

基于深度学习的糖尿病眼底病变分级方法研究

蒋鹏¹, 何勇¹, 姚凯学¹, 胡加德²

(1. 贵州大学 计算机科学与技术学院, 贵州 贵阳 550025;

2. 贵州科海新技术发展有限公司, 贵州 贵阳 550002)

摘要:针对目前糖尿病视网膜病变图像人工分级困难、识别精度差等问题,研究了一种基于复合缩放模型的糖尿病视网膜病变图像分级方法,用于检测糖尿病视网膜病变,达到辅助医师诊断的目的。首先对糖尿病眼底病变图像进行了详细分析,经过图像预处理后,采用图像混合增强图像特征,同时使用图像翻转、图像加噪、调节对比度等不同方式扩充训练集,最后利用迁移学习和基于复合缩放模型方法进行糖尿病眼底图像病变程度分级。经过多组实验表明,该方法对糖尿病视网膜病变图像分级的准确率高达92%,Kappa系数高达0.88。所提出的方法无需指定病变的特征就能够达到高精度的病变分级,相较于 AlexNet、LeNet、CompactNet、ResNet 等模型两项指标有显著提升,对糖尿病视网膜病变诊断能提供有效的科学依据,同时对研究其他眼底病变图像也有一定参考价值。

关键词:糖尿病;视网膜病变;深度学习;迁移学习;特征增强;复合缩放模型

中图分类号:TP183

文献标识码:A

文章编号:1673-629X(2021)12-0193-05

doi:10.3969/j.issn.1673-629X.2021.12.032

Research on Classification of Diabetic Retinopathy Based on Deep Learning

JIANG Peng¹, HE Yong¹, YAO Kai-xue¹, HU Jia-de²

(1. School of Computer Science and Technology, Guizhou University, Guiyang 550025, China;

2. Guizhou Kehai New Technology Development Co., Ltd., Guiyang 550002, China)

Abstract: In order to solve the problems of difficulty in manual classification and poor recognition accuracy of diabetic retinopathy (DR) image, a DR image classification method based on composite scaling model is studied, which is used to detect DR and assist doctors in diagnosis. Firstly, the image of diabetic fundus lesions is analyzed in detail. After image preprocessing, the image features are enhanced by image blending. At the same time, the training set is expanded by image flipping, image noise adding and contrast adjusting. Finally, the degree of diabetic fundus lesions is graded by using transfer learning and composite scaling model. After a number of experiments, the accuracy of the proposed method for diabetic retinopathy image classification is as high as 92%, and the Kappa coefficient is as high as 0.88. It is showed that the proposed method can achieve high-precision classification of diabetic retinopathy without specifying the characteristics of the disease. Compared with AlexNet, LeNet, CompactNet, ResNet and other models, the proposed method has significantly improved the accuracy of diabetic retinopathy image classification. The diagnosis of the disease can provide an effective scientific basis, and it also has a certain reference value for the study of other fundus lesions.

Key words: diabetes; retinopathy; deep learning; transfer learning; feature enhancement; compound scaling model

0 引言

糖尿病会引起视网膜病变,在几乎所有的 I 型糖尿病病人和超过 60% 的 II 型糖尿病病人都会发生视网膜病变,对人体视力影响极大,糖尿病视网膜病变属于目前造成失明和视力受损的主要原因之一^[1]。但处于糖尿病视网膜病变早期的患者,其视力或完全不受影响,

患者无明显感受,然而目前没有治疗视网膜病变晚期的良好方案,需要在病变早期进行预防^[2-3],在此前提下,对眼底图像病变程度的识别十分重要。然而除糖尿病外,动脉硬化、高血压、肾炎、妊娠、白血病、贫血等都会引起眼底变化^[4],传统的人工识别需要具备专业、丰富的医学经验。因此,一种能有效、客观地观测和筛

收稿日期:2020-12-31

修回日期:2021-05-06

基金项目:贵州省科学技术科技成果应用及产业化计划项目(黔科合成果[2018]4002)

作者简介:蒋鹏(1995-),男,硕士研究生,CCF 会员(E8600G),研究方向为嵌入式与物联网;何勇,副教授,研究方向为嵌入式与无线传感器网络;姚凯学,教授,研究方向为嵌入式与物联网。

