

基于 Improved DR-Net 算法的糖尿病视网膜病变识别与分级

郑雯, 沈琪浩, 任佳*

浙江理工大学机械与自动控制学院, 浙江 杭州 310018

摘要 针对糖尿病性视网膜图像数据集的不均衡、组织形态的特征提取不充分、分级准确率不高等问题, 本文提出一种基于 DR-Net 模型的改进识别算法, 即 Improved DR-Net。选用 Kaggle 失明检测竞赛数据集 APTOS 2019 Dataset, 采用多种数据增强策略扩充数据集, 并引入 Eye-PACS 数据集进行无偏修正, 同时采用高斯滤波等形态学方法增强眼底图像特征; 对 ResNext50 聚合残差结构进行预训练, 通过迁移学习对基线模型进行参数及结构微调; 引入空洞卷积代替普通卷积, 融合注意力机制进一步优化模型性能。测试结果表明, 本文所提的 Improved DR-Net 模型大大提高了糖尿病视网膜病变分级的准确率: 阳性预测值 97.9%, 阴性预测值 98.03%, 准确率达到 98.04%, 远高于同类算法。结合深度学习技术辅助视网膜病变的筛查, 对于视网膜病变的早期自动筛查具有一定的指导意义。

关键词 图像处理; 糖尿病视网膜; 深度学习; 形态学处理; 聚合残差网络; 迁移学习; 注意力机制

中图分类号 TP183

文献标志码 A

doi: 10.3788/AOS202141.2210002

Recognition and Classification of Diabetic Retinopathy Based on Improved DR-Net Algorithm

Zheng Wen, Shen Qihao, Ren Jia*

School of Mechanical Engineering and Automation, Zhejiang Sci-Tech University, Hangzhou, Zhejiang 310018, China

Abstract In this paper, an improved DR-Net recognition algorithm based on DR-Net model is proposed to solve the problems of unbalanced diabetic retina image dataset, insufficient feature extraction of tissue morphology, and low classification accuracy of diabetic retinopathy. The Kaggle APTOS 2019 contest dataset is selected, which is expanded with various data enhancement strategies, and the Eye-PACS dataset is introduced for unbiased correction. Moreover, morphological methods such as Gaussian filtering are used to intensify the fundus image characteristics. Then the aggregated residual structure of ResNext50 is pre-trained and the parameters and structure of the baseline model are fine-tuned through transfer learning. In addition, the cavity convolution is introduced to replace the ordinary convolution, and the attention mechanism is also involved to further optimize the model performance. The test results show that the improved DR-Net model greatly improves the accuracy of diabetic retinopathy classification: the positive and negative predictive values reach 97.9% and 98.03%, respectively, with the accuracy being up to 98.04%, which is much higher than those of similar algorithms. In short, the screening of retinopathy with the assistance of deep learning technology is of guiding significance for the research of early automatic screening for retinopathy.

Key words image processing; diabetic retina; deep learning; morphological processing; aggregated residual network; transfer learning; attention mechanism

OCIS codes 100.3008; 100.4996; 170.3880

收稿日期: 2021-04-25; 修回日期: 2021-05-25; 录用日期: 2021-06-03

基金项目: 浙江省公益技术研究项目(LGG20F030007)、国家自然科学基金(51876194)

通信作者: *jren@zstu.edu.cn

1 引言

随着我国人民生活水平的提高,糖尿病的患病人数在近些年急剧上升。据统计,截至 2020 年我国糖尿病患者人数高达 1.3 亿,糖尿病也成为我国第二大慢性病^[1-2]。对糖尿病患者而言,糖尿病性视网膜病变(DR)是糖尿病最常见的慢性并发症之一^[3],约有 80% 的糖尿病患者正在面临这种眼部并发症的困扰。据研究人员估计发现,糖尿病的致盲率比其他致盲率高出 3 倍,已成为威胁人们眼部健康的首要病症。该病变主要是由于患者胰岛素代谢异常,从而导致眼部组织及血管微循环的改变,及时发现并精准治疗对于减少糖尿病患者失明情况十分重要。目前,医疗领域对 DR 的诊断主要通过医生分辨二维(2D)彩色眼底图片级别的方式进行,它对医生的专业水平要求比较高,需要经过长年累月的经验积累才能做出较为准确的诊断。我国能够达到这个标准的医生群体数量较小,无法满足国内巨大的 DR 患者群体的需求,因此 DR 的智能诊断和辅助医治极具研究价值^[4]。

近年来,基于计算机视觉的 DR 智能诊断技术受到广泛关注,针对视觉图像的处理方法主要有两类:传统机器学习方法和深度神经网络算法。目前在 DR 智能诊断过程中主要存在如下困难:1)不同 DR 级别的患者样例比例差异大,导致采集的 DR 图像数据集中级别样本分布不均;2)病灶过于细微,且无法忽略拍摄环境的光照不均产生的影响,导致特征难以提取。针对该问题,国内外学者开展了相关研究:郑绍华等^[5]提出了一种基于微动脉瘤和视网膜出血自动检测的非增生性 DR 图像分级方法,在红色通道上利用自适应阈值、形态学技术及区域主动轮廓模型对视盘中心进行精确定位,在绿色通道预处理并提出基于 Radon 变换的新方法对视网膜出血和微动脉瘤进行识别,有效地抑制了血管或背景纹理对病灶特征的干扰,算法对膜内出血和微动脉瘤的检测均得到较高的灵敏度和特异性,具有较高的实用价值。传统机器学习虽然取得良好的效果,但面对大量数据时则难以突破,随着深度学习的发展,卷积神经网络已被广泛应用于医疗领域^[6]的图像分类。郑婷月等^[7]利用跳跃连接构建编码器-解码器结构全卷积神经网络,将高层语义信息和低层特征信息进行融合,采用类别平衡损失函数解决正负样本不平衡问题,通过残差块进一步学习细节和纹理特征,并利用空洞卷积构建多尺度空间金字

塔池化结构,实验结果证明所提方法相较于其他方法性能更优。孙雨琛等^[8]在保证模型深度的前提下,减少模型参数,提升收敛速度,为解决梯度消失和梯度爆炸等问题引入残差模块,利用数据扩充和 Dropout 有效避免了过拟合现象,所提网络对 DR 等级分类任务的识别率达到 91%,相对于 LeNet、AlexNet 和 CompactNet 等网络模型均有 20% 以上识别率的提升。综上可知:深度网络可以有效提取图像隐含特征,在医学影像识别领域的应用取得了很大进展,但医疗影像识别技术尚处于起步阶段,在许多病理诊断辅助方面并未得到相应的广泛应用。

