3 个经典分类器,通过最大投票算法来给出分级结果。虽然最终取得不错的效果,但是从该研究的混淆矩阵可以看出,假阳性的最大数量属于 0 标签(无 DR),这意味着存在着异常且大量的标记为 0 的图像,模型存在过度拟合的问题。

Aljehane^[30]提出一种基于 Inception 架构的 DR 分级模型(IMFO-INDR),将 Inception-V4 模型用于提取特征,并通过飞蛾火焰优化算法(MFO)^[31]对其中涉及的超参数进行优化调整,使用 MESSIDOR^[12]对 IMFO-INDR 模型进行实验验证,准确率达 99%,特异性为 99.1%。虽然 IMFO-INDR 模型对小数据集的性能表现非常出色,但是对于大数据集的性能表现有待研究,可以引入对比实验,创建新的验证集进行验证,模型存在过度拟合。

Bhardwaj 等^[32]针对 DR 分级存在的小病变检测困难的问题,提出了一种基于象限的 QIV-3 模型,该模型使用 Inception-V3 模型作为基础架构,将 4 个 Inception-V3 模型并行组合,以获得更多的自由参数并产生更好的精度,使用 4 个不同象限的眼底图像的裁剪提供对这 4 个模型的输入,提高了对细节特征的提取能力,最终在 MESSIDOR^[12]的性能表现优秀,准确率为 93.33%、交叉熵损失为 0.291。但是,该模型在着重关注提取细节特征的同时意味着需要消耗更多的计算成本和时间成本,尽管该模型在训练时引入了迁移学习和 4 个模型并行训练的方法来加快模型的收敛,但 4 倍的参数量增加了模型的内存开销。

综合来看,以 InceptionNet 为骨干的模型在降低模型参数量这一方面具有一定的优势,其中 QIV-3^[32]模型就是利用了 InceptionNet 的这一特点。此外,IMFO-INDR^[30]模型虽然取得优良的性能,但是作者并没有用其他数据集做验证对比实验,模型的泛化能力有待验证。

2.3 基于 ResNet 的 DR 分级模型

ResNet^[33]研究团队发现了“退化现象”,并提出了“残差连接”的思想来解决这一现象,ResNet 将浅层特征信息直接连接到深层,利用新的表达方式实现层间的恒等映射。ResNet 的每层特征仅需要在前续层特征的基础上学习新的特征,残差连接中反向传播的导数

需要加上恒等项,这样缓解了网络退化和梯度弥散问题。残差技术的出现推动了深层网络模型的发展,深层网络具有更好的逐层学习能力和更强的模型表达能力。因此,基于深度学习的 DR 分级方法更多以 ResNet 作为基础网络架构。

ResNet 网络可以结合迁移学习的策略用来 DR 的分级任务,Sallam 等^[34]提出了一种以 ResNet34 为骨干网络的高效 CNN 分级模型,使用其在 ImageNet 数据集上的训练权重来初始化模型权重,在 APTOS 2019^[8]进行训练,最后在 IDRiD^[11]上进行测试,其特异性为 98.5%,平均准确度为 94.9%。迁移学习多用于针对小样本的数据集^[35],Elsawah 等^[36]设计了一个基于 ResNet50 的 DR 分级架构,该架构由 3 个流水线阶段组成:数据扩充和数据平衡的预处理阶段、ResNet50 模型结合迁移学习的特征提取阶段、SVM 分类器的分类阶段,选用小样本数据集 IDRiD^[11]作为实验数据集,总体分类准确率为 93.3%,但其他性能指标并没有被提及。虽然迁移学习更适用于小样本数据集,但通过对比文献^[34]和文献^[36]的性能表现,发现选用小样本数据集 IDRiD^[11]作为训练集限制了 ResNet 作为一种深层网络的自身优势。

ResNet 模型结合注意力机制也被广泛应用于 DR 分级任务,SEA-Net^[37]以 ResNet 为基础架构来提取特征,然后引入了具有 1×1 卷积层和池化层序列的空间注意力模块来进行降维,为了兼顾学习特征的通道关系,又引入了 SE 模块来重新校准通道特征图,用于 DR 细粒度分级,该模型在 EyePACS^[10]上有着出色的性能表现。为了捕捉 DR 类别内之间的细微病变,Tian 等^[38]设计了一个细粒度注意力网络(FA+FC+Net),骨干网络 ResNet50 在 ImageNet 上预训练后来进行特征提取,改进的注意力 FA 模块能够有效区分不同类别之间具有重要意义的细微特征,并清除冗余信息。KC 模块选用 YOLO-V5 作为分级器,它能够快速捕获多个细节病变,同时 KC 模块结合了专业的医学先验知识,经过广泛的实验验证,该模型在多个数据集的上性能表现都十分出色。综合文献^[37,38]来看,ResNet 结合注意力机制的模型善于捕捉 DR 的细节特征,更容易区分 DR 各类别间的细微差距,提高了 DR 的分级精度,但是在模型轻量化方面表现出了一定的局限性,尤其是 SEA-Net^[37]模型在着重关注提升精度的同时没有兼顾模型的体量。

针对糖尿病视网膜病变的复杂性,基于多尺度融合的思想, Li 等^[39]基于 ResNet50 设计了一个损失注意金字塔网络 (LAPN) 来整合不同分辨率的眼底图像, LAPN 为不同分辨率的图像建立了独立的网络路径,同时也设计了一种融合图像特征和病变激活图的病变注意力模块 (LAM), 在 EyePACS^[10]和 MESSIDOR^[12]上实验证明,该方法在 DR 分级方面有着出色的性能表现. 其中, LAPN 提供的独立网络路径,在低分辨率网络中,使用弱监督定位方法获得病变的位置信息,同时引导高分辨网络集中在病变区域,并且注意力模块 (LAM) 构建低分辨率特征和高分辨率特征之间互补关系,在模型可解释性方面,该方法可以提供良好的可解释病变区域,并且这些区域在多级网络中具有病变一致性. 此外,针对数据集的不均衡问题,郑雯等^[40]基于 ResNet50 模型设计了一种 Improved-DR-Net,该模型将普通卷积层替换为空洞卷积层,同时融入了注意力机制,在 APTOS 2019^[8]和 EyePACS^[10]上平均准确率为 98.04%. 综合文献^[39,40]来看,该类方法适用于图像尺寸不统一和类别样本数量不均衡的 DR 数据集.

