



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象 图形学报

2020
08
VOL.25

ISSN1006-8961
CN11-3758/TB



三维建模和步态识别 P1539



第25卷第8期（总第292期）

2020年8月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院遥感与数字地球研究所

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 顾行发

编辑出版《中国图象图形学报》编辑出版委员会

通信地址 北京市海淀区北四环西路19号

邮 编 100190

电子信箱 jig@radi.ac.cn

电 话 010-58887035

网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司

（邮政信箱：北京399信箱 邮编：100048）

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing and Digital Earth, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Gu Xingfa

Editor, Publisher Editorial and Publishing Board of Journal of
Image and Graphics

Address No. 19, North 4th Ring Road West, Haidian District,
Beijing, P. R. China

Zip code 100190

E-mail jig@radi.ac.cn

Telephone 010-58887035

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China))

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

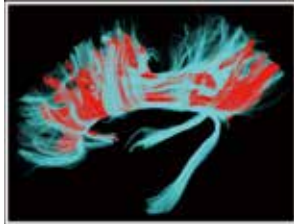
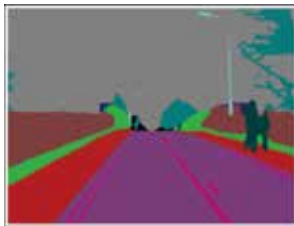
ISSN 1006-8961

CODEN ZTTXFZ

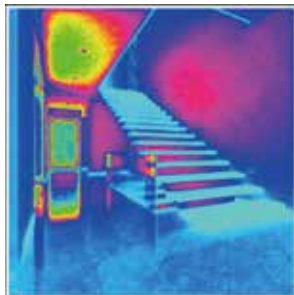
国外发行代号 M1406

国内邮发代号 82-831

国内定价 60.00元

神经纤维跟踪算法研究进展
(第1513页)

结合注意力机制的双路径语义分割(第1627页)



突变策略下多通路Metropolis光照传播(第1658页)

综述

神经纤维跟踪算法研究进展

李茂, 何建忠, 冯远静 1513

多模态生物特征提取及相关性评价综述

杨雪鹤, 刘欢喜, 肖建力 1529

图像分析和识别

异常步态3维人体建模和可变视角识别

罗坚, 黎梦霞, 罗诗光 1539

扩展点态卷积网络的点云分类分割模型

张新良, 付陈琳, 赵运基 1551

特征一致性约束的视频目标分割

征煜, 陈亚当, 郝川艳 1558

增量角度域损失和多特征融合的地标识别

毛雪宇, 彭艳兵 1567

部件检测和语义网络的细粒度鞋类图像检索

陈前, 刘骊, 付晓东, 刘利军, 黄青松 1578

图像理解和计算机视觉

图注意力网络的场景图到图像生成模型

兰红, 刘秦邑 1591

跨层多模型特征融合与因果卷积解码的图像描述

罗会兰, 岳亮亮 1604

Haar-like特征双阈值Adaboost人脸检测

刘禹欣, 朱勇, 孙结冰, 王一博 1618

结合注意力机制的双路径语义分割

翟鹏博, 杨浩, 宋婷婷, 余亢, 马龙祥, 黄向生 1627

自学习规则下的多聚焦图像融合

刘子闻, 罗晓清, 张战成 1637

车路视觉协同的高速公路防碰撞预警算法

蔡创新, 高尚兵, 周君, 黄子赫 1649

计算机图形学

突变策略下多通路Metropolis光照传播

贺怀清, 赵煜桢, 刘浩翰, 王旭 1658

融合DNN与AABB-圆形包围盒自碰撞检测

靳雁霞, 程琦甫, 张晋瑞, 齐欣, 马博, 贾瑶 1674

医学图像处理

压缩感知下溃疡性结肠炎辅助诊断

杨海清, 孙道洋 1684

特征重用和注意力机制下肝肿瘤自动分类

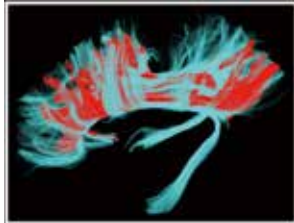
冯诺, 宋余庆, 刘哲 1695

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

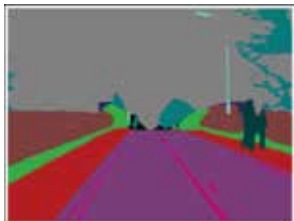
张子振, 刘明, 朱德江 1708

CONTENTS

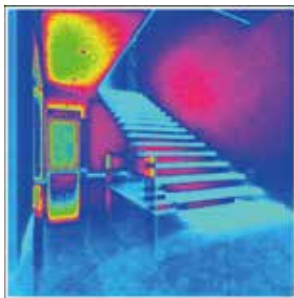
JOURNAL OF IMAGE AND GRAPHICS



Research progress of neural fiber tracking(P1513)



Two-path semantic segmentation algorithm combining attention mechanism(P1627)



Improved multiplexed Metropolis light transport based on fusion mutation strategy(P1658)

Review

Research progress of neural fiber tracking

Li Mao, He Jianzhong, Feng Yuanjing 1513

Extraction and relevance evaluation for multimodal biometric features

Yang Xuehe, Liu Huanxi, Xiao Jianli 1529

Image Analysis and Recognition

Parametric 3D body modeling and view-invariant abnormal gait recognition

Luo Jian, Li Mengxia, Luo Shiguang 1539

Extended pointwise convolution network model for point cloud classification and segmentation