查糖尿病性视网膜病变的方法是十分有意义的。

目前,糖尿病视网膜眼底图像分级面临特征差异小,特征位置分布不均,干扰因素多等难题。深度学习可以有效提取图像隐含特征,学习识别特征层次结构,在医学图像领域内有广泛的应用^[5]。因此,该文采用深度学习相关算法基于复合缩放模型建立神经网络,使用预处理后的眼底图像,训练神经网络模型进行糖尿病性视网膜病变程度分级,其结果可以作为诊断参考,同时将来还可以用来检测其他疾病,比如青光眼和黄斑变性。

1 数据集介绍与数据分析

该文使用了公开可用的 Kaggle 数据集。Kaggle 数据集由免费的视网膜病变筛查平台 EyePACS 提供。该数据集中有 5 590 幅高分辨率彩色左右双眼的眼底图像,由专业的临床医生对每张图片上的糖尿病视网膜病变进行了 0 到 4 的评分,分别对应没有、一般、中度、严重、增生五种糖尿病视网膜病变程度。该文将数据集样本依上对其标注分为两部分,其中训练集 3 662 张图片,测试集 1 928 张图片,再由训练集分出验证集,其占比为 0.2。

从整体数据集上分析,由数据集中提取每张图像的分级标签进行统计,可以发现该数据集的糖尿病视网膜病变等级分布十分不均,其中正常的眼底图像(等级 0)的占比达到 49.3%,属于有偏数据集,如图 1 所示。

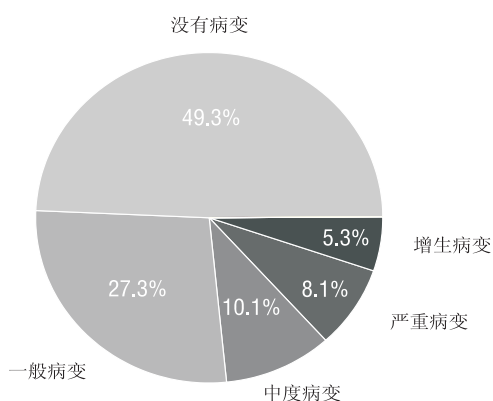


图 1 数据集中不同视网膜病变程度分布饼状图

眼底主要由玻璃体、视网膜、视神经、黄斑区等结构构成。当眼底发生病变时,上述相应的结构及其颜色都会发生相应的改变^[6],在眼底图像中会显现血噬菌体、硬性渗出物、小动脉瘤、出血、水肿和棉絮斑点等不同程度的症状。不同的病变等级,变性的范围或大或小,同时,伴随视神经的萎缩,其颜色也会有所不同^[7-8]。

如图 2 所示,经由上述分析中总结得出,主要的图像特征有:渗出物、变性面积、视神经颜色。但同时图

像特征的具体的结构和位置信息很难依靠经验总结^[9-10],图像潜在的其他特征有待挖掘。

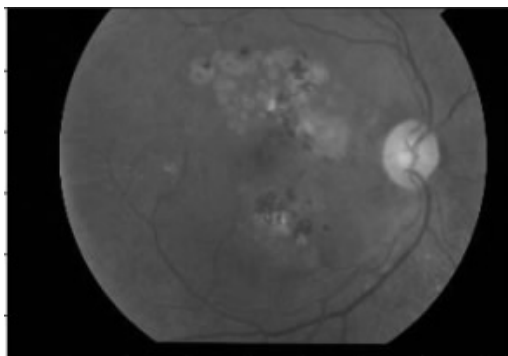


图 2 数据集中严重的视网膜病变图样

2 数据预处理

数据预处理对于训练神经网络十分重要,对最终模型的影响毋庸置疑。经观察分析,可发现数据集中的眼底图像由于采集图像时环境条件所限,图像数据有两个可以改进的地方。第一,图像中出现的黑边,这属于干扰信息,可以进行裁剪,消除其影响。第二,通过对比不同的照片,可以发现眼底图像有不同的亮度,有些图像非常昏暗,从而导致图像中主要特征与周围的差异不明显,需要减少光照的影响。以上数据集暴露的问题,主要来源于硬件设备,问题同时存在于实际之中。为克服以上的影响,该文研究了以下的数据预处理方法,流程如图 3 所示。