为了研究 ResNet50 的各个关键组件对其性能的影响, Huang 等^[41]以输入分辨率、目标函数、学习率、数据预处理、先验知识、集成方法作为关键变量,利用控制变量的方法,将这些关键变量的不同组合在 EyePACS^[10]上进行了广泛实验,来评估每个变量对模型性能的影响,最终选出各个关键组件的最佳组合,取得较好的性能效果 ($0.81 < \kappa = 0.8631 < 1$). 该研究的特点是没有从模型的架构改进出发,作者另辟蹊径,在模型本身并没有任何令人眼花缭乱的改进,而是着手于基础组件的优化调整,但是最终取得了意想不到的性能.

ResNet 网络及其衍生模型具有减缓梯度消失和网络退化的特点,因此 ResNet 网络更能在较大规模的 DR 数据集上发挥出自身的优势,但是该类模型在轻量化方面发展有一定局限性. 此外,受文献^[41]的启发,对于 DR 分级的模型改进工作,回归到框架的基础组件上未免不是一种提升模型性能的新思路.

2.4 基于 DenseNet 的 DR 分级模型

DenseNet^[42]是将 ResNet 的跨层连接的思路进一步扩展应用,但相比 ResNet, DenseNet 提出了一种更激进的密集连接机制,即互相连接所有的层,实现了各层特征图的特征复用,同时加强网络各层之间的特征传递并降低邻层之间的依赖性,其中网络每层输入的

特征图由前续所有层的信息组合而成,不同长短的跨层连接提供紧凑且有区别的特征,有效地缓解了深层网络中因梯度消失而难以优化的问题, DenseNet 还利用全部特征进行最终预测来提升模型鲁棒性^[43]. 这些特点让 DenseNet 在参数和计算成本更少的情况下实现比 ResNet 更好的性能,因此近几年涌现出许多基于 DenseNet 的 DR 分级模型.

Rahman 等^[44]使用 DenseNet-169 在 ImageNet 上预训练,使用预训练的权重初始化模型的权重,并且引入预训练的 DenseNet-121 和 ResNet50 做对比实验,实现结果表明, DenseNet-169 在 APTOS 2019^[8]上测试的分类准确率为 96.54%,灵敏度和特异性分别为 96.23%、99.21%,优于另外两个网络的性能. AbdelMaksoud 等^[45]结合迁移学习的方法,提出了基于 EyeNet 和 DenseNet-101 的混合网络 (E-DenseNet),预处理阶段选用 APTOS 2019^[8]、EyePACS^[10]、IDRiD^[11]、MESSIDOR^[12]这 4 个 DR 数据集对模型进行训练,并且 4 个数据集作为独立验证集对模型的性能进行验证,平均准确率为 91.35%、平均敏感性为 96%、特异性为 69%. 综合文献^[44,45]来看,文献^[44]针对数据集的各类别样本的不均衡问题,作者用重采样的方法来扩充数据集,然而并没有引入其他 DR 分级数据集进行实验,模型的扩展性和可移植性并没有得到证明,从而不能真正应用到临床实践中. 对于文献^[45], E-DenseNet 模型对不同风格的眼底图像都具有不错的分级能力,且不限数据集的规模,具有很强的可移植性,符合 DR 临床诊断的实际情况 (患者的眼底图像的采集环境存在差异),但是该模型的平均特异性过低,也就意味着容易出现假阳性的情况 (误诊率偏高).

张彤等^[46]利用二阶马尔代夫的思想来简化 DenseNet-121,从而设计出一个 2-DenseNet 网络,并且作者借鉴了残差网络中的 Shortcut 连接思想,将全局池化层和第 1 个卷积模块的输出作为第 2 个卷积模块的输入,第 2 个卷积模块的输出用于调整原始网络的输出,同时嵌入了注意力模块,在 EyePACS^[10]上与其他经典网络架构进行对比实验验证,准确率为 97.2%、灵敏度为 92.1%、特异性为 96.7%. 2-DenseNet 在实现了特征复用和缓解梯度退化的同时,还减少了参数量,从而极大减少了运算量,加快了模型的收敛速度,并且该网络注重细粒特征的提取. 但是,该网络的参数量减少是相对于 DenseNet 而言,对比其他经典轻量化架构, 2-DenseNet

参数开销仍然较大。

总体而言,基于 DenseNet 的 DR 分级模型的改进思路大多是在保证 DenseNet 本身优势的前提下,去降低网络的参数量,2-DenseNet 的参数量相比 DenseNet 降低了 30%。但是,该类模型在模型复现方面仍然存在困难,轻量化应用是该类模型的发展方向。

2.5 基于 CapsNet 的 DR 分级模型

针对 CNN 的缺陷, CapsNet^[47]采用动态路由算法,取代传统卷积神经网络的最大池化法,避免图像因池化导致关键特征信息的丢失,并且它使用浅层网络特征,让 Capsule 层嵌套在其他层中,从而解决了只能增加层数来加深网络深度的问题^[48]。相比于 CNN, Capsule 网络不需要大规模数据集的训练,能够在多个角度识别物体,在 DR 分级中展现了良好的性能。