Zhang Xinliang, Fu Chenlin, Zhao Yunji 1551

Video object segmentation algorithm based on consistent features

Zheng Yu, Chen Yadang, Hao Chuanyan 1558

Landmark recognition based on ArcFace loss and multiple feature fusion

Mao Xueyu, Peng Yanbing 1567

Fine-grained shoe image retrieval by part detection and semantic network

Chen Qian, Liu Li, Fu Xiaodong, Liu Lijun, Huang Qingsong 1578

Image Understanding and Computer Vision

Image generation from scene graph with graph attention network

Lan Hong, Liu Qinyi 1591

Image caption based on causal convolutional decoding with cross-layer multi-model feature fusion

Luo Huilan, Yue Liangliang 1604

Improved Adaboost face detection algorithm based on Haar-like feature statistics

Liu Yuxin, Zhu Yong, Sun Jiebing, Wang Yibo 1618

Two-path semantic segmentation algorithm combining attention mechanism

Zhai Pengbo, Yang hao, Song Tingting, Yu Kang, Ma Longxiang, Huang Xiangsheng 1627

Multi-focus image fusion with a self-learning fusion rule

Liu Ziwen, Luo Xiaoqing, Zhang Zhancheng 1637

Freeway anti-collision warning algorithm based on vehicle-road visual collaboration

Cai Chuangxin, Gao Shangbing, Zhou Jun, Huang Zihe 1649

Computer Graphics

Improved multiplexed Metropolis light transport based on fusion mutation strategy

He Huaqing, Zhao Yuzhen, Liu Haohan, Wang Xu 1658

Self-collision detection algorithm based on fused DNN and AABB-circular bounding box

Jin Yanxia, Cheng Qifu, Zhang Jinrui, Qi Xin, Ma Bo, Jia Yao 1674

Medical Image Processing

Adjunctive diagnosis of ulcerative colitis under compressed sensing

Yang Haiqing, Sun Daoyang 1684

Automatic classification of liver tumors by combining feature reuse and attention mechanism

Feng Nuo, Song Yuqing, Liu Zhe 1695

Automatic recognition and classification of diabetic retinopathy images by combining an attention mechanism and an efficient network

Zhang Zizhen, Liu Ming, Zhu Dejiang 1708

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2020)08-1708-11

论文引用格式: Zhang Z Z, Liu M and Zhu D J. 2020. Automatic recognition and classification of diabetic retinopathy images by combining an attention mechanism and an efficient network. Journal of Image and Graphics, 25(08): 1708-1718 (张子振, 刘明, 朱德江. 2020. 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. 中国图象图形学报, 25(08): 1708-1718) [DOI: 10.11834/jig.190644]

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

张子振, 刘明, 朱德江

1. 河北大学电子信息工程学院, 保定 071000; 2. 河北省数字医疗工程重点实验室, 保定 071000

摘要: **目的** 糖尿病视网膜病变(diabetic retinopathy, DR)是一种病发率和致盲率都很高的糖尿病并发症。临床中,由于视网膜图像不同等级之间差异性小以及临床医生经验的不同,会出现误诊、漏诊等情况,目前基于人工 DR 的诊断分类性能差且耗时费力。基于此,本文提出一种融合注意力机制(attention mechanism)和高效率网络(high-efficiency network, EfficientNet)的 DR 影像自动分类识别方法,以此达到对病变类型的精确诊断。**方法** 针对实验中 DR 数据集存在的问题,进行剔除、去噪、扩增和归一化等处理;利用 EfficientNet 进行特征提取,采用迁移学习的策略用 DR 的数据集对 EfficientNet 进行学习训练,提取深度特征。为了解决病变之间差异小的问题,防止网络对糖尿病视网膜图像的特征学习时出现错分等情况,在 EfficientNet 输出结果上加入注意力机制;根据网络提取的特征在深度分类器中进行分类,将视网膜图像按等级进行五分类。**结果** 本文方法的分类精度、敏感性、特异性和二次加权(kappa)值分别为 97.2%、95.6%、98.7% 和 0.84,具有较好的分类性能及鲁棒性。**结论** 基于融合注意力机制的高效率网络(attention EfficientNet, A-EfficientNet)的 DR 分类算法有效地提高了 DR 筛查效率,解决了人工分类的手动提取特征的局限性,在临床上对医生诊断起到了辅助作用,能更有效地防治此类恶性眼疾造成严重视力损伤,甚至失明。

关键词: 糖尿病视网膜病变分类;高效率网络;注意力机制;深度学习;迁移学习;深度特征

Automatic recognition and classification of diabetic retinopathy images by combining an attention mechanism and an efficient network

Zhang Zizhen, Liu Ming, Zhu Dejiang

1. College of Electronic and Information Engineering, Hebei University, Baoding 071000, China;
2. Key Laboratory of Digital Medical Engineering of Hebei Province, Baoding 071000, China

Abstract: Objective Diabetic retinopathy (DR) is a diabetic complication with high incidence and blindness rate. The diagnosis of DR through color fundus images requires experienced clinicians to identify the presence and significance of many small features. This requirement, along with a complex grading system, makes the diagnosis a difficult and time-consuming task. The small difference among various grades of retinal images and the limited experience of clinicians usually lead to misdiagnosis and missed diagnosis. The large population of diabetic patients and their massive screening require-

收稿日期: 2019-12-11; 修回日期: 2020-01-19; 预印本日期: 2020-01-26

基金项目: 国家自然科学基金项目(61473112); 国家青年基金项目(61703133); 河北省青年基金项目(F2018201070); 河北省青年拔尖人才项目(BJ2019044); 河北省科技支撑计划重点项目(F2017201222); 河北大学研究生创新资助项目(hbu2019ss069)