针对糖尿病性视网膜图像数据集的不均衡、特征提取不充分等问题,本文引入 Eye-PACS 数据集,扩充数据后采用不同比例抽取,将其修正为无偏数据集,并组合多种形态学方法对眼底图像的特征进行增强提取。针对模型分级准确率不高的问题,本文改进传统方法并优化迁移模型,提出了一种针对糖尿病视网膜病变识别与分类的深度神经网络模型,Improved DR-Net 的主要改进内容如下:

1) 利用空洞卷积代替普通卷积,使其在相同参数量和计算量下拥有更大的感受野,解决训练过程中下采样引起的图像分辨率降低、信息丢失等问题^[9];

2) 融合注意力机制^[10](CBAM),使得模型可以淡化图中无用信息的重要性,加强对病灶特点的关注,优化模型性能,进一步提高了网络 DR 分级的准确率;

3) 采用迁移学习模式,先在 ImageNet 数据集上进行预训练,学习到形状、颜色等基本特征后,通过迁移学习进一步学习病灶的细节特征。

2 Improved DR-Net 网络结构及实现

2.1 数据集介绍

本文采用 Kaggle 失明检测竞赛数据集 (APTOS 2019 Dataset, www.kaggle.com/c/aptos2019-blindness-detection)^[11] 以及 Eye-PACS 数据集^[12]开展研究,其中 APTOS 数据集包含 896 张 2D 彩色眼底图, Eye-PACS 数据集包含 35126 张 2D 彩色眼底图。目前基于 DR 可进行诊断的病灶有 4 种,分别为:微动脉瘤、出血点、硬性渗出物和软性渗出物,在 2D 彩色眼底图中的各自呈现情况如图 1 所示。其中血管汇集处的高亮区域为视盘,另一边的深暗处为中心凹;微动脉瘤以深红色点的样子呈现,是血液中的微小突起,其形成原因是血液中的

红细胞携氧量变低造成了眼部毛细血管的缺氧导致血管壁变薄;在微动脉瘤附近暗淡的小斑点状便是出血点,是由于脆弱的血管壁破裂而造成的;当渗漏的血液未能被身体完全吸收,造成有部分脂质和蛋白质残留时,眼内就会逐渐形成高亮点状的硬性渗出物;由于硬性渗出物导致了眼部血管被阻塞,血液无法继续被供给到后面更为纤细的血管处,未能得到供给的血管坏死形成如棉絮状的软性渗出物。在糖尿病性视网膜病变的后期,为了应对眼球内部长期缺氧的状况,身体不得不通过生成新的血管来维持血液和氧气的供给,这就导致了在后期眼底图中会出现更多的血管,这样的异变也会被反映到眼底图上。

根据目前医疗领域的分级标准将 DR 分为 5 级,各级症状如图 2 所示。轻度 DR(Mild_DR)只有红色的病变,也就是有小出血点或者是很小的微动脉瘤;中度 DR(Moderate_DR)就是在红色小点病灶之上又添上一些黄色的病灶,出现了一些黄白色

的点状的硬性渗出物;重度 DR(Severe_DR)在红、黄病灶之上又添加了白色的像棉团一样的软性渗出物;增殖性 DR(Proliferate_DR)已经产生视网膜的新生血管,它会导致玻璃体、视网膜的出血或明显的增殖,牵拉导致视网膜脱离^[13]。

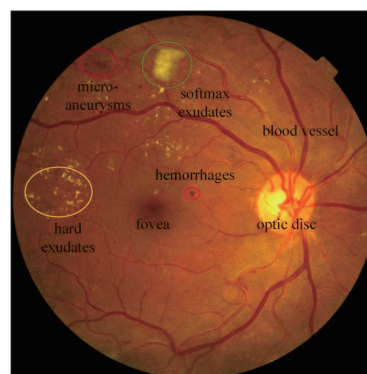


图 1 2D 彩色眼底图示例

Fig. 1 Example of 2D color fundus

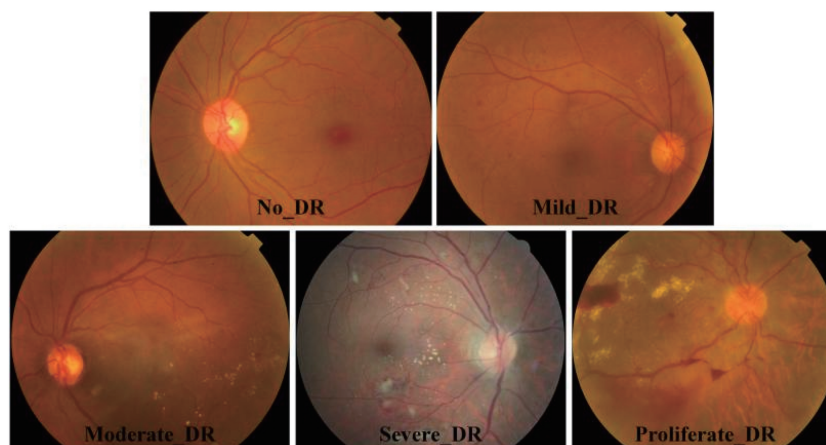


图 2 DR 分级样本示例

Fig. 2 Examples of DR grading sample

2.2 图像预处理

2.2.1 数据清洗及无偏修正

通过分析数据集发现各级别样本数量不均衡现象严重,会导致模型对数据量少的类别的检测性能较低,因此本文引入 Eye-PACS 数据集将其修正为无偏数据集。首先对原数据集进行清洗,剔除因受到光线等因素影响产生的不完整或者噪声严重的无效图片;通过不同角度的旋转、翻转、平移等数据增强手段扩充数据至 5 倍,再按照不同的比例抽取各个级别的数据进行修正,最后将两个数据集合并生成新的数据。数据清洗及无偏修正过程的分布情况详见表 1。

2.2.2 特征增强

临床医疗人员能够从彩色眼底图像中提取视网

膜血管,通过对血管形态状况的分析达到诊断这类疾病的目的。但受眼底图像采集技术的限制,图像中往往存在大量噪声,再加之视网膜血管自身结构复杂多变,使得视网膜眼底图像的特征不够清晰^[14]。因此,利用形态学方法增强视网膜眼底图像的特征,对辅助识别 DR 级别具有重要作用。

在预处理之前,先对原始图像进行 RGB 颜色通道的分离,如图 3 所示。通过对比观察及进一步的实验验证,发现其 G 通道拥有最佳的病灶与图像背景对比度特征,而 R 和 B 通道的信息几乎可以忽略,同时参考领域内相关文献^[15]的做法,本文选取 G 通道图像进行处理,整个处理过程效果如图 4 所示。