Kumar 等^[49]提出了一种改良的胶囊网络,使用卷积层和主胶囊层从眼底图像中提取特征,然后用类别囊层和 Softmax 层估计图像属于特种类别的概率,选用 MESSIDOR^[12]的 806 张眼底图像进行训练,90 张用作验证,304 张用于验证网络的性能,平均准确率为 97.98%,召回率为 96.11%。可见, Capsule 网络可以在小规模数据集展现了其强大的学习能力,但是动态算法内部的迭代时耗长,迭代次数多,降低了识别效率,尽管准确率优于其他网络,但识别速度仍有很大的提升空间。

为了克服胶囊网络的局限性, Kalyani 等^[50]以 CapsNet 为基础架构,将最后一层(全连接层)替换成支持向量机(SVM),选用 APTOS 2019^[8]中的 90% 作为训练集,10% 用于测试。最终取得 96% 的平均准确率、94% 的特异性,但敏感性仅为 79%。可见, CapsNet 与 SVM 相结合的集成方法拥有在不均衡数据集上进行有效分类的能力,并且 SVM 模块通过避免所有层中的权重重新计算而加快了模型的收敛。但是,敏感性表现较低,临床应用中容易出现漏诊的情况。

对比文献^[49,50]发现,基于 CapsNet 的 DR 分级模型在小数据集上的性能表现更好,未来的改进工作应更倾向于提高动态路由算法的效率。

2.6 基于 EfficientNet 的 DR 分级模型

EfficientNet^[51]提供了一种多维缩放方法,交叉使用 3×3 和 5×5 的卷积核进行特征提取,提出了组合缩放系数,可同时对宽度、深度和图片分辨率进行参数调整^[52],同时兼顾了精度和速度。以 EfficientNet 为骨干网络的 DR 分级模型在分级准确度上取得了优异成果。

Chetoui 等^[53]提出了一种 RA-EfficientNet 网络,在 EfficientNet 中添加了残差注意力块,从而提取了更多的特征,解决了病变之间差异很小的问题。在 APTOS 2019^[8]和 EyePACS^[10]上测试精度分别是:93.55%、89.29%,并且参数量方面相比于 EfficientNet 仅高出 5.2%,综合精度提升来说,增长 5.2% 的参数量是可以接受的。Zhao 等^[54]在 EfficientNet 中嵌入注意力模块来抑制无关信息,并且引入生成对抗网络^[55]进行数据增强和扩充,该方法在 APTOS 2019^[8]和 MESSIDOR-2^[13]上的平均准确率分别为:92.25%、87.18%。综合文献^[53,54]来看,基于 EfficientNet 的 DR 分级模型在不同数据集上性能差异较大。

王明智等^[56]提出基于 EfficientNet-B3 的迁移学习模型,并引入代价敏感正则化,能够使模型在多分类任务中惩罚不同类型的错误,从而更容易区别不同类别之间的差异,在 EyePACS^[10]和 MESSIDOR-2^[13]上得到 0.86 和 0.88 的二次加权 kappa 分数。Naveenkumar 等^[57]将 EfficientNet-B3 和 Ben 模块相结合,并与 EfficientNet-B4 和 EfficientNet-B5 做对比实验, EfficientNet-B3 结合 Ben 的性能表现优于另外两种方法,在汕头国际眼科中心的数据集上得到 96.8% 的平均准确率。虽说该方法在精度方面有其独特的优势,但在模型的可移植性方面具有一定局限性。

总体来看,基于 EfficientNet 的 DR 分级模型缺乏可重复性,对数据集的质量要求颇高,因此提高模型的泛化能力应该着手于多样化数据集的开发。

2.7 小结

上述 6 类经典分类网络经过优化改进,通过融合 DR 图像中各个层次和各个区域的特征信息,不断提高 DR 分级模型的分级精度。但是深层神经网络的特征提取依赖于大规模 DR 数据集,当前的大多数 DR 数据集可解释性差,导致 DR 分级模型难以广泛地推广到临床应用当中,因此未来的研究重点应放在提高模型的解释性上。表 3 总结归纳了各个 DR 分级模型的性能及优缺点。

3 多融合网络在 DR 分级中的应用

多融合网络^[58]是指通过联立两个及以上的网络架构,实现对 DR 分级的深度学习方法。按照联立形式可分为多网络级联和多网络集成。多网络融合能结合每个网络的优势有针对性地用于 DR 分级任务。