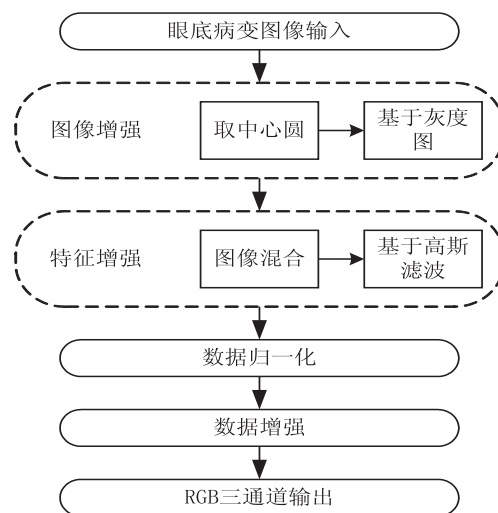


图 3 数据预处理流程

2.1 图像预处理

针对图像中消除“黑边”的影响问题^[11],该文采取裁剪的方法。首先取图像中心圆做掩模,保护眼球的有效区域,将图像转为灰度图,设定好灰度阈值,分别选三通道阈值以下区域进行裁剪,最后将三通道堆叠合成彩色图像,从而消除图像中的“黑边”。整个过程如图 4 所示。

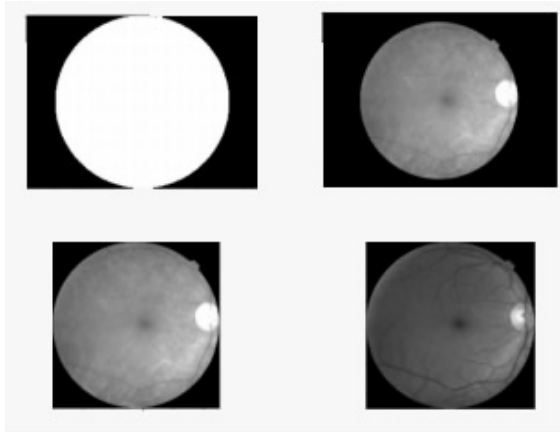


图 4 视网膜眼底图像裁剪过程

2.2 特征增强

针对消除图像中光照的影响,提升特征的对比度的问题,该文利用通过高斯滤波器后的眼底图像和原图像进行图像混合的方法,整个过程如图 5 所示。

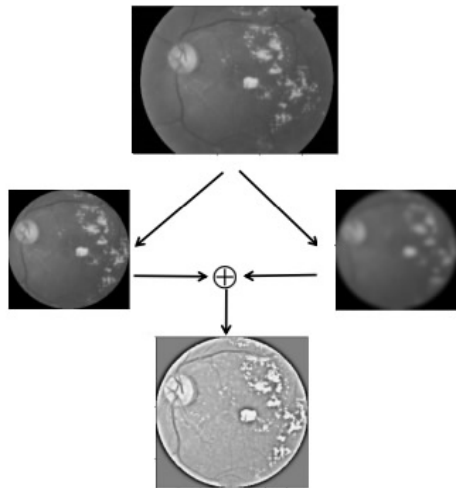


图 5 基于高斯滤波器的图像混合

从图 5 所示,可以明显地观察到一些图像特征如:斑点、渗出物、絮状物和血管等。说明在使用了该方法后,图像中特征和特征结构更加清晰,同时削弱了图像中光照的影响,保留了视神经的颜色这一特征。

2.3 数据增强及归一化

本次实验所用的 Kaggle 数据集样本图像有限且自身内部各病变程度等级的图像分布差异大,且图像分辨率大小不一,普遍很高。在此情况下,该文首先对准备输入神经网络的图像进行归一化,经实验调整图像分辨率为固定的 380×380 ,采用该尺寸能有效地提升模型的收敛速度,且尽量使原图像特征得以保留。同时,采用对图像依概率进行水平、左右翻转、旋转等数据增强方法来按比例地增加训练样本,以达到各病变等级分布均匀的目的,形成无偏数据集。该方法起到了提高神经网络模型的泛化能力,避免模型出现过拟合情况的作用。

3 基于复合模型缩放算法的分级模型

3.1 迁移学习

在现实生活中,数据集样本的收集受隐私等条件问题影响,导致可用于实验的病变样本很少且样本分级不均。同时,对数据样本的识别标记工作需要专业人员进行,导致已经标记的样本更少。由于上述原因,使用神经网络模型解决糖尿病视网膜病变分级问题前,对神经网络的训练非常容易导致模型过拟合,不能达到理想的分级效果。通常,面对已标记医疗图像资源少这一领域难题,相关专家使用迁移学习的方法来解决。该文采用在 Imagenet 项目上已经训练好的 Efficient-B4 模型预训练参数,再通过 Kaggle 数据集进行二次训练,这样使得模型有更好的初始化参数,显著提高模型训练效率,最终模型的泛化能力也会更好。