表 1 数据集修正及分布

Table 1 Dataset correction and distribution

Dataset	No_DR	Mild_DR	Moderate_DR	Severe_DR	Proliferate_DR
APTOS 2019	413	96	262	49	76
Data cleaning	363	90	192	41	72
Enhanced & extraction	1089	720	768	672	864
Eye-PACS	25810	2443	5292	873	708
Data cleaning	25584	2317	5040	554	686
Enhanced & extraction	2325	2317	2520	2219	2058
Proposed dataset	3414	3037	3288	2891	2922

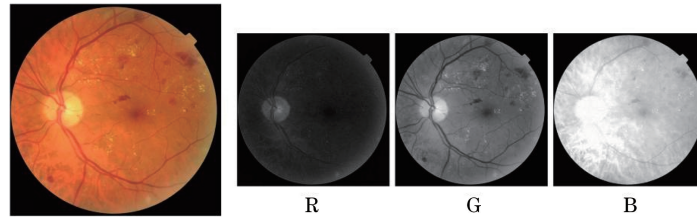


图 3 原图与 RGB 通道输出图像

Fig. 3 Original image and RGB channel output image

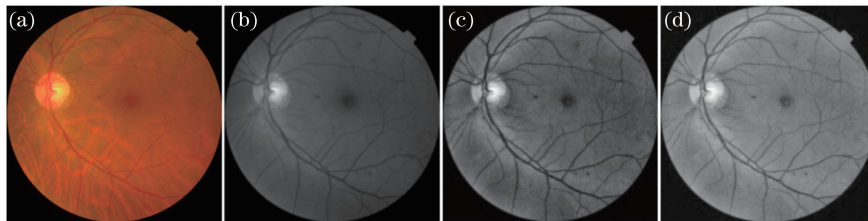


图 4 图像预处理过程及手段。(a)原始图像;(b)滤波器去噪;(c)对比度增强;(d)光均衡

Fig. 4 Image preprocessing process and means. (a) Original image; (b) filter denoising; (c) contrast enhancement; (d) light equalization

1) 高斯滤波去噪

高斯滤波是一种简单快速的线性平滑滤波去噪算法,对抑制服从正态分布的噪声有明显效果,其计算公式为

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \quad (1)$$

$$P_i(x, y) = G(x, y)P(x, y), \quad (2)$$

式中: $G(x, y)$ 表示高斯函数; $P_i(x, y)$ 表示高斯函数与原图卷积的结果,即高斯滤波去噪后的图像,如图 4(b)所示。其原理是离散化高斯函数,把离散点上的高斯函数值作为加权值,对 G 通道的眼底图像上每个像素点的灰度值做一定范围内的加权平均,用所得的函数值替换原来像素点的灰度值,即可消除噪声^[15]。

2) 对比度增强

对眼底图像进行高斯滤波处理后,病灶与背景的对比度减弱,边界变得模糊,因此需要对眼底图像进行对比度增强,突出眼底图像中病灶信息,同时弱

化干扰信息。本文采用对比度受限的自适应直方图均衡化(CLAHE)操作,限制直方图映射函数的斜率,同时对图像子块进行自适应直方图均衡化处理^[16](AHE),即计算局部直方图,然后重新分配像素值以提高图像对比度,其算法过程如下。

模板在图像上逐行移动,且模板的中心 $c(x_0, y_0)$ 对应图像上的点 $f(x_0, y_0)$,计算模板 W 区域的直方图均衡化变化关系: $g(x, y) = T[f(x, y)]$,计算模板中心点 $c(x_0, y_0)$ 的均衡化对应像素值: $g(x_0, y_0) = T[f(x_0, y_0)]$,用 $g(x_0, y_0)$ 替代 $f(x_0, y_0)$ 逐行计算得到整幅图像的自适应直方图均衡化图像,如图 4(c)所示。该算法更适合于改进图像的局部对比度以及获得更多的图像细节。自适应直方图均衡化效果的确优于一般性直方图均衡化,但易导致局部对比度提高过大从而图像失真,因此需要限制直方图映射函数的斜率^[17]。

3) 光均衡化处理

由于光照不均匀及眼底各结构对光反应的差

别,获取的眼底图像会存在光照不均匀的情况,使眼底图像存在明暗不均的区域,因此需要进行光均衡化处理,将眼底图像上的每个像素点的灰度值重新计算。一般采用的方法是将图像灰度均值搬到图像灰度中值附近,该方法操作简单但容易使图像出现光晕等过度增强现象,导致图像某些细节退化甚至消失。本文采用底帽变换^[18]操作来实现光均衡化处理,即用一个结构元通过闭运算操作从一张图像中删除不均衡背景部分,做差计算得到一张仅保留已删除分量的图像,从而消除光照不均匀现象,其计算公式为

$$B_{\text{hat}}(f) = (f \circ b) - f, \quad (3)$$

式中: f 为图像; b 表示闭操作。由图 4(d)可以看

到底帽变换后的图像删除了大部分非均匀背景,解决了光照不均匀的问题,实现了眼底图像的光均衡化。

2.3 Improved DR-Net 网络模型设计

本文提出的 Improved DR-Net 网络模型如图 5 所示,其结构包括空洞卷积、聚合残差模块(ResNext block)以及注意力机制。将预处理后的眼底图像放入 Improved DR-Net 模型中训练,在训练数据不足的情况下,迁移学习通过加载在 ImageNet 数据集上预训练的 ResNext50 权重,在开始优化梯度之前,该模型有一个较好的权重初始化,以此来训练自己的模型。接下来将详细介绍各模块的结构及功能。