表3 经典分类网络用于DR分级的汇总表

基础网络	改进方法及机制	数据集	性能指标	优势	局限性	轻量化
VGG	VGG-16、VGG-19、VGG-S +迁移学习+数据扩充 ^[18]	EyePACS	ACC=95.68%、 SP=97.43%、 SE=86.47%、 AUC=0.9786	计算量较低	使用大量池化层丢失了大量空间信息	否
	VGG-19改进、SVM、统计特征选择快+PFTL ^[20]	MESSIDOR IDRiD	ACC=88.9%、 SP=89.86%、 SE=89.89%、 AUC=0.89	模型有较强的泛化的能力	交叉损失熵较大	否
	VGG-16改进、注意力模块+高斯模糊+CLAHE+数据扩充 ^[21]	EyePACS	ACC=89.2%	模型的倾向性低,提高了计算效率	模型的泛化能力低	否
	VGG-NIN、SPP ^[25]	EyePACS	ACC=83.8%、 AUC=0.95	多尺度特征融合	对0类的倾向性强	否
	VGG-16改进、U-Net+迁移学习 ^[27]	DRIVE	ACC=91.72%、 Loss=0.2195	避免了小数据集带来的过拟合问题	小数据集限制了模型的实验效果	否
Inception	Inception-V3改进 +迁移学习+CLAHE ^[28]	私人数据集1 MESSIDOR-2	ACC=93.49%、 SP=93.45%、 SE=96.93%、 AUC=0.9905	数据集种类丰富,节省内存	对图片质量要求较为严格	否
	Inception改进+SVM&KNN &Logisti+最大投票 ^[29]	MESSIDOR EyePACS DDR	ACC=95.4%、 RECALL=95.4%、 F1 score=95.5%	数据集种类丰富,模型的可移植性强	假阳性的最大数量属于0标签,模型面临过度拟合	否
	IMFOI-INDR ^[30]	MESSIDOR	ACC=99%、SP=99.1%	参数量小,易部署	模型难复现	是
	Inception-V3+基于象限法 +迁移学习 ^[32]	MESSIDOR	ACC=93.3%、 交叉损失熵为0.291	解决了小数据集引起的过拟合问题	参数量大导致内存开销大	否
ResNet34	ResNet34改进+迁移学习 ^[34]	APTOS 2019 IDRiD	ACC=94.9%、 SP=98.5%、 SE=95%、 F1 score=93.2%	高灵敏度、高特异性	要求高分辨率图像	否
	ResNet50+SVM+迁移学习 ^[36]	IDRiD	ACC=93.3%	—	只提及准确率,其他指标没有给出	否
	ResNet50+注意力机制 ^[37]	EyePACS	AUC=0.876	注重细粒度特征	所给性能指标单一	否
	ResNet50+注意力机制 +YOLO-V5 ^[38]	APTOS 2019 MESSIDOR DDR	APTOS 2019: ACC=99.04%; MESSIDOR: ACC=94.1%; DDR: ACC=86.17%	模型的可移植性强	模型体量大,参数量大,不易部署	否
	ResNet50+损失注意力金字塔+病变注意力 ^[39]	MESSIDOR EyePACS	ACC=93.2%、 kappa=0.891、 AUC=0.97	模型的可移植性强	模型体量大,参数量大,不易部署	否
	ResNet50改进+注意力机制 +迁移学习 ^[40]	APTOS 2019 EyePACS	ACC=98.04%、 SP=98.03%、 SE=97.91%	模型的可移植性强	模型体量大,参数量大,不易部署	否
	控制变量法分析ResNet50的各个组件 ^[41]	EyePACS	kappa=0.8631	研究架构的基础组件对性能影响	只针对DR数据集,不易复现	是
DenseNet	DenseNet-121/DenseNet-169 +迁移学习+数据扩充 ^[44]	APTOS 2019	ACC=96.54%、 SP=99.21%、 SE=96.23%	—	模型的泛化能力弱	否
	DenseNet-101+EyeNet +迁移学习 ^[45]	MESSIDOR EyePACS DDR APTOS 2019	ACC=91.35%、 SP=69%、 SE=96%	不要求数据集的规格和图像质量	平均特异性较低,容易出现假阳性(误诊)	否

表3(续) 经典分类网络用于DR分级的汇总表

基础网络	改进方法及机制	数据集	性能指标	优势	局限性	轻量化
	马尔代夫思想简化DenseNet-121+注意力机制+数据扩充 ^[46]	EyePACS	ACC=97.2%、 SP=96.7%、 SE=92.1%	参数量相对较小, 运行时间相对更快	实验数据单一	是
CapsNet	改进CapsNet+Softmax+CLAHE ^[49]	MESSIDOR	ACC=97.98%、 RECALL=96.7%、 F1 score=97.19%	在有限小数据集上能进行有效的高性能分类	处理图像耗时, 且对图像的规模和质量有要求	否
	改进CapsNet+SVM ^[50]	APTOS 2019	ACC=96%、 SP=94%、 SE=79%	高准确率、高特异性	灵敏度低, 容易漏诊的情况	否
EfficientNet	RA-EfficientNet+迁移学习 ^[53]	EyePACS APTOS 2019	EyePACS: ACC=89.29%; APTOS 2019: ACC=93.55%	计算量相对较低, 有很好的稳健性	在不同数据集上的性能差异较大	否
	改进EfficientNet+注意力+生成对抗网络扩充数据 ^[54]	APTOS 2019 MESSIDOR-2	APTOS 2019: ACC=92.25%; MESSIDOR-2: ACC=87.18%	有很好的稳健性	在不同数据集上的性能差异较大	否
	改进EfficientNet-B3+迁移学习+代价敏感正则化 ^[56]	APTOS 2019 MESSIDOR-2	APTOS 2019: kappa=0.86; MESSIDOR-2: kappa=0.88	综合性能好	模型可解释性差	否
	EfficientNet-B3+BN ^[57]	私人数据集2	ACC=96.8%、 RECALL=93.4%、 F1 score=92%	小模型, 易部署	实验数据集单一	是

注: “—”代表没有显著性优缺点; “ACC”代表准确率; “SE”代表灵敏度; “SP”代表特异性; “RECALL”代表召回率; “私人数据集1”来自上海中山医院; “私人数据集2”来自汕头国际眼科中心

3.1 多网络级联

Khalid 等^[59]用 VGG-16、DenseNet-121 和 ResNet50 分别对眼底图像进行特征提取, 引入最大投票算法来合并各个网络的预测概率, 在 DDR^[14]上的平均准确率为 81.02%、F1 得分为 79.83%。该模型的特点是不同模型的优缺点可以互相弥补, 但是模型的过量堆积必将导致参数量的增加, 增加了计算成本和内存开销。