Supported by: National Natural Science Foundation of China (61473112)

ments have generated interest in a computer-aided and fully automatic diagnosis of DR. The present diagnostic classification of artificial DR is poor, time consuming, and laborious. On this basis, an automatic classification method for DR with a high-efficiency network (EfficientNet) incorporating an attention mechanism (A-EfficientNet) is proposed in this study. The goal of accurate classification is achieved by using the strategy of transfer learning to guide the classification model to learn the features among lesions. **Method** The characteristics of DR dataset present that noise in images and labels is unavoidable in raw dataset. This condition indicates a high demand for the robustness of our classification system. Preprocessing of retinal images mainly includes retinal image culling, noise reduction, enhancement, and normalization of retinal images. The images after such preprocessing will improve the performance of the classification network, and the computing power of the computer will also be accelerated. We use EfficientNet as the basic network for feature extraction. The migration-learning strategy is adopted to learn and train A-EfficientNet with DR datasets. We design an attention mechanism in EfficientNet to address the problem of small differences among lesions in the dataset and prevent the network from misclassifying fine-grained features in retinal image categories. The mechanism not only can extract features from fundus images but also can focus on the lesion area. The EfficientNet classification model integrated with the attention mechanism (A-EfficientNet) can match images with image label categories to achieve the classification task of retinal images. With fully connected layers and softmax, the model is used to learn for classifying DR status as DRfree, mild DR, moderate DR, severe DR, and proliferative DR. **Result** Experimental results show that, in the case of insufficient samples of retinal images, the use of transfer learning and data enhancement strategies can help the classification model extract deep features for classification. A-EfficientNet can learn the features in the training samples and pay attention to the differences among various categories of features to achieve accurate classification. The reference to the attention mechanism greatly improves the performance of the classification network. In summary, a high-efficiency network (EfficientNet) incorporating an attention mechanism is proposed in this study. The network plays a positive role in extracting lesion features at various stages of images. The classification accuracy of the model for lesions reaches 97.2%. To improve the classification efficiency, the results of the two-category model indicates sensitivity of 95.6% on the high-specificity operating point and specificity of 98.7% on the high-sensitivity operating point. The result shows that the model achieves a kappa score of 0.840, which is higher than that of the existing non ensemble model. Therefore, the classification model based on the attention mechanism and EfficientNet can accurately distinguish the lesion types of retinopathy. Transfer-learning strategies and data augmentation provide accurate information and assistance for the accurate classification of fundus images. A-EfficientNet reflects the good classification performance of the model and the robustness of the network-learning ability. **Conclusion** In this study, we use A-EfficientNet to realize automatic classification of retinal images. We also propose a new classification framework for DR, which mainly benefits from four parts, namely, the image-preprocessing stage, the transfer-learning strategy, the introduction of data enhancement, and A-EfficientNet for feature extraction. The dataset used in the experiment has problems, such as noise, artifacts, and unsuccessful focusing. After such problematic data are removed, the dataset is denoised and normalized to facilitate the classification network for calculating and learning the data. When the training samples are insufficient, the introduction of data enhancement meets the data requirements of the deep learning classification model, satisfies the model's learning of the data, and improves the generalization ability of the model. The introduction of an attention mechanism based on EfficientNet can improve the classification performance of the algorithm. Therefore, the automatic classification algorithm for DR incorporating the attention mechanism of EfficientNet can effectively improve the efficiency of DR screening, avoid the limitations of artificial feature extraction and image classification, and solve the problem of over fitting caused by insufficient sample data. The experimental results show that this method can be used to diagnose DR. It can operate efficiently without marking the information of suspicious lesions and thus avoids the time-consuming work of labeling lesions by experts and reduces false positives. The classification model can also be referenced in the classification tasks of other images.

Key words: mutil-classification of diabetic retinopathy; high-efficiency network (EfficientNet); attention mechanism; deep learning; transfer learning; deep features

0 引言

糖尿病视网膜病变(diabetic retinopathy, DR)是由于长期患有糖尿病引起的一种眼底视网膜并发症(Laud 等,2017)。据世界卫生组织统计,全球约有4.25亿人患有糖尿病,其中约1/3的糖尿病患者发生了糖尿病视网膜病变。据研究显示(Flaxman 等,2017),DR的早期阶段几乎没有迹象或症状,致使许多患者错过了治疗的最佳时间。此外,由于DR检查人数过多,给医生带来诊断压力,导致病变诊断得不到及时反馈,从而出现漏诊、误诊等问题。因此,DR的早期发现和及时治疗至关重要。糖尿病视网膜病变在不同阶段会出现不同的病理特征,最终会对眼部造成损伤而导致失明。计算机辅助自动诊断在临床上具有很大潜力,可以在短时间内准确检测DR,帮助医生提高DR的筛查率并减少失明的情况。

随着研究学者对糖尿病视网膜病变的关注,已开发了多种自动诊断系统。由于眼科医生在评估眼底照片时往往会关注与DR相关的一些典型病变,如微动脉瘤、出血和硬性渗出,所以许多研究主要集中在自动检测和分割这些病变区域或计算一些数值指标。Shahin 等人(2012)开发了一种系统,将视网膜眼底图像自动分类为是否有增值性糖尿病视网膜病变,采用形态学处理方法提取血管面积和渗出区等病理特征,计算熵和同质性两个指标,将这些特征输入浅层神经网络,获得了88%的灵敏度和100%的特异性。Casanova 等人(2014)提出一种利用随机森林的算法区别是否为DR患者,准确率超过90%,并根据分级的眼底照片和系统数据评估DR风险。然而,这些方法都依赖临床眼科医生进行人工标注,需要花费大量精力用图像处理方法对手工制作的特征进行手动提取,使模型的分类任务变得更加复杂和不稳定。