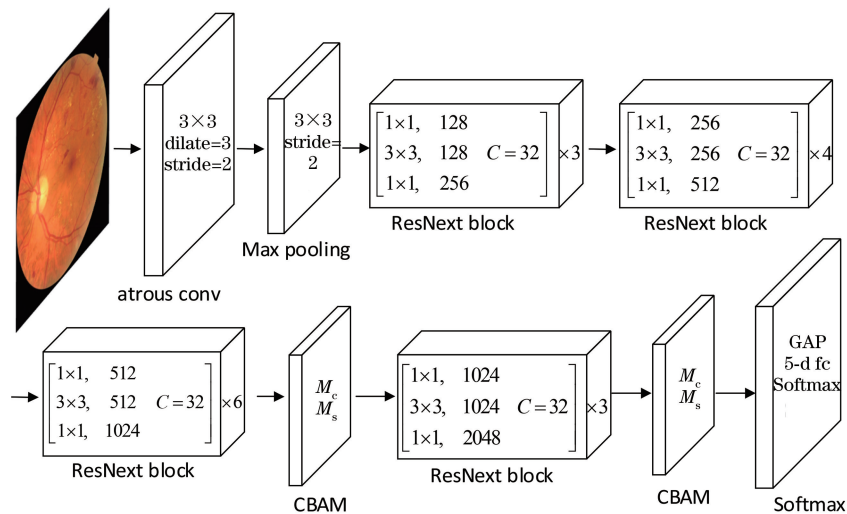


图 5 Improved DR-Net 网络模型架构

Fig. 5 Architecture of improved DR-Net network model

2.3.1 空洞卷积

随着网络层数的不断加深,特征图的空间分辨率会逐渐降低。为了降低这种影响,本文在网络的卷积层中引入了空洞卷积策略^[9],空洞卷积的原理如图 6 所示,相比原来的标准卷积,空洞卷积多了一个超参数称之为扩张率,指的是卷积核各点之前的间隔数量。扩张率为 1 时的卷积核,相当于网络进行正常的 3×3 卷积操作;扩张率为 2 时的卷积核,相当于 5×5 正常卷积核的感受野;扩张率为 3 时的卷积核,相当于 7×7 正常卷积核的感受野。本文将扩张卷积引入代替一般性的卷积,该策略可以在不增加算法参数复杂度的情况下增大网络的感受野,从而提高网络捕捉特征信息的能力^[19]。普通卷积和扩张卷积特征图尺寸的计算公式分别为

$$\text{Floor}\left(\frac{F - k + 2p}{s}\right) + 1, \quad (4)$$

$$\text{Floor}\left(\frac{F - [k + (k - 1)(d - 1) + 2p]}{s}\right) + 1, \quad (5)$$

式中: F 为原特征图尺寸; $k \times k$ 表示卷积核大小; p 为填充; s 为步长; d 表示扩张率。

2.3.2 聚合残差结构

通过实验结果对比分析,本文选取基于聚合残差模块的 ResNext50 作为基线模型^[20],该网络同时采用了 VGGNet 网络堆叠的思想^[21]和 Inception 网络的拆分-转换-合并策略^[22],除了深度、宽度外,将并行的路径数提取为第三维度,称为基数,将卷积神经网络的通道进行拆分,通过增加基数来提高网络的性能,使得网络在不增加参数复杂度的前提下提高准确率。经验表明,无论在数据集 ImageNet 还是 COCO 上的研究结果都比 ResNet 表现更优。

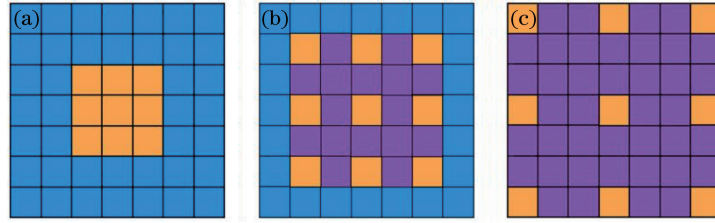


图 6 不同扩张率下的空洞卷积。(a) $d=1$; (b) $d=2$; (c) $d=3$

Fig. 6 Dilated convolution under different expansion rates. (a) $d=1$; (b) $d=2$; (c) $d=3$

图 7(a) 是 ResNet 的一个 block; 图 7(b) 中 ResNext 分支的个数就是基数, 每个分支均包括输入通道数、卷积核大小、输出通道数, 与图 7(a) 结构有同样的参数个数; 本文采用的是与其等价的

grouped convolutions 结构, 如图 7(c) 所示, 每个 group 的输入输出 channels 均为 4, 最后把 channels 合并, 当基数为 32 时中间层的通道数为 128, 该结构相比一般的残差结构可以有效地减少计算量。

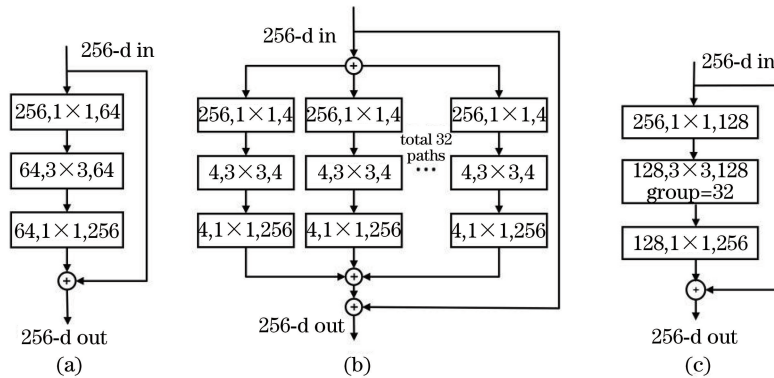


图 7 聚合残差结构图

Fig. 7 Architecture of aggregated residual

2.3.3 注意力机制的引入

深度学习中的注意力机制从本质上讲和人类的选择性视觉注意力机制类似, 其主要作用是从卷积层提取到的各种特征中找出更为关键的特征信息, 从而辅助网络对目标的检测。

注意力机制^[10]从通道和空间两个维度计算特

征图的注意力权重, 然后将获得的权重与输入的特征图依次相乘来进行特征的自适应学习, 结构如图 8 所示, 其中 \otimes 表示逐元素相乘。实验表明将通道注意力模块和空间注意力模块串行放置且通道模块在前的组合方式优于其他组合效果。

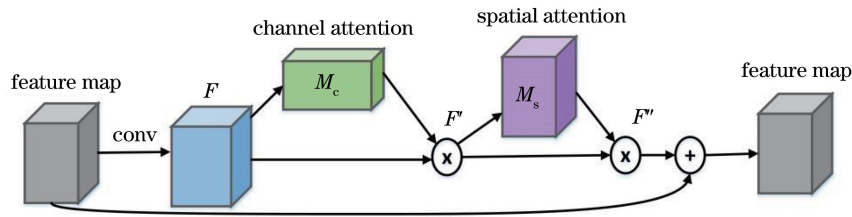


图 8 注意力机制结构图

Fig. 8 Structure diagram of attention mechanism

通道注意力模块的输入是一个 $H \times W \times C$ 的特征图 F , 最大池化编码了最显著的部分, 而平均池化编码了全局的统计信息, 为有效计算通道注意力, 该模块同时使用平均池化和最大池化得到两个 $1 \times 1 \times C$ 的通道描述, 再将它们分别送入一个两层的神经网络, 第一层神经元个数为 C/r , 第二层神经元个

数为 C 。最后, 将得到的两个特征相加后经过一个 Leaky_ReLU 激活函数得到权重系数 M_c , 计算公式为

$$M_c(F) = \sigma\{\text{MLP}[\text{AvgPool}(F)] + \text{MLP}[\text{MaxPool}(F)]\} = \sigma\{W_1[W_0(F_{\text{avg}}^c)] + W_1[W_0(F_{\text{max}}^c)]\}, \quad (6)$$