Gour 等^[60]将 CapsNet 和 Xception 串联在一起, 提出了一种名为 XCapsNet 的模型, 同时结合迁移学习方法和 CLAHE 预处理技术, 该模型在 APTOS 2019^[8]和 MESSIDOR^[12]上对 DR 图像的二元分类和多元分类分别实现了 98.91% 和 83.06% 的准确率。该模型仅在 DR 病变的检测方面有着很好的性能表现, 对于 DR 严重程度的分级仍有很大的提升空间。

Jian 等^[61]提出了一种三级联网络 (Triple-DRNet), 其中 DR-NET 主要负责有无 DR 的二元检测, PDR-Net 负责将视网膜血管分割图作为医学先验知识关联到模型中, NPDR-Net 执行多元分级, 并引入了两个注意力模块, 在 APTOS 2019^[8]上的平均准确率为 92.08%,

AUC 曲线面积为 0.990 1。该模型将分级任务划为 3 个阶段, 从而提高了对细粒病变特征的捕捉能力, 但 NPDR-Net 的分级性能仍有提升空间, 并且网络的参数量仍然可以进一步压缩。

3.2 多网络集成

为了提高诊断效率, Qian 等^[62]采用 ResNet 和 DenseNet 设计出 AD2Net, 同时引入注意力机制, 在 EyePACS^[10]上取得 83.2% 的平均准确率, kappa 系数为 0.8。虽然作者进行了数据集的扩充, 扩充方式较为单一, 该模型仍然表现出了倾向性, 导致只有正常类的准确率超过 90%, 其他 4 类的准确率均低于 80%。

Bhardwaj^[63]等基于 Inception 和 ResNet 设计一种集成网络 QEIRV-2, 继承并扩展了文献^[32]的思路, 采用 4 个 QEIRV-2 并行的方式, 且结合迁移学习、象限裁剪、直方图均衡、光盘定位等方法, 在 MESSIDOR^[12]上进行训练, 并选用 IDRiD^[11]来做性能验证, 平均准确率为 93.33%。该模型相比 QIV-3^[32], 准确率并没有明显提升, 但 QEIRV-2 的泛化能力得到了提高。

Yasin 等^[64]提出一个基于 Inception 和 ResNet 的

混合模型,且该模型采用了多网络级联和多网络集成两种融合方式,并结合迁移学习的方法,在 IDRiD^[11]上的准确率为 96.73%,且模型的计算量小,但并没有提

及其他性能指标,模型的泛化能力也没有引入对照实验进行验证。

多网络融合 DR 分级模型的性能对比见表 4。

表 4 基于多网络融合的 DR 分级方法对比表

融合方式	基础网络	改进方法机制	数据集	性能指标	优点	局限性
级联	VGG-16、ResNet50 DenseNet-121	三网络独立进行特征提取,最大投票算法将这些特征耦合	DDR	ACC=81.02%、 RECALL=81.02%、 F1 score=0.7983	—	参数量大,内存开销大
	CapsNet、Xception	迁移学习+CLAHE预处理	APTOS 2019 MESSIDOR	ACC=83.06%	模型不限制数据集的规模大小	准确率有很大的提升空间
	ResNet、PDR-Net NPDR-Net	三网络级联+两个高效的注意力模块+高斯模糊	APTOS 2019	ACC=92.08%、 AUC=0.9901	有很强的细粒特征提取能力	参数量大
集成	ResNet、DenseNet	多网络集成+高效注意力模块	EyePCAS	ACC=83.2%、 kappa=0.8	特征复用,参数量相对较小	轻微过拟合
	Inception、ResNet	四网络级联并行+迁移学习	MESSIDOR IDRiD	ACC=93.33%	模型的可移植性强	模型体量大,不易部署
	Inception、ResNet	级联和集成并存+迁移学习	IDRiD	ACC=96.73%	高准确率,低计算量	泛化能力待验证

注:“—”代表没有显著性优缺点

4 总结与展望

综上所述,基于深度学习的糖尿病视网膜病变严重程度分级研究已经取得了不错效果。本节首先对 DR 分级的研究现状进行总结,然后结合研究现状对当前面临的挑战和展望进行阐述。

4.1 总结

分类模型的改进思路主要有两个方向,一是高精度方向;二是模型轻量化方向。然而近年来的 DR 分级模型大部分集中在高精度方向,不管是全监督学习方法中的增加感受野、多特征融合、多模型融合,还是半监督学习方法中的迁移学习、生成对抗网络,其主要目的都在追求高精度指标,虽然当前的 DR 分级研究在高精度已经取得了不错的成果,但对模型轻量化的研究相对欠缺。

模型性能很大程度上受制于数据集,当前的 DR 数据集由于来源不同、采集环境和采集设备不同、标记医生的专业能力不同,DR 数据集大多存在图像标签模糊、图像尺寸多规格、各类别间的不平衡等问题,针对这些问题,如果没有妥善的数据预处理方案,模型的性能将受到限制。

4.2 挑战与展望

DR 分级面临的挑战如下。

(1) 目前的 DR 分级模型普遍存在模型泛化能力差

和模型体量大等问题,然而实际临床应用中不仅要求 DR 分级模型具有高精度,还要兼顾其泛化能力和轻量化。

(2) DR 数据集多为小样本,且当前高质量 DR 数据集极为稀缺,限制了 DR 分级模型的发展。

(3) 深度学习的“黑箱”性质,导致 DR 分级模型的模型可解释性差,进一步导致 DR 分级模型很难在临床应用中得到广泛推广。

为应对上述挑战,未来的 DR 分级研究应侧重于:

(1) 研究者应在保证高精度的同时,应把降低模型参数量和计算复杂度作为研究重点,并着重提高模型的泛化能力。

(2) 针对小样本问题,迁移学习和对抗生成网络是最好的应对方案。开发高质量的 DR 数据集是推动 DR 分级模型发展的最强动力。

(3) 结合可视化技术,模拟医生的临床诊断,分析每个特征对模型决策的贡献,在保证诊断准确率的同时兼顾模型的可解释性。

参考文献

- International Diabetes Federation. IDF diabetes atlas. <https://www.diabetesatlas.org>. (2021-12-6)[2023-5-10].
- Yin L, Zhang DL, Ren Q, *et al*. Prevalence and risk factors of diabetic retinopathy in diabetic patients: A community

- based cross-sectional study. *Medicine*, 2020, 99(9): e19236. [doi: [10.1097/MD.00000000000019236](https://doi.org/10.1097/MD.00000000000019236)]
- 3 Wilkinson CP, Ferris III FL, Klein RE, *et al.* Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*, 2003, 110(9): 1677–1682. [doi: [10.1016/S0161-6420\(03\)00475-5](https://doi.org/10.1016/S0161-6420(03)00475-5)]
- 4 Sebastian A, Elharrouss O, Al-Maadeed S, *et al.* A survey on deep-learning-based diabetic retinopathy classification. *Diagnostics*, 2023, 13(3): 345. [doi: [10.3390/diagnostics13030345](https://doi.org/10.3390/diagnostics13030345)]
- 5 Tsiknakis N, Theodoropoulos D, Manikis G, *et al.* Deep learning for diabetic retinopathy detection and classification based on fundus images: A review. *Computers in Biology and Medicine*, 2021, 135: 104599. [doi: [10.1016/j.combiomed.2021.104599](https://doi.org/10.1016/j.combiomed.2021.104599)]
- 6 Ismail HR, Hassan MM. Bayesian deep learning methods applied to diabetic retinopathy disease: A review. *The Indonesian Journal of Electrical Engineering and Computer Science*, 2023, 30(2): 1167–1177. [doi: [10.11591/ijeecs.v30.i2.pp1167-1177](https://doi.org/10.11591/ijeecs.v30.i2.pp1167-1177)]
- 7 范家伟, 张如如, 陆萌, 等. 深度学习方法在糖尿病视网膜病变诊断中的应用. *自动化学报*, 2021, 47(5): 985–1004. [doi: [10.16383/j.aas.c190069](https://doi.org/10.16383/j.aas.c190069)]
- 8 Maggie K, Dane S. APTOS 2019 blindness detection. Kaggle. <https://kaggle.com/competitions/aptos2019-blindness-detection>. (2019-06-28).
- 9 Alyoubi WL, Abulkhair MF, Shalash WM. Diabetic retinopathy fundus image classification and lesions localization system using deep learning. *Sensors*, 2021, 21(11): 3704. [doi: [10.3390/s21113704](https://doi.org/10.3390/s21113704)]
- 10 Cuadros J, Bresnick G. EyePACS: An adaptable telemedicine system for diabetic retinopathy screening. *Journal of Diabetes Science and Technology*, 2009, 3(3): 509–516. [doi: [10.1177/193229680900300315](https://doi.org/10.1177/193229680900300315)]
- 11 Porwal P, Pachade S, Kamble R, *et al.* Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research. *Data*, 2018, 3(3): 25. [doi: [10.3390/data3030025](https://doi.org/10.3390/data3030025)]
- 12 Decencière E, Zhang XW, Cazuguel G, *et al.* Feedback on a publicly distributed image database: The Messidor database. *Image Analysis & Stereology*, 2014, 33(3): 231–234.
- 13 Abràmoff MD, Folk JC, Han DP, *et al.* Automated analysis of retinal images for detection of referable diabetic retinopathy. *JAMA Ophthalmology*, 2013, 131(3): 351–357. [doi: [10.1001/jamaophthalmol.2013.1743](https://doi.org/10.1001/jamaophthalmol.2013.1743)]
- 14 Guo XX, Li X, Lin QF, *et al.* Joint grading of diabetic retinopathy and diabetic macular edema using an adaptive attention block and semisupervised learning. *Applied Intelligence*, 2023, 53(13): 16797–16812.
- 15 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述. *计算机学报*, 2017, 40(6): 1229–1251.
- 16 Ker J, Wang LP, Rao J, *et al.* Deep learning applications in medical image analysis. *IEEE Access*, 2018, 6: 9375–9389. [doi: [10.1109/ACCESS.2017.2788044](https://doi.org/10.1109/ACCESS.2017.2788044)]
- 17 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *Proceedings of the 3rd International Conference on Learning Representations*. San Diego: ICLR, 2015.
- 18 Chandrasekaran R, Loganathan B. Retinopathy grading with deep learning and wavelet hyper-analytic activations. *The Visual Computer*, 2023, 39(7): 2741–2756. [doi: [10.1007/s00371-022-02489-z](https://doi.org/10.1007/s00371-022-02489-z)]
- 19 Pan SJ, Tsang IW, Kwok JT, *et al.* Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks*, 2011, 22(2): 199–210. [doi: [10.1109/TNN.2010.2091281](https://doi.org/10.1109/TNN.2010.2091281)]
- 20 Bhardwaj C, Jain S, Sood M. Transfer learning based robust automatic detection system for diabetic retinopathy grading. *Neural Computing and Applications*, 2021, 33(20): 13999–14019. [doi: [10.1007/s00521-021-06042-2](https://doi.org/10.1007/s00521-021-06042-2)]
- 21 Geetha S, Parashar M, Abhishek JS, *et al.* Diabetic retinopathy grading with deep visual attention network. *International Journal of Online and Biomedical Engineering*, 2022, 18(9): 160–177.
- 22 Reza AM. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, 2004, 38(1): 35–44. [doi: [10.1023/B:VLSI.0000028532.53893.82](https://doi.org/10.1023/B:VLSI.0000028532.53893.82)]
- 23 Zhu XZ, Cheng DZ, Zhang Z, *et al.* An empirical study of spatial attention mechanisms in deep networks. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE, 2019. 6687–6696.
- 24 Meddeb M, Cagnazzo M, Pesquet-Popescu B. Region-of-interest-based rate control scheme for high-efficiency video coding. *APSIPA Transactions on Signal and Information Processing*, 2014, 3: e16.
- 25 Khan Z, Khan FG, Khan A, *et al.* Diabetic retinopathy detection using VGG-NIN a deep learning architecture. *IEEE Access*, 2021, 9: 61408–61416. [doi: [10.1109/ACCESS.2021.3074422](https://doi.org/10.1109/ACCESS.2021.3074422)]