3.2 复合模型缩放算法

复合模型缩放算法^[12]的核心思想是神经网络模型的三大主要指标(深度、宽度、分辨率)并不是孤立的,其中有必然的联系。通过综合考虑三者之间的平衡可以使得神经网络模型有更好的表现,能让现有神经网络模型在不牺牲参数量和计算量的情况下,有效提升模型的准确率。

$$N = \bigodot_{i=1, \dots, s} F_i^{L_i}(X_{\langle H_i, W_i, C_i \rangle}) \quad (1)$$

$$\text{Accuracy}(N(d, w, r)) \quad (2)$$

$$N(d, w, r) = \bigodot_{i=1, \dots, s} \widehat{F}_i^{d_i, L_i}(X_{\langle r_i, \hat{H}_i, r_i, \hat{W}_i, w_i, \hat{C}_i \rangle}) \quad (3)$$

$$\text{Memory}(N) \leq \text{target}_{\text{memory}} \quad (4)$$

$$\text{FLOPS}(N) \leq \text{target}_{\text{flops}} \quad (5)$$

一个卷积神经网络可由式(1)表达,其中 X 为输入张量, F 为卷积操作,而其算法的优化问题可表达为式(2)~式(5),在现有基础模型的约束条件下,平衡基线模型的深度、宽度、分辨率三项指标,提升最佳准确率、控制参数量和计算量,达到优化模型的目的。

$$d = \alpha^\varphi \quad (6)$$

$$w = \beta^\varphi \quad (7)$$

$$r = \gamma^\varphi \quad (8)$$

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \quad (9)$$

$$\alpha \geq 1, \beta \geq 1, \gamma \geq 1 \quad (10)$$

在复合模型缩放算法中,整个算法在约束条件式(10)下,通过使用复合系数 φ (如式(6)~式(9))对模型统一进行缩放,其中 d 为神经网络的深度, w 为神经网络的宽度, r 为分辨率。

复合模型缩放算法同时可以应用于其他现有的分类模型中,其泛化效果十分不错,能够有效提高模型的效率。模型的缩放效率十分依赖于基线模型,该文采用谷歌团队开发的基线模型 EfficientNet-B4,可分为 9 个 stage,而每个 stage 由多个 MBConvBlock 组成,整个

神经网络模型由 1 个 Stem、32 个 MBCConvBlock 和最后连接层堆叠而成,其模型结构如表 1 所示。

表 1 EfficientNet-B4 神经网络模型结构

Stage i	Operator F_i	Resolution $H_i \times W_i$	Channels C_i
1	Conv3×3	380×380	48
2	MBCConvBlock	190×190	24
3	MBCConvBlock	95×95	32
4	MBCConvBlock	48×48	56
5	MBCConvBlock	24×24	112
6	MBCConvBlock	24×24	160
7	MBCConvBlock	12×12	272
8	MBCConvBlock	12×12	448
9	Conv&Pooling&FC	12×12	1 792

其中关键模块为 MBCConvBlock,又可以分为多个 Block,Block 的通用结构如图 6 所示。

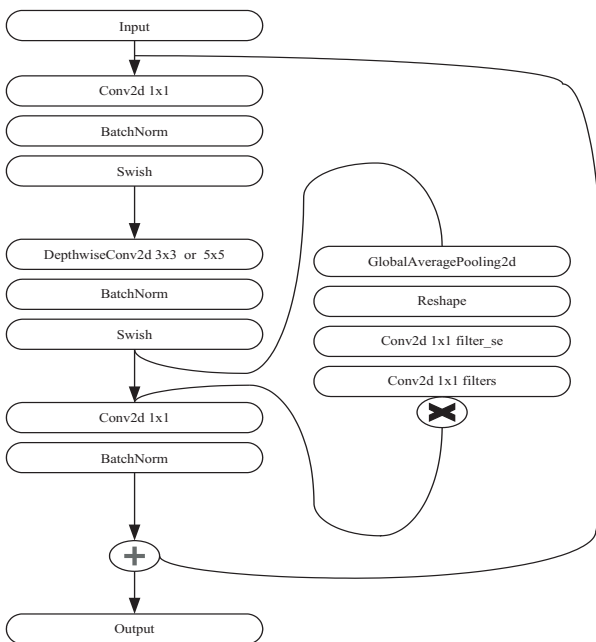


图 6 Block 通用结构

其 Block 模块结构借鉴了残差结构和反向残差结构,同时引用了注意力机制。首先利用卷积(conv2d)提升 channel,再进行深度可分离卷积(depthwiseconn2d)操作,之后通过卷积(conv2d)降低 channel,同时在该模块后增加了一个对通道的注意力机制,最后利用卷积(conv2d)降维后加上一个残差结构。其中整个过程使用 swish 激活函数。