式中: $W_0 \in \mathbf{R}^{C/r \times C}$; $W_1 \in \mathbf{R}^{C \times C/r}$; σ 表示 LeakyReLU 函数; MLP 为多层感知机。CBAM 模块将通道注意力权重系数 M_c 和原来的特征 F 相乘即可得到缩放后的新特征, 即 $F' = M_c(F) \otimes F$, 生成空间注意力模块需要的输入特征 F 。与通道注意力相似, 空间注意力模块的输入为上一模块所得大小为 $H \times W \times C$ 的特征图 F' , 为计算空间注意力, 该模块沿着通道轴应用平均池化和最大池化操作得到两个 $H \times W \times 1$ 的通道描述图, 再将它们通过 concatenation 操作合并特征通道生成一个有效的特征层。通过一个 7×7 的卷积层和 LeakyReLU 激活函数生成大小为 $H \times W \times 1$ 的空间注意力图 M_s , 该空间注意力图编码了需要关注或抑制的位置, 计算公式为

$$M_s(F) = \sigma\{f^{7 \times 7}[\text{AvgPool}(F); \text{MaxPool}(F)]\} = \sigma\{f^{7 \times 7}[F_{\text{avg}}^s; F_{\text{max}}^s]\}, \quad (7)$$

式中: F_{avg}^s 大小为 $H \times W \times 1$; F_{max}^s 大小为 $H \times W \times 1$; $f^{7 \times 7}$ 表示滤波器大小为 7×7 的卷积运算。将权重系数 M_s 和特征 F' 相乘即可得到缩放后的新特征, 即 $F = M_s(F') \otimes F'$ 。

3 实验与讨论

本文实验环境为 Windows10 操作系统, 服务器 CPU i7 9700K, 显卡 GPU GTX 1080Ti, 运行内存 16 G, 编程环境为 Python, 深度学习框架为 Pytorch。本文采用随机梯度下降 (SGD)^[23] 作为 DR-Net 模型的优化算法, 使用动态学习率 (Multi-Sep-LR), 初始学习率 $\alpha = 0.003$, 动态梯度 gamma 为 0.2, 共训练 400 个 epoch, 学习率在 epoch 为 160、240、320 时改变。同时, 采用交叉熵损失函数进行迭代, batch_size 设置为 64, 模型初始值为 ResNext 在 ImageNet 上预训练后的参数^[24]。

根据修正后 15552 张眼底图片, 70% 训练集用于神经网络模型的参数训练, 27% 测试集用于反馈迭代, 3% 验证集用于模型评估。实验结果准确率达到 98.04%, 测试结果如图 9 所示, 测试集的准确率及误差曲线如图 10 所示, 验证集的混淆矩阵如图 11 所示。由图 10(a) 可见在 epoch=160 时准确率明显提高, 当 epoch=240 之后准确率处于平稳缓慢的上升。

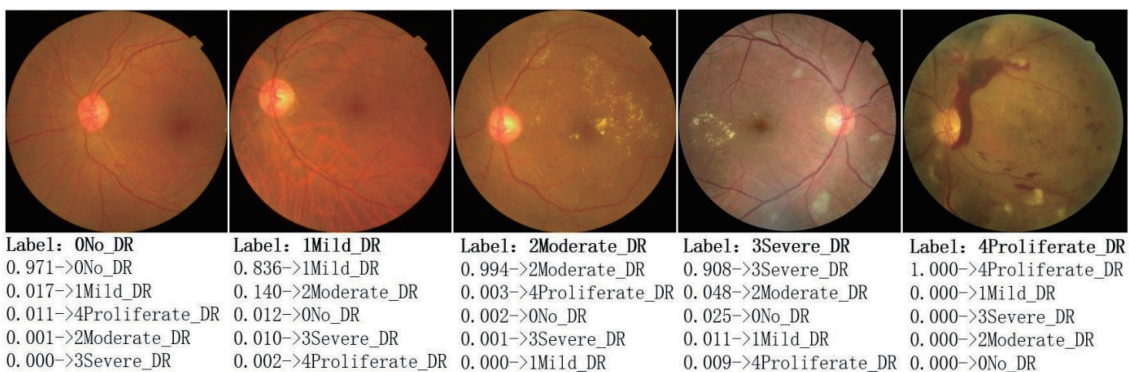


图 9 Improved DR-Net 测试结果示例

Fig. 9 Examples of improved DR-Net test results

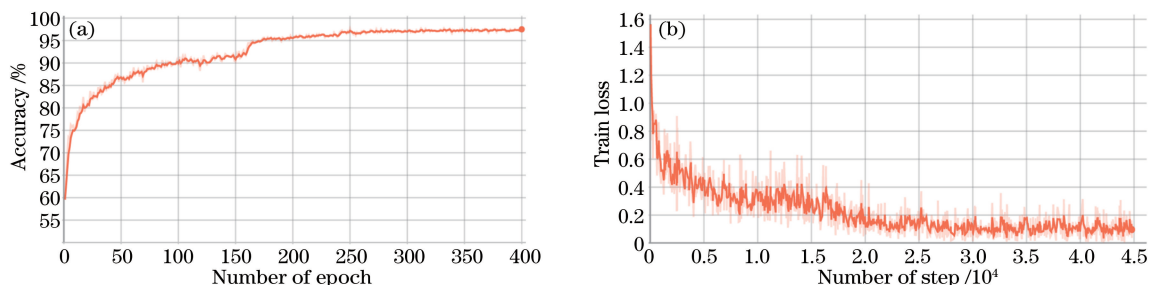


图 10 Improved DR-Net 训练过程的准确率及误差曲线。(a) 准确率; (b) 训练损失

Fig. 10 Accuracy and error curve in training process of improved DR-Net. (a) Accuracy; (b) train loss

3.1 评价标准

本文采用准确率、灵敏度、特异性、阳性预测值 (PPV) 以及阴性预测值 (NPV) 作为算法评价指

标^[25]。

正确率表示对于给定的测试数据集, 正确分类的样本数与总样本数的比,

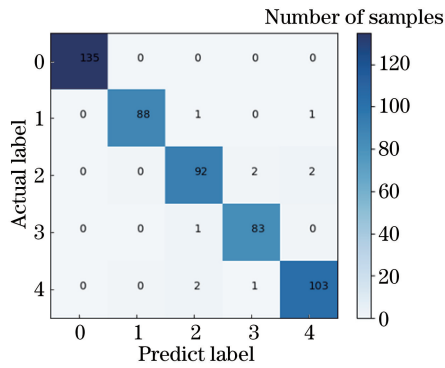


图 11 Improved DR-Net 模型混淆矩阵测试结果
Fig. 11 Confusion matrix test results of improved DR-Net model

$$R_{accuracy} = \frac{R_{TP} + R_{TN}}{R_{TP} + R_{TN} + R_{FP} + R_{FN}} \quad (8)$$