深度学习在视网膜图像检测和分类任务上取得了显著成果(Krizhevsky 等,2017;李琼 等,2018;Voets 等,2018;Li 等,2017;Gargeya 和 Leng,2017;Wang 等,2018;Graham,2015;Doshi 等,2016;Zhou 等,2018)。Gulshan 等人(2016)采用InceptionV3深度模型,利用包括12.8万多幅眼底图像的数据集检测DR,得益于大量的训练数据和眼底专家对眼底

图像的筛选,在两个不同测试集上获得的AUC(area under curve)值非常高。Li 等人(2017)采用迁移学习的方法对DR进行二分类检测,将预训练的卷积神经网络(convolutional neural network,CNN)模型进行微调,在对视网膜图像特征提取的基础上,通过支持向量机对提取的特征进行训练分类,在MESSIDOR(methods for evaluating segmentation and indexing techniques dedicated to retinal ophthalmology)(French Ministry,2014)和DRiDB数据库上进行验证,实现了二分类任务。Gargeya 等人(2017)利用残差网络(ResNet)(He 等,2016)加决策树分类器训练二分类模型来区分患病和正常健康的图像,在MESSIDOR-2等数据集上进行了验证。Wang 等人(2018)用DenseNet121网络进行特征提取,通过机器学习提升树算法进行预测,经过数据集测试,验证了模型具有更好的性能。Graham(2015)等在Kaggle(2015)竞赛上对DR分类取得了不错的结果。

然而,以上均属糖尿病视网膜病变的二分类方法,虽然取得的分类结果较好,但是没有深入到具体病变分类任务中,对临床医生的诊断很难起到准确的预测作用。

糖尿病视网膜病变有5个等级的演变过程,由于病理图像类别之间的特点,每个阶段的差异很小,因此,在不需要人工标注的情况下对DR进行自动的五分类任务具有一定的挑战性。

目前,对DR进行五分类检测对临床有积极的辅助作用。Doshi 等人(2016)利用深度卷积神经网络训练数据集,实现了DR的诊断和五分类任务,并在Kaggle数据集上进行了验证。由于采用GPU(graphics processing unit)加速训练,使得训练速度得到提升,但是效果没有得到显著提高。Zhou 等人(2018)利用多单元结构对高像素的眼底影像进行多任务学习,通过分类和回归的形式预测标签,并在EyePACs数据集上验证了该方法的有效性,但是仅在高分辨率图像上有明显效果,对分辨率低的图像无法精确识别与分类。Pratt 等人(2016)利用数据扩增和卷积神经网络在Kaggle数据集上进行了DR的五分类,取得了一定进步。

基于上述分析,为了提高对DR的分类准确性和高效性,本文提出一种融合注意力机制的高效率网络(attention EfficientNet, A-EfficientNet)的分类

算法对眼底图像进行自动分类,利用 EfficientNet 网络(Tan 和 Le,2019)通过优化网络宽度、网络深度和增大分辨率来达到提升指标的的优点,在达到现有分类网络准确率的情况下,大幅度地减少模型参数数量和计算量,提高了模型的泛化能力。特别地,在网络训练过程中加入注意力机制,根据不同病变的注意力映射与输入的特征映射做乘操作来对提取的特征进行自适应学习,从而对眼底细粒度图像病变之间的差异性起到了关注作用,抑制其他无关信息对网络学习的影响。

1 本文方法

基于融合注意力机制和高效率网络的糖尿病视网膜病变分类模型如图 1 所示,通过对糖尿病视网膜病变图像进行数据预处理和数据增强等操作来提高数据量,以此完成分类模型的特征提取与分类。

1.1 迁移学习

迁移学习是机器学习的一种方法,是将一个任务训练得到的模型移植到其他任务的训练上来。受迁移学习(Pan 和 Yang,2010)的影响,在训练数据不足的情况下,通过加载在 ImageNet 数据集上预训练的 EfficientNet 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。考虑到眼底图像数据集和 ImageNet 数据集之间的巨大差异,实验过程中对网络层的训练是从每一层重新开始训练的。

1.2 高效率网络

对糖尿病视网膜病变的图像进行预处理后,利用高效率网络对预处理图像进行训练。增大网络深度(Huang 等,2018)是训练许多神经网络经常使用的方法,因为这样能捕捉更丰富、更复杂的特征并且适应新任务来进行学习。然而,增加网络的深度会带来梯度消失的问题。网络的宽度即特征图的通道数,增加网络宽度(Zagoruyko 和 Komodakis,2016),即特征图通道数增多,更多的卷积核可以得到更多丰富的特征,增强了网络的表征能力。小尺寸模型对网络的宽度要求小,更宽的网络往往能够学习到更加丰富的特征,并且很容易训练。但是对于网络结构过宽且深度较浅的网络在特征提取过程中很难学习到更高层次的特征。卷积神经网络对于具有高分辨率的输入图像也可以捕捉细粒度特征,这样能