4 结果与分析

4.1 实验设置

该文在 Pytorch 深度学习框架下进行实验,本次实验所采用的优化策略及超参数见表 2。为了尽量保留

图像的细节特征,综合 Efficient-B4 模型的最佳输入分辨率,该文采用 380 * 380 图像输入。

表 2 神经网络模型配置

方案及参数	名称
基础模型	Efficient-B4
损失函数	Cross-entropy Loss
优化方案	Adam
学习率方案	初始值为 0.05,等间隔衰减
Batch_size	32

4.2 评价标准

采用平均准确率 Accuracy 曲线来观察神经网络的过拟合情况,提前停止神经网络的训练。其计算方法如式(11)、式(12)所示,首先计算每个分级的准确率,再计算整体平均准确率。同时采用 Kappa 系数(式(13))来评估训练好的神经网络。

$$Accuracy = \frac{m_{correct}}{m_{total}} \quad (11)$$

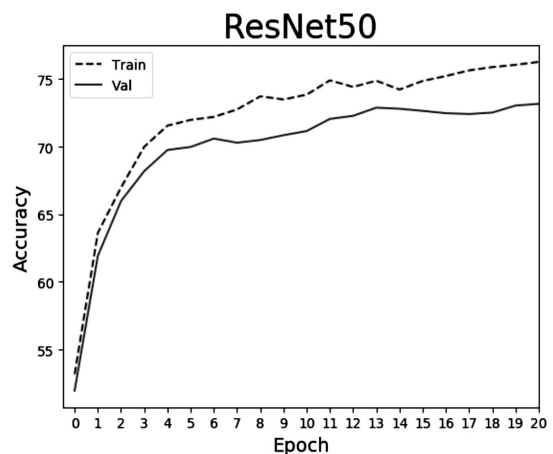
$$Accuracy_{avg} = \frac{\sum_{i=1}^N Accuracy_i}{N} \quad (12)$$

$$K = \frac{\rho_o - \rho_e}{1 - \rho_e} \quad (13)$$

其中平均准确率 Accuracy,能直接反映对病变程度正确分级的比率,但面对本数据集样本数量不均衡的情况,需要采用 Kappa 系数检验其一致性问题,当 Kappa 系数大于 0.81 时,说明其一致性强。

4.3 实验结果分析

为验证该文使用的方法和其他常规方法的分级效果,进行了两次模型训练,训练过程如图 7 所示。在相同条件下,一次使用了未预处理眼底图像对 ResNet 模型进行训练,另外使用了该文所研究的预处理方法后的眼底图像来训练 EfficientNet-B4 模型。两者就训练后进行结果对比。



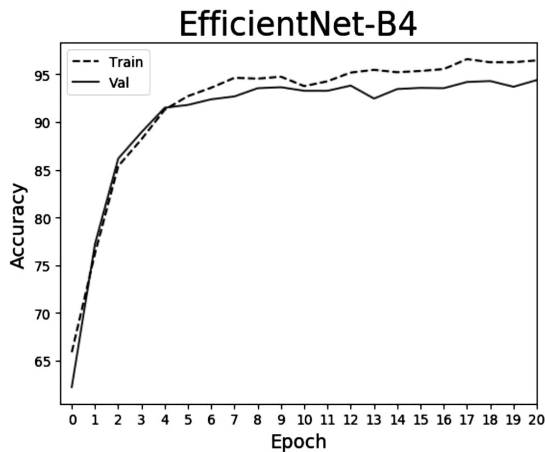


图 7 ResNet50 与 EfficientNet-B4 的 Accuracy 曲线

利用该文所研究的方法对糖尿病视网膜病变程度等级分级的最后平均准确率为 92.46%, Kappa 系数 0.88。结果表明,相对于采用传统的 K 近邻算法、支持向量机(SVM)的平均准确率更高^[13]。