敏感度(也称召回率)用来表示实际为阳性的样本中判断为阳性的比例,敏感度的值越大说明“漏检”越少,

$$R_{sensitivity} = \frac{R_{TP}}{R_{TP} + R_{FN}} \quad (9)$$

特异度是指实际为阴性的样本中,判断为阴性的比例,特异度的值越大说明“误检”(FP)越少,

$$R_{specificity} = \frac{R_{TN}}{R_{TN} + R_{FP}} \quad (10)$$

阳性预测值为预测阳性中准确的占比,表示为

$$R_{PPV} = \frac{R_{TP}}{R_{TP} + R_{FP}} \quad (11)$$

阴性预测值为预测阴性中准确的占比,表示为

$$R_{NPV} = \frac{R_{TN}}{R_{TN} + R_{FN}} \quad (12)$$

式中: R_{TP} 表示正确预测正样本数量; R_{FP} 表示将正样本错误判断为负样本数量; R_{TN} 表示正确预测负样本数量; R_{FN} 表示将负样本错误判断为正样本数量。

3.2 实验结果

本文将多种主流模型包括 Densenet121、Mobilenetv2、Efficient-net b5、ResNext50 等进行了对比测试,各种模型对应的混淆矩阵如图 12 所示。结果表明:在相同实验环境下,ResNext50 对糖尿病视网膜病变分级的准确率最高。

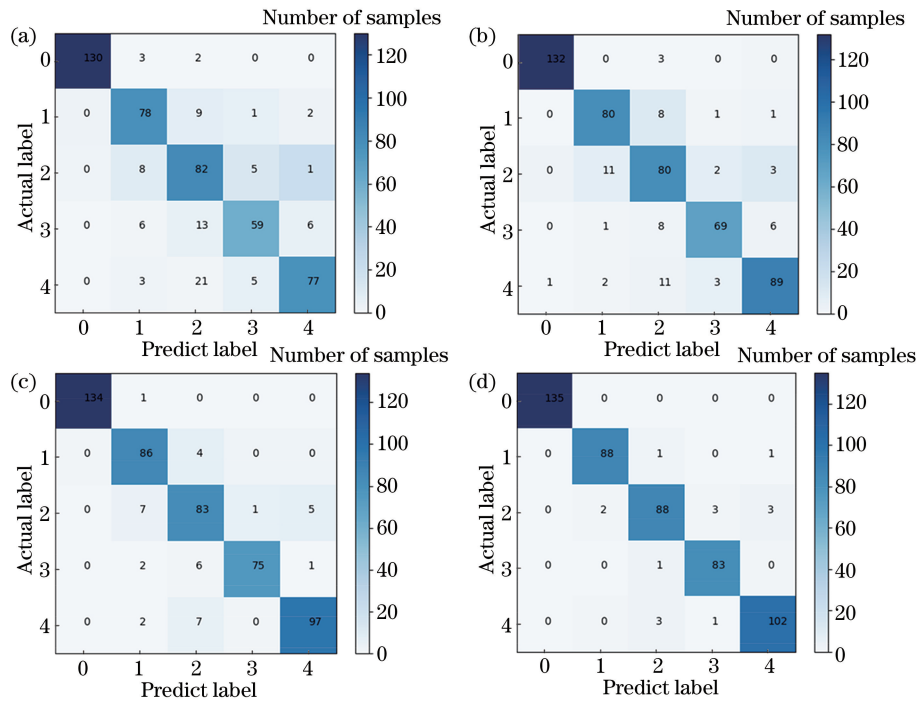


图 12 主流模型对 DR 分级的混淆矩阵对比实验结果。(a) Densenet121;(b) Mobilenetv2;
(c) Efficient-net b5;(d) ResNext50

Fig. 12 Comparison of confusion matrix among mainstream models on DR classification. (a) Densenet121;
(b) Mobilenetv2; (c) Efficient-net b5; (d) ResNext50

同时,本文对比了 DR-Net (ResNext50) 与 Improved DR-Net 模型的性能,如图 13 所示。实验证明,Improved DR-Net 模型相较于 DR-Net 模型,

五项指标均有不同程度的提升,对糖尿病视网膜病变的自动分级起到重要作用。

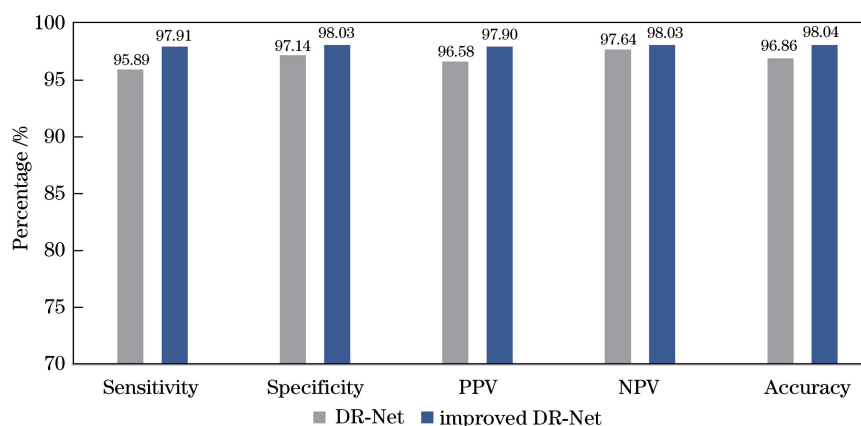


图 13 DR-Net 与 Improved DR-Net 模型性能对比

Fig. 13 Comparison of performance on DR-Net and improved DR-Net models

3.3 与其他算法的对比实验结果

在糖尿病视网膜病变图像智能诊断及识别领域,很多学者提出了多样的解决方案;Li 等^[26]采用迁移学习的方法,从其他相关任务的大数据集中学习知识,并在目标数据集中实现高精度的分类效果,解决了眼底影像数据匮乏的问题;为了更好地诊断轻微 DR,防治严重病情的发生;谢云霞等^[27]提出了改进 Faster R-CNN 网络和子图分割相结合的 DR 分级方法,子图分割可解决视盘区域对于病灶识别的干扰问题,深度残差网络以解决病灶在高分辨率眼底图像中占比小而导致的特征难以获取的问题,采用样本挖掘方法解决正负难易样本不平衡的问题,在公开数据集 EyePACS 进行实验,所提方法在 DR 分期中精确率很高;van Grinsven 等^[28]采用了动态选择抽样策略解决了模型训练过程中正样本过多、数据集不平衡的问题,将训练精力动态地集中在训练难度较大的样本上,从而加快模型训练任务,提高整体性能;李轩屹等^[29]为对视网膜图像的 DR 严重程度进行更准确的分类,提出一种多尺度 ResNext 网络模型。使用不同分辨率眼底图像交替作为输入层数据,采用加权融合深浅层特征作为全连接层分类的信息,