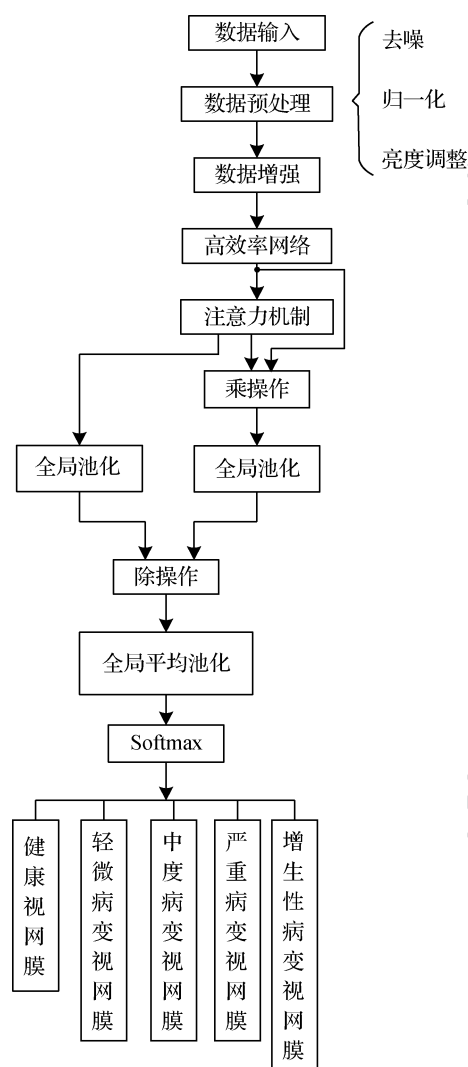


图1 融合注意力机制的高效率分类模型

Fig.1 Classification network combining attention mechanism and EfficientNet model

丰富网络的感受野来提升网络。上述网络的宽度、深度及图像的分辨率3个指标都可以提高精度,但对于较大的模型,精确度会降低,由于数据集中的图像具有高分辨率,对于高分辨率的图像,通过增加网络深度可以帮助网络在训练过程中学习到更多与像素信息相似的特征。相应地,当分辨率更高时,可以增加网络宽度,目的是在高分辨率图像中捕获更多的图像维度特征。这些问题表明,需要协调和平衡不同维度之间的关系,而不是常规的单维度缩放。EfficientNet 成功地将网络宽度、深度及提高图像的分辨率通过缩放系数对分类模型进行3个维度的缩放,自适应地优化网络结构。这样在训练过程中,大幅减少了训练参数,降低了计算复杂度。

本文将 EfficientNet 用于糖尿病视网膜病变的特征提取,网络计算为

$$N = \bigotimes_{i=1,2,\dots,s} F^{L_i}(X_{[H_i, W_i, C_i]}) \quad (1)$$

式中, N 表示分类网络, $\bigotimes$ 代表卷积操作, X 表示输入张量, F 表示基础网络层, i 表示卷积层数, L_i 表示网络的深度。该网络通过调整 3 个维度(高度 H , 宽度 W , 通道数 C) 进行优化, 需要找到最优的 3 个维度的缩放参数。以此在满足模型参数和计算量达到最大化时, 使模型的准确率有所提高。模型的最大精度记作 $Acc_{\max}(N(d, w, r))$, 具体公式如式(2)。

$$N(d, w, r) = \bigotimes_{i=1,2,\dots,s} \hat{F}^{d \times L_i}(X_{[r \times \hat{H}_i, r \times \hat{W}_i, w \times \hat{C}_i]}) \quad (2)$$

式中, 深度 $d = \alpha^e$, 宽度 $w = \beta^e$, 分辨率 $r = \gamma^e$ 。其中变量 α 、 β 、 γ 的关系为

$$\alpha^2 \times \beta^2 \times \gamma^2 \approx 2, \alpha \geq 1, \beta \geq 1, \gamma \geq 1 \quad (3)$$

为了得到能满足式(2)的 3 个维度参数, 使用复合参数 φ 对网络的深度、宽度以及分辨率进行优化。首先固定 $\varphi = 1$, 然后通过网格搜索找到满足式(3)的最优 α 、 β 、 γ 参数, 经过实验调整得出 $\alpha = 2.2$, $\beta = 1.6$, $\gamma = 1.16$ 。

通过式(1)–(3), 利用 EfficientNet 网络进行图像的特征提取, 能够将眼底图像的特征进行多维度融合。

1.3 注意力机制

医学图像往往含有许多无关信息, 可能会干扰决策。在糖尿病视网膜病变五分类任务中, 像小动脉瘤这样的微小特征、渗出物等类别之间差异小的细粒度图像对于医生的 DR 分级是至关重要的。加入注意力机制后, 在进行特征提取时能够让网络自主选择需要关注的区域, 这样在学习的过程中, 能将糖尿病视网膜图像的病变类型进行关注与学习, 可以对模型提取的特征进行选择, 以此获取更多具有描述性的信息, 这样能模仿临床医生专注于关键特征, 对临床诊断起到辅助作用。注意力机制网络主要是把 EfficientNet 的输出特征映射 F 作为输入, 然后放入含有 3 个 1×1 的卷积层中并且加入激活函数 ReLU 和 Sigmoid, 将输入转换成非线性特征。具体网络如图 2 所示。

加入该网络的目的就是生成注意力映射 A 。将特征映射 F 与注意力映射 A 相乘, 目的是生成图

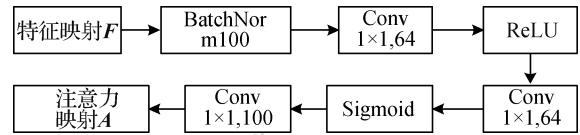


图2 注意力机制

Fig. 2 Attention mechanism

像的掩码 M 。为了减少网络的参数及避免过拟合, 在图像掩码 M 和注意力映射 A 上分别使用全局平均池化(global average pooling, GAP)。最后, 使用除法操作获得图像的权重, 并过滤掉不相关的信息。注意力机制的输出为

$$O = GAP(A^l) / GAP(A^l \times F^l) \quad (4)$$

式中, A^l 和 F^l 分别代表第 l 层注意力映射和第 l 层特征映射, $GAP()$ 为全局平均池化操作。

1.4 训练算法

为了防止网络拟合和提高测试集的预测性能, 在全连接层加入 dropout (Srivastava 等, 2014)。该方法在训练过程中能随机剔除网络中的单元, 减少局部特征带给网络结果的误差, 使模型的泛化性更强。同时, 在模型中应用早停法和微调, 缩短训练和调试时间, 提高模型的性能。此外, 在训练深度学习模型时, 优化器的选择也很重要。一个好的优化器能显著加快训练进程, 同时避免结果出现局部优化。本文采用具有良好自适应学习率的 Adam 优化器 (Kingma 和 Ba, 2014) 来对分类模型进行深度优化。