- 26 He KM, Zhang XY, Ren SQ, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904–1916. [doi: [10.1109/TPAMI.2015.2389824](https://doi.org/10.1109/TPAMI.2015.2389824)]
- 27 Beevi SZ. Multi-level severity classification for diabetic retinopathy based on hybrid optimization enabled deep learning. *Biomedical Signal Processing and Control*, 2023, 84: 104736. [doi: [10.1016/j.bspc.2023.104736](https://doi.org/10.1016/j.bspc.2023.104736)]
- 28 Li F, Liu Z, Chen H, *et al.* Automatic detection of diabetic retinopathy in retinal fundus photographs based on deep learning algorithm. *Translational Vision Science & Technology*, 2019, 8(6): 4.
- 29 Devi TA, Yamana S, Reddy GH, *et al.* An ensemble learning approach to detect the presence of diabetic retinopathy based on Inception module. *Proceedings of the 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*. Salem: IEEE, 2023. 960–965.
- 30 Aljehane NO. An intelligent moth flame optimization with inception network for diabetic retinopathy detection and grading. *Proceedings of the 2nd International Conference on Computing and Information Technology (ICCIT)*. Tabuk: IEEE, 2022. 370–373.
- 31 Ewees AA, Sahlol AT, Amasha MA. A bio-inspired moth-flame optimization algorithm for arabic handwritten letter recognition. *Proceedings of the 2017 International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO)*. Prague: IEEE, 2017. 154–159.
- 32 Bhardwaj C, Jain S, Sood M. Diabetic retinopathy severity grading employing quadrant-based Inception-V3 convolution neural network architecture. *International Journal of Imaging Systems and Technology*, 2021, 31(2): 592–608. [doi: [10.1002/ima.22510](https://doi.org/10.1002/ima.22510)]
- 33 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 770–778.
- 34 Sallam MS, Asnawi AL, Olanrewaju RF. Diabetic retinopathy grading using ResNet convolutional neural network. *Proceedings of the 2020 IEEE Conference on Big Data and Analytics (ICBDA)*. Kota Kinabalu: IEEE, 2020. 73–78.
- 35 陈志威, 赵奎, 曹吉龙, 等. 基于迁移学习的伤口图像分割. *计算机系统应用*, 2022, 31(8): 259–264. [doi: [10.15888/j.cnki.csa.008640](https://doi.org/10.15888/j.cnki.csa.008640)]
- 36 Elswah DK, Elnakib AA, Moustafa HED. Automated diabetic retinopathy grading using ResNet. *Proceedings of the 37th National Radio Science Conference (NRSC)*. Cairo: IEEE, 2020. 248–254.
- 37 Zhao ZY, Chopra K, Zeng Z, *et al.* Sea-Net: Squeeze-and-excitation attention net for diabetic retinopathy grading. *Proceedings of the 2020 IEEE International Conference on Image Processing (ICIP)*. Abu Dhabi: IEEE, 2020. 2496–2500.
- 38 Tian M, Wang HQ, Sun YX, *et al.* Fine-grained attention & knowledge-based collaborative network for diabetic retinopathy grading. *Heliyon*, 2023, 9(7): e17217. [doi: [10.1016/j.heliyon.2023.e17217](https://doi.org/10.1016/j.heliyon.2023.e17217)]
- 39 Li X, Jiang YC, Zhang JS, *et al.* Lesion-attention pyramid network for diabetic retinopathy grading. *Artificial Intelligence in Medicine*, 2022, 126: 102259. [doi: [10.1016/j.artmed.2022.102259](https://doi.org/10.1016/j.artmed.2022.102259)]
- 40 郑雯, 沈琪浩, 任佳. 基于 Improved DR-Net 算法的糖尿病视网膜病变识别与分级. *光学学报*, 2021, 41(22): 2210002.
- 41 Huang YJ, Lin L, Cheng PJ, *et al.* Identifying the key components in ResNet-50 for diabetic retinopathy grading from fundus images: A systematic investigation. *Diagnostics*, 2023, 13(10): 1664. [doi: [10.3390/diagnostics13101664](https://doi.org/10.3390/diagnostics13101664)]
- 42 Huang G, Liu Z, van der Maaten L, *et al.* Densely connected convolutional networks. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 2261–2269.
- 43 马金林, 裘硕, 马自萍, 等. 新型冠状病毒肺炎的深度学习诊断方法综述. *计算机工程与应用*, 2022, 58(12): 51–65. [doi: [10.3778/j.issn.1002-8331.2109-0110](https://doi.org/10.3778/j.issn.1002-8331.2109-0110)]
- 44 Rahman MT, Dola A. Automated grading of diabetic retinopathy using DenseNet-169 architecture. *Proceedings of the 5th International Conference on Electrical Information and Communication Technology (EICT)*. Khulna: IEEE, 2021. 1–4.
- 45 AbdelMaksoud E, Barakat S, Elmogy M. A computer-aided diagnosis system for detecting various diabetic retinopathy grades based on a hybrid deep learning technique. *Medical & Biological Engineering & Computing*, 2022, 60(7): 2015–2038.
- 46 张彤, 孟亮. 基于注意力神经网络的糖尿病视网膜病变识别. *计算机工程与科学*, 2022, 44(3): 479–485. [doi: [10.3969/j.issn.1007-130X.2022.03.013](https://doi.org/10.3969/j.issn.1007-130X.2022.03.013)]
- 47 Choi J, Seo H, Im S, *et al.* Attention routing between capsules. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop*. Seoul: IEEE, 2019. 1981–1989.
- 48 贺文亮, 朱敏玲. 胶囊神经网络研究现状与未来的浅析. *计*