使用迁移学习方法对网络模型参数进行初始化,避免发生过拟合问题,结果表明,相对于传统模型,该模型方法可以更准确地对 DR 严重程度进行分类;张子振等^[30]提出一种融合注意力机制和高效率网络的 DR 自动分类方法,对 Kaggle (2015) 和 MESSIDON (French Ministry, 2014) 数据集进行剔除、去噪、扩增和归一化等处理,采用迁移学习策略对 Efficient Net 进行学习及训练提取深度特征,为解决病变之间差异小的问题加入注意力机制,分类精度达到 97.2%,具有较好的分类性能及鲁棒性。

表 2 给出了本文算法 Improved DR-Net 与上述算法在各项指标下的对比结果。结果表明:本文算法在各项指标中均具有较明显的优势,性能均衡,且识别准确率提高到了 98.04%。进一步分析原因,可总结如下:1) Improved DR-Net 算法通过引入 Eye-PACS 数据集,结合数据增强技术较好地解决了数据集中类别样本不均衡的问题;2) 引入高斯滤波、自适应直方图均衡化、底帽变换为主的形态学方法,能够提取类别有效特征;3) 网络结构的改进,以及参数预训练、迁移学习的引入保证了网络结构和参数的最优化,进一步提升了模型的性能。

表 2 现有 DR 分级算法对比结果

Table 2 Comparison results of existing DR classification algorithms

unit: %

Related research	Sensitivity	Specificity	PPV	NPV	Accuracy
Densenet121	82.25	83.28	83.60	84.59	83.36
Mobilenetv2	87.22	87.91	90.46	87.98	88.06
Ref. [26]	97.11	86.03	93.41	96.63	92.01
Ref. [27]	97.72	87.11	89.17	97.45	92.41
Efficient-net b5	92.39	92.91	92.52	92.62	92.55
Ref. [28]	84.80	90.40	92.72	92.92	92.95
Ref. [29]	90.50	95.60	94.17	90.96	96.90
Ref. [30]	95.60	98.70	98.66	95.73	97.20
Improved DR-Net	97.91	98.03	97.90	98.03	98.04

4 结 论

本文针对糖尿病视网膜病变细胞分级准确率不均衡、检测效率低等现象提出一种基于 DR-Net 模型的改进识别算法。选用 ResNext50 聚合残差结构进行预训练,通过迁移学习对基线模型进行微调,融合注意力机制 CBAM 优化模型性能,根据数据集分布偏差大、数量小等特点,使用数据增强技术扩充数据集避免模型过拟合,并修正数据集为无偏数据提高模型的鲁棒性,同时在预处理中采用底帽变换等多种形态学方法进行眼底图像的特征增强,利用空洞卷积代替普通卷积,解决图像语义分割问题中下采样会降低图像分辨率、丢失信息的问题。实验结果表明,本文所提的 DR-Net 模型大大提高了糖尿病视网膜病变分级的准确性,远高于同类算法,对于促进视网膜病变早期自动筛查系统的研究具有一定的参考价值。

参 考 文 献

- [1] He B L, He Y. Research advance in the correlation between diabetes mellitus and the nonretinal eye diseases[J]. *International Eye Science*, 2021, 21(4): 623-627.
何蓓蕾, 何媛. 糖尿病与非视网膜眼部并发症相关性的研究进展[J]. *国际眼科杂志*, 2021, 21(4): 623-627.
- [2] Liu Y H, Gao L. Current status, problems and prospects of treatment research on diabetic retinopathy[J]. *Chinese Journal of Ocular Fundus Diseases*, 2016, 32(2): 206-210.
刘玉华, 高玲. 糖尿病视网膜病变治疗研究现状, 问题与展望[J]. *中华眼底病杂志*, 2016, 32(2): 206-210.
- [3] Zhu D. Clinical observation on acupuncture combined with hexue mingmu tablets in treatment of diabetic retinopathy [J]. *Journal of Hubei University of Chinese Medicine*, 2018, 20(4): 87-89.
朱丹. 针刺联合和血明目片治疗糖尿病视网膜病变的临床观察[J]. *湖北中医药大学学报*, 2018, 20(4): 87-89.
- [4] Dong H Y. Research on image classification method of diabetic retinopathy based on machine learning [D]. Beijing: Beijing University of Posts and Telecommunications, 2019: 1-60.
董慧妍. 基于机器学习的糖尿病性视网膜病变图像分级研究[D]. 北京: 北京邮电大学, 2019: 1-60.
- [5] Zheng S H, Pan L, Chen J, et al. Grading method for non-proliferative diabetic retinopathy images based on microaneurysms and hemorrhages automatic detection [J]. *Chinese Journal of Scientific Instrument*, 2014, 35(1): 59-67.
郑绍华, 潘林, 陈健, 等. 微动脉瘤与出血自动检测的 NPDR 图像分型方法[J]. *仪器仪表学报*, 2014, 35(1): 59-67.
- [6] Fan J W, Zhang R R, Lu M, et al. Applications of deep learning techniques for diabetic retinal diagnosis [J]. *Acta Automatica Sinica*, 2021, 47(5): 985-1004.
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[J]. *自动化学报*, 2021, 47(5): 985-1004.
- [7] Zheng T Y, Tang C, Lei Z K. Multi-scale retinal vessel segmentation based on fully convolutional neural network [J]. *Acta Optica Sinica*, 2019, 39(2): 0211002.
郑婷月, 唐晨, 雷振坤. 基于全卷积神经网络的多尺度视网膜血管分割[J]. *光学学报*, 2019, 39(2): 0211002.
- [8] Sun Y C, Liu Y H, Zhang D F, et al. Diagnosis method of diabetic retinopathy based on deep learning [J]. *Laser & Optoelectronics Progress*, 2020, 57(24): 241701.
孙雨琛, 刘宇红, 张达峰, 等. 基于深度学习的糖尿病视网膜病变诊断方法[J]. *激光与光电子学进展*, 2020, 57(24): 241701.
- [9] Shang Q F, Shen W, Shuai S Y. Semantic segmentation of high resolution remote sensing image based on deep learning [J]. *Computer Systems & Applications*, 2020, 29(7): 180-185.
尚群锋, 沈炜, 帅世渊. 基于深度学习高分辨率遥感影像语义分割[J]. *计算机系统应用*, 2020, 29(7): 180-185.
- [10] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11211: 3-19.
- [11] Khalifa N E M, Loey M, Taha M H N, et al. Deep transfer learning models for medical diabetic retinopathy detection[J]. *Acta Informatica Medica*, 2019, 27(5): 327-332.
- [12] Cuadros J, Bresnick G. EyePACS: an adaptable telemedicine system for diabetic retinopathy screening [J]. *Journal of Diabetes Science and Technology*, 2009, 3(3): 509-516.
- [13] Wang P. A study on the correlation between choroidopathy and retinopathy in diabetic patients [D]. Jinan: Shandong University, 2019: 19-48.
王平. 糖尿病患者脉络膜病变与视网膜病变相关性