在 Adam 优化器中, 损失函数 L 的梯度为 g , 1 阶和 2 阶矩估计分别为 s 和 r , 在设置的固定学习步长 ε 下进行最优解的优化, 具体推导为

$$g \leftarrow + \frac{1}{m} \nabla_{\theta} \sum_i L(f(x_i; \theta), y_i) \quad (5)$$

$$s \leftarrow \rho_1 s + (1 - \rho_1) g \quad (6)$$

$$r \leftarrow \rho_2 r + (1 - \rho_2) g \otimes g \quad (7)$$

$$\hat{s} \leftarrow \frac{s}{1 - \rho_1 t} \quad (8)$$

$$\hat{r} \leftarrow \frac{r}{\rho_2} \quad (9)$$

$$\Delta \theta = - \varepsilon \frac{\hat{s}}{\sqrt{\hat{r}} + \delta} \quad (10)$$

$$\theta \leftarrow \theta + \Delta \theta \quad (11)$$

式中, θ 表示参数变量, m 为样本总数, 函数 f 表示网络的输出, x_i 和 y_i 分别代表函数 f 的输入和输出, ρ_1 和 ρ_2 为超参数。具体设置为 $\rho_1 = 1.0$, $\rho_2 =$

0.999, $\varepsilon = 0.001$, $\delta = 10^{-8}$ 。

2 数据集

2.1 数据及预处理

实验使用 Kaggle (2015) 和 MESSIDON (French Ministry, 2014) 数据集。Kaggle 数据集来自免费提供视网膜筛选的平台 EyePacs, 由多家医院提供, 包括了在各种成像条件下拍摄的几万幅高分辨率彩色眼底图像。MESSIDOR 数据是来自法国国防研究部

建立的眼底筛查项目, 数据图像由 3 个不同的眼科机构提供。

根据病变程度将视网膜图像分为 5 类, 如图 3 所示, 依次为健康、轻微病变、中度病变、严重病变和增生性病变的视网膜图像, 但是某些图像中包含伪影、失焦以及曝光不足或过度等情况, 这些质量差的无意义图像 (如图 4 所示) 都会影响图像中的像素强度值, 并产生与分类级别无关的不必要变化。因此, 为了规范这些图像, 减少冗余信息和环境伪影, 对眼底图像进行预处理是必要的。

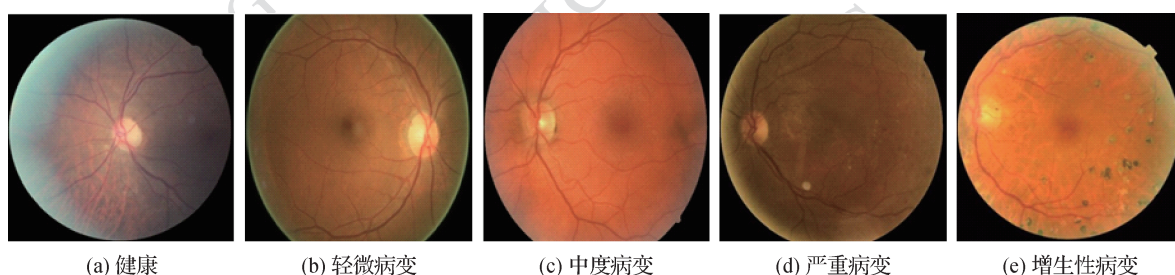


图3 糖尿病视网膜图像

Fig. 3 Diabetic retina images ((a) no DR; (b) mild DR; (c) moderate DR; (d) sever DR; (e) proliferative DR)

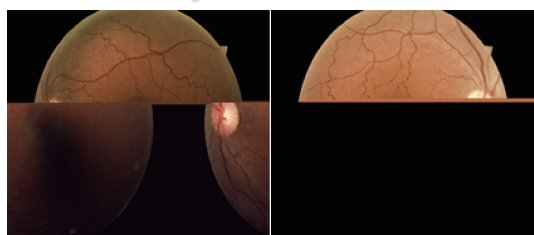


图4 糖尿病视网膜病变无意义图像

Fig. 4 Diabetic retinopathy meaningless images

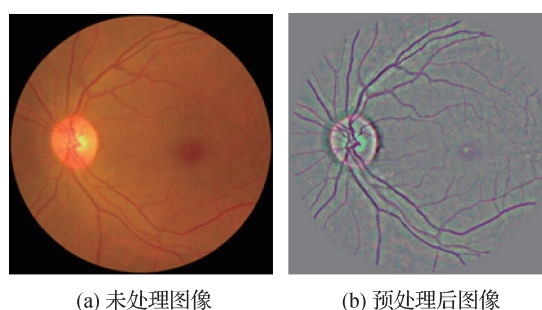
2.2 归一化

在对图像进行预处理时, 首先要剔除质量差的图像, 其次, 因为图像周边有黑色背景区域, 这一部分图像信息对训练网络的精度没有任何作用, 因此有必要将黑色边缘去掉。最后, 针对亮度不一致情况, 对图像进行亮度均衡和对比度均衡化处理, 具体为

$$I_i(x, y; \sigma) = \alpha(I(x, y) - G(x, y; \sigma) * I(x, y)) + \beta \quad (12)$$