同时,该文对比了不同卷积神经网络模型对糖尿病视网膜病变检测的准确率,如 LeNet、AlexNet 和 CompactNet 神经网络模型^[14-15],对比结果如表 3 所示。

表 3 神经网络模型对比

Models	Params/M	Accuracy	Kappa
LeNet	0.72	0.42	-
AleNet	60	0.62	-
CompactNet	14.16	0.69	-
ResNet50	25.57	0.71	0.68
文中模型	19	0.92	0.88

由表 3 可知,虽然各个神经网络模型所用数据样本来源一致,然而最终模型得到的准确率却大相径庭,说明所用的预处理方法和采用的神经网络模型对准确率的影响巨大。相对前人采用的前三个模型,ResNet50 网络模型引入残差模块,解决了由于模型网络过深带来的梯度弥散问题,但是没有结合网络深度、宽度、分辨率三者考虑。而该文采用了复合模型缩放算法,复合缩放后的模型更加倾向关注于具有更多对象细节的相关位置,对于视网膜病变区域及特征捕捉更好,最终结果表明,该方法平均准确率更高、泛化能力更好。

5 结束语

该文充分考虑了数据集特点(样本数少、样本标签分布不均、样本规格不统一)和数据特征(特征的种类、特征的分布、特征结构),采用了图像混合的预处理方法,利用了基于复合模型缩放算法的神经网络模

型,实现视网膜病变等级分类准确率达 92%,Kappa 系数 0.88。实验结果表明该方法对视网膜病变等级分级效果提升显著,可对糖尿病视网膜病变的诊断提供科学依据,同时对研究其他眼底病变图像也有一定参考价值。

后续工作将考虑不同级别间病变特征差异小的因素,导致少数处于分级界限边缘的样本分级时有偏差,考虑将目前多分类问题转化为多标签多分类问题。

参考文献:

- [1] 甘柳珠,邹凯,陈辉.糖尿病视网膜病变危险因素的研究进展[J].右江民族医学院学报,2015,37(5):743-744.
- [2] 钱菁.糖尿病性视网膜病变(DR)的早期综合防治研究[J].哈尔滨医药,2017,37(2):190-191.
- [3] 张丹,胡燃燃.普罗布考辅助视网膜光凝对增殖早期 DR 疗效及炎症递质的影响[J].基层医学论坛,2020,24(19):2693-2695.
- [4] 许迅.我国眼底病研究发展现状、进展和努力方向[J].中华眼科杂志,2014,50(11):801-803.
- [5] 肖璐璐,窦晓燕.人工智能在眼部疾病中的应用及其挑战[J].国际眼科杂志,2020,20(7):1197-1201.
- [6] 张婧妍.基于中医三消探究糖尿病视网膜病变眼底形态特征的差异性研究[D].沈阳:辽宁中医药大学,2020.
- [7] 张芳,赵东旭,肖志涛,等.眼底图像质量分类综述[J].计算机辅助设计与图形学学报,2020,32(3):501-512.
- [8] 范围,熊祥伟,白莲,等.DR 早期视网膜神经纤维层和角膜神经的变化及其相关性[J].国际眼科杂志,2017,17(4):772-774.
- [9] MO J,ZHANG L. Multi-level deep supervised networks for retinal vessel segmentation[J]. Int J Comput Assist Radiol Surg,2017,12(12):2181-2193.
- [10] YI Y, ADEL M, BOURENNANE S. Retinal vessel segmentation using a probabilistic tracking method[J]. Pattern Recognition,2012,45(4):1235-1244.
- [11] 杨帆,陈睿诗,莫阳,等.基于深度学习的视网膜病变眼底图视盘自动定位与分割研究[J].贵州医科大学学报,2020,45(4):432-437.
- [12] TAN Mingxing. EfficientNet: rethinking model scaling for convolutional neural networks[EB/OL]. (2019-05-28). <https://arxiv.org/abs/1905.11946>.
- [13] 潘燕红,潘林.基于 SVM 的眼底图像硬性渗出检测[J].计算机与现代化,2014(4):33-37.
- [14] 丁蓬莉,李清勇,张振,等.糖尿病性视网膜图像的深度学习神经网络分类方法[J].计算机应用,2017,37(3):699-704.
- [15] 李琼,柏正尧,刘莹芳.糖尿病性视网膜图像的深度学习分类方法[J].中国图象图形学报,2018,23(10):1594-1603.