- 的研究[D]. 济南: 山东大学, 2019: 19-48.
- [14] Xu G Z, Wang Y W, Hu S, et al. Retinal vascular segmentation combined with PCNN and morphological matching enhancement [J]. Opto-Electronic Engineering, 2019, 46(4): 74-85.
徐光柱, 王亚文, 胡松, 等. PCNN 与形态匹配增强相结合的视网膜血管分割[J]. 光电工程, 2019, 46(4): 74-85.
- [15] Ma Q K. Design and development of fundus image inspection and analysis system [D]. Guangzhou: Jinan University, 2018: 1-62.
马青柯. 眼底图像检测和分析系统的设计与开发[D]. 广州: 暨南大学, 2018: 1-62.
- [16] Wu C M. Studies on mathematical model of histogram equalization [J]. Acta Electronica Sinica, 2013, 41(3): 598-602.
吴成茂. 直方图均衡化的数学模型研究[J]. 电子学报, 2013, 41(3): 598-602.
- [17] Yang W Z, Xu Y L, Qiao X, et al. Method for image intensification of underwater sea cucumber based on contrast-limited adaptive histogram equalization [J]. Transactions of the Chinese Society of Agricultural Engineering, 2016, 32(6): 197-203.
杨卫中, 徐银丽, 乔曦, 等. 基于对比度受限直方图均衡化的水下海参图像增强方法[J]. 农业工程学报, 2016, 32(6): 197-203.
- [18] Yuan W Q, Xue D. Review of tunnel lining crack detection algorithm based on machine vision [J]. Chinese Journal of Scientific Instrument, 2017, 38(12): 3100-3111.
苑玮琦, 薛丹. 基于机器视觉的隧道衬砌裂缝检测算法综述[J]. 仪器仪表学报, 2017, 38(12): 3100-3111.
- [19] Tang X L, Du Y M, Liu Y W, et al. Image recognition with conditional deep convolutional generative adversarial networks [J]. Acta Automatica Sinica, 2018, 44(5): 855-864.
唐贤伦, 杜一铭, 刘雨微, 等. 基于条件深度卷积生成对抗网络的图像识别方法[J]. 自动化学报, 2018, 44(5): 855-864.
- [20] Xie S N, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 5987-5995.
- [21] Ke H J, Chen D, Li X L, et al. Towards brain big data classification: epileptic EEG identification with a lightweight VGGNet on global MIC [J]. IEEE Access, 2018, 6: 14722-14733.
- [22] Alom M Z, Hasan M, Yakopcic C, et al. Improved inception-residual convolutional neural network for object recognition [J]. Neural Computing and Applications, 2020, 32(1): 279-293.
- [23] Ji Z Y, Zhang X J, Fu Z, et al. Performance-awareness based dynamic batch size SGD for distributed deep learning framework [J]. Journal of Computer Research and Development, 2019, 56(11): 2396-2409.
纪泽宇, 张兴军, 付哲, 等. 分布式深度学习框架下基于性能感知的 DBS-SGD 算法[J]. 计算机研究与发展, 2019, 56(11): 2396-2409.
- [24] Ge Y, Jiang S L, Ye F M, et al. Remote sensing image retrieval using pre-trained convolutional neural networks based on ImageNet [J]. Geomatics and Information Science of Wuhan University, 2018, 43(1): 67-73.
葛芸, 江顺亮, 叶发茂, 等. 基于 ImageNet 预训练卷积神经网络的遥感图像检索[J]. 武汉大学学报·信息科学版, 2018, 43(1): 67-73.
- [25] Tang J. Research on automatic grading and focus recognition system of myopic macular disease based on color fundus photos [D]. Beijing: Peking Union Medical College, 2020: 23-35.
汤加. 基于眼底彩照的近视性黄斑病变自动分级和病灶识别系统研究[D]. 北京: 北京协和医学院, 2020: 23-35.
- [26] Li X G, Pang T T, Xiong B, et al. Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification [C]//2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), October 14-16, 2017, Shanghai, China. New York: IEEE Press, 2017: 1-11.
- [27] Xie Y X, Huang H Y, Hu J B. Staging and lesion detection of diabetic retinopathy based on deep convolution neural network [J]. Journal of Computer Applications, 2020, 40(8): 2460-2464.
谢云霞, 黄海于, 胡建斌. 基于深度卷积神经网络的糖尿病视网膜病变分期及病灶检测[J]. 计算机应用, 2020, 40(8): 2460-2464.
- [28] van Grinsven M J J P, van Ginneken B, Hoyng C B, et al. Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images [J]. IEEE Transactions on Medical Imaging, 2016, 35(5): 1273-1284.
- [29] Li X Y, Zhu X J. Classification of diabetic retinopathy with multi-scale depth CNN [J]. Computer Engineering and Design, 2020, 41(11): 3229-3234.
李轩屹, 朱晓军. 多尺度深度 CNN 的糖尿病视网膜

- 病变分类[J]. 计算机工程与设计, 2020, 41(11): 3229-3234.
- [30] Zhang Z Z, Liu M, Zhu D J. Automatic recognition and classification of diabetic retinopathy images by combining an attention mechanism and an efficient network[J]. Journal of Image and Graphics, 2020, 25(8): 1708-1718.
张子振, 刘明, 朱德江. 融合注意力机制和高效网络的糖尿病视网膜病变识别与分类[J]. 中国图象图形学报, 2020, 25(8): 1708-1718.
- [31] Wang C J, Sun T, Chen J. A novel target recognition method based on fast retina key point descriptor for laser active imaging[J]. Chinese Journal of Lasers, 2014, 41(3): 0309004.
王灿进, 孙涛, 陈娟. 基于快速视网膜关键点局部不变特征的激光主动成像目标识别[J]. 中国激光, 2014, 41(3): 0309004.
- [32] Lian C M, Zhong S C, Zhang T F, et al. Transfer learning-based classification of optical coherence tomography retinal images [J]. Laser & Optoelectronics Progress, 2021, 58(1): 011702.
连超铭, 钟舜聪, 张添福, 等. 光学相干断层扫描视网膜图像的迁移学习分类[J]. 激光与光电子学进展, 2021, 58(1): 011702.