式中, $G(x, y; \sigma)$ 代表高斯平滑函数, σ 是标准差, 用来评估图像的背景, 以此来提高对比度, β 是强度, 增强后大多数图像的像素值在 $[0, 255]$ 。在实验调试过程中这些参数的值设置为 $\alpha = 4$, $\sigma = r/30$, $\beta = 128$ 。

处理前后的图像如图 5 所示。由于原始数据的视网膜图像尺寸太大, 按照网络训练图像的大小要求, 将输入图像都设置为 456×456 像素。



(a) 未处理图像 (b) 预处理后图像

图5 预处理前后的糖尿病视网膜图像

Fig. 5 Diabetic retinal images before and after pretreatment
((a) unprocessed image; (b) preprocessed image)

2.3 数据增强

数据集中每种类别数据数量如表 1 所示, 可以看出, 数据类别存在严重不均衡问题, 无法高精度地解决医学图像的识别问题。因此, 除了预处理步骤外, 利用数据增强来大量增加训练数据, 以提高模型的泛化性能。对每幅图像均随机旋转 $0 \sim 90^\circ$ 、随机水平和垂直翻转、随机水平和垂直移动, 增加样本量少的类别数量, 解决样本不均衡问题。

表1 糖尿病视网膜病变分级数量

Table 1 Number of grades of diabetic retinopathy

类别	DR 程度	数量/幅	所占比例/%
0	健康视网膜	25 810	73.5
1	轻微病变视网膜	2 443	7.0
2	中度病变视网膜	5 292	15.1
3	严重病变视网膜	873	2.5
4	增生性病变视网膜	708	2.0

3 实验结果与分析

3.1 实验设置

实验在 Intel 2667V4CPU、32 GB 内存、NVIDIA Tesla K80 显卡的 Ubuntu16.04 服务器上进行调试,使用的模型框架为 tensorflow,分类网络的参数配置如表2所示。

表2 分类网络的参数配置

Table 2 Parameter configuration of classification network

名称	类型	输出大小($H \times W \times C$)
Input	input_3	$456 \times 456 \times 3$
Model	efficientnet	$15 \times 15 \times 2\ 048$
BatchN	batch_normalization_234	$15 \times 15 \times 2\ 048$
Dropout	dropout_4	$15 \times 15 \times 2\ 048$
Conv2D	conv2d_316	$15 \times 15 \times 64$
Conv2D	conv2d_317	$15 \times 15 \times 16$
Conv2D	conv2d_318	$15 \times 15 \times 8$
Conv2D	conv2d_319	$15 \times 15 \times 1$
Conv2D	conv2d_320	$15 \times 15 \times 2\ 048$
Multiply	multiply_80	$15 \times 15 \times 2\ 048$
Glo	global_average_pooling2d_3	$1 \times 1 \times 2\ 048$
Glo	global_average_pooling2d_4	$1 \times 1 \times 2\ 048$
Lambda	rescaleGAP	$1 \times 1 \times 2\ 048$
Dropout	dropout_5	$1 \times 1 \times 2\ 048$
Dense	dense_3	$1 \times 1 \times 128$
Dropout	dropout_6	$1 \times 1 \times 128$
Dense	dense_4	$1 \times 1 \times 5$

目前,糖尿病视网膜分类通常采用 Inception V3 网络(Zeng 等,2019;庞浩和王枫,2017;Rakhlin,

2017)进行数据训练。为了更全面地对比试验结果,本文设计了4组实验,分别利用 InceptionV3 和 EfficientNet 网络及 InceptionV3 和 EfficientNet 网络加注意力机制训练网络。将 Kaggle 数据集的 35 126 幅图像和 MESSIDOR 数据集的 1 200 幅图像(Decencière 等,2014)用于训练,且每幅图像定义一个标签,按每幅图像的病变类型,选择 $\{0,1,2,3,4\}$ 中的一个。将其中 7 025 幅图像作为验证集,5 096 幅图像作为测试集,剩下的图像用作训练集。为了验证模型的有效性,对实验数据做了二分类实验,检测是否患病的有效性。

3.2 实验指标

实验选用特异性(specificity, SP)、敏感性(sensitivity, SE)、准确率(accuracy, AC)、混淆矩阵(confusion matrix)和 kappa 系数等指标对结果进行评价,计算为

$$SP = \frac{TN}{TN + FP} \quad (13)$$

$$SE = \frac{TP}{TP + FN} \quad (14)$$

$$AC = \frac{TP + TN}{TP + FP + FN + TN} \quad (15)$$

式中, TP 、 FP 、 FN 、 TN 分别代表真阳性、假阳性、假阴性、真阴性。

混淆矩阵是评价分类模型对类别细分的一种可视化结果,横轴为模型预测的类别数量统计,纵轴为真实数据标签的数量统计。

本文实验采用二次加权 kappa 系数(Fleiss 等,1969),通过将模型预测的结果与医生提供的标注类别标签比较来评价模型分类的一致性。具体计算为

$$\kappa = 1 - \frac{\sum_{i,j} w_{i,j} o_{i,j}}{\sum_{i,j} w_{i,j} ((\sum_j o_{i,j}) \times (\sum_i o_{i,j})) / n} \quad (16)$$