- 算机工程与应用, 2021, 57(3): 33–43. [doi: [10.3778/j.issn.1002-8331.2009-0209](https://doi.org/10.3778/j.issn.1002-8331.2009-0209)]
- 49 Kumar G, Chatterjee SK, Chattopadhyay C. Drdnet: Diagnosis of diabetic retinopathy using capsule network (workshop paper). Proceedings of the 6th IEEE International Conference on Multimedia Big Data (BigMM). New Delhi: IEEE, 2020. 379–385.
- 50 Kalyani G, Janakiramaiah B, Karuna A, *et al.* Diabetic retinopathy detection and classification using capsule networks. Complex & Intelligent Systems, 2023, 9(3): 2651–2664.
- 51 Tan MX, Le QV. EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning. Long Beach: PMLR, 2019. 6105–6114.
- 52 徐常转, 吴云, 蓝林, 等. 融合注意力机制与多任务学习的DR分级模型. 计算机工程与应用, 2021, 57(24): 212–218. [doi: [10.3778/j.issn.1002-8331.2007-0385](https://doi.org/10.3778/j.issn.1002-8331.2007-0385)]
- 53 Chetoui M, Akhlooufi MA. Explainable diabetic retinopathy using EfficientNet. Proceedings of the 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). Montreal: IEEE, 2020. 1966–1969.
- 54 Zhao KF, Zhao WT, Xie J, *et al.* Generative adversarial networks and improved EfficientNet for imbalanced diabetic retinopathy grading. Proceedings of the 6th International Conference on Cognitive Systems and Signal Processing. Suzhou: Springer, 2022. 345–359.
- 55 邹秀芳, 朱定局. 生成对抗网络研究综述. 计算机系统应用, 2019, 28(11): 1–9. [doi: [10.15888/j.cnki.csa.007156](https://doi.org/10.15888/j.cnki.csa.007156)]
- 56 王明智, 马志强, 赵锋锋, 等. 基于代价敏感正则化和 EfficientNet 的糖尿病视网膜病变分类方法. 液晶与显示, 2022, 37(12): 1626–1635. [doi: [10.37188/CJLCD.2022-0161](https://doi.org/10.37188/CJLCD.2022-0161)]
- 57 Naveenkumar M, Srithar S, Maheswaran T, *et al.* Diabetic retinopathy disease classification using EfficientNet-B3. Innovative Data Communication Technologies and Application. Singapore: Springer, 2022. 809–821.
- 58 张选, 胡晓娟. 基于 GoogLeNet 和 ResNet 的深度融合神经网络在脉搏波识别中的应用. 计算机系统应用, 2019, 28(10): 15–26. [doi: [10.15888/j.cnki.csa.007110](https://doi.org/10.15888/j.cnki.csa.007110)]
- 59 Khalid S, Abdulwahab S, Rashwan HA, *et al.* Robust yet simple deep learning-based ensemble approach for assessing diabetic retinopathy in fundus images. Proceedings of the 5th International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT). Aligarh: IEEE, 2022. 1–5.
- 60 Gour M, Jain S, Kaushal S. XCapsNet: A deep neural network for automated detection of diabetic retinopathy. International Journal of Imaging Systems and Technology, 2023, 33(3): 1014–1027. [doi: [10.1002/ima.22842](https://doi.org/10.1002/ima.22842)]
- 61 Jian MW, Chen HY, Tao C, *et al.* Triple-DRNet: A triple-cascade convolution neural network for diabetic retinopathy grading using fundus images. Computers in Biology and Medicine, 2023, 155: 106631. [doi: [10.1016/j.combiomed.2023.106631](https://doi.org/10.1016/j.combiomed.2023.106631)]
- 62 Qian ZX, Wu CJ, Chen H, *et al.* Diabetic retinopathy grading using attention based convolution neural network. Proceedings of the 5th IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). Chongqing: IEEE, 2021. 2652–2655.
- 63 Bhardwaj C, Jain S, Sood M. Deep learning-based diabetic retinopathy severity grading system employing quadrant ensemble model. Journal of Digital Imaging, 2021, 34(2): 440–457. [doi: [10.1007/s10278-021-00418-5](https://doi.org/10.1007/s10278-021-00418-5)]
- 64 Yasin S, Iqbal N, Ali T, *et al.* Severity grading and early retinopathy lesion detection through hybrid Inception-ResNet architecture. Sensors, 2021, 21(20): 6933. [doi: [10.3390/s21206933](https://doi.org/10.3390/s21206933)]

(校对责编: 孙君艳)