式中, $\mathbf{o}$ 是 5×5 矩阵; n 表示样本个数; $o_{i,j}$ 代表真实的 i 类错分为 j 类的数量; $w_{i,j}$ 为加权系数,通过加权惩罚错分项来提高分类效果,加权系数计算为

$$w_{i,j} = \frac{(i-j)^2}{(N-1)^2} \quad (17)$$

3.3 实验结果

本文采用混淆矩阵对 Softmax 的分类结果进行

统计评估,根据式(13)~(17)分别计算糖尿病视网膜病变的五分类和二分类的各项评估指标,得到的五分类的混淆矩阵如表3所示。可以看出,大部分数据都集中在对角线上,二次加权 kappa 值为 0.84,验证了模型分类的一致性。对于二分类任务,特异性为 98.7%,敏感性为 95.6%,准确率为 97.2%。

表3 融合注意力机制的 EfficientNet 分类混淆矩阵

Table 3 Classification confusion matrix incorporating EfficientNet and attention mechanism

真实标签	预测类别				
	0	1	2	3	4
0	4 535	240	312	3	6
1	117	302	59	3	7
2	287	201	449	70	51
3	4	3	32	90	45
4	0	1	41	16	83

注:加粗字体表示各列最优结果。

3.4 实验对比分析

3.4.1 与 Pratt 等人(2016)的方法相比

在对分类的实验结果进行对比时,采用混淆矩阵的形式来说明。由于标签0数据较多,与标签0类数据相比,其他4类病变的数据较少,在数据增强后,将本文分类方法与 Pratt 等人(2016)利用 CNN 的方法对数据较少的4种病变的分类准确率进行比较,对比结果如表4所示。可以看出,本文方法对4种病变的分类准确率都高于 Pratt 等人(2016)的方法。此外,采用本文方法得到的五分类混淆矩阵(如表3所示)虽然存在类别之间的错分情况,但都是分成相邻的类别。采用 Pratt 等人(2016)方法产生的混淆矩阵如表5所示,可以看出该方法将大部分的样本错分为正常类,这种情况在临床中会将患者诊断为正常,错过最佳的诊断机会,最终会导致失明。通过与 Pratt 等人

表4 Pratt 等人(2016)和本文方法对四分类 DR 的分类准确率

Table 4 Classification accuracy of methods of Pratt et al. (2016) and ours for four-class DR

方法	DR 类别			
	1	2	3	4
Pratt 等人(2016)	0	23.3	7.8	44.3
本文	61.8	42.4	51.7	58.8

注:加粗字体表示各列最优结果。

表5 Pratt 等人(2016)方法产生的混淆矩阵

Table 5 Confusion matrix generated by method of Pratt et al. (2016)

真实标签	预测类别				
	0	1	2	3	4
0	3 456	0	145	1	34
1	344	0	27	0	1
2	543	0	179	5	40
3	40	0	63	10	15
4	28	0	23	3	43

的结果对比,验证了本文模型方法的优越性。

3.4.2 与其他网络结构相比

EfficientNet 加入注意力机制产生的分类结果的混淆矩阵如表3所示。为进一步验证分类网络加入注意力机制后的效果,在 InceptionV3 网络中加入注意力机制,产生的分类结果的混淆矩阵如表6所示。从表3和表6可以看出,大部分数据均落在对角线上,验证了模型适用于 DR 图像分类任务。但具体到每一类验证时,可以看出本文提出的融合注意力机制的 EfficientNet 分类网络(A-EfficientNet)的效果明显优于融合注意力机制的 Inception V3 分类网络。

表6 融合注意力机制的 InceptionV3 网络产生的混淆矩阵

Table 6 Confusion matrix generated by InceptionV3 network incorporating the attention mechanism

真实标签	预测类别				
	0	1	2	3	4
0	3 466	878	620	63	69
1	233	240	12	1	2
2	327	389	269	35	41
3	26	29	50	42	27
4	3	12	37	25	64

表7是分别利用 InceptionV3 网络和 EfficientNet 及 InceptionV3 网络和 EfficientNet 加注意力机制训练网络的实验结果。可以看出,在加入注意力机制的情况下,InceptionV3 和 EfficientNet 模型的准确率分别达到 80.4% 和 97.2%。在不加入注意力机制的情况下,由于视网膜图像与自然图像之间特征差异大,虽然 EfficientNet 分类模型在 0~2 类的准

确性高于 InceptionV3 网络分类模型,但是并不能有效提取 DR 图像中的特征信息,因为在彩色视网膜图像的 3 通道中,通常绿色通道分量图像中包含较多的生理结构以及病变信息,而自然图像中每个通道的分量相对比较平均,从而导致分类模型对其他类别的分类准确性低。因此,利用 EfficientNet 在深度、宽度、像素等方面进行特征提取,加上注意力机制对细粒度图像的关注,提高了分类模型对 DR 的

分类效果。EfficientNet 加入注意力机制后预测类别出现的概率及图像类别分类的效果如表 8 和图 6 所示,可以看出,加入注意力机制后对每一个类别都有预测,真实标签与预测的类别概率基本一致。该分类模型明确表征出 EfficientNet 自适应优化网络结构的优势,体现出注意力机制对不同等级病变之间的有效特征提取与分类,表明了 A-EfficientNet 分类网络的优越性。

表 7 实验结果对比
Table 7 Comparison of experimental results

网络结构	准确率/%	特异性/%	敏感性/%	kappa 值
InceptionV3	78.3	88.9	78.2	0.722
InceptionV3 + attention	80.4	90.7	83.5	0.782
EfficientNet	92.5	97.5	91.8	0.80
A-EfficientNet	97.2	98.7	95.6	0.84

注:加粗字体表示各列最优结果。

表 8 EfficientNet 加入注意力机制后预测类别出现的概率
Table 8 Probability of predicting categories after adding attention mechanism to EfficientNet

真实标签	预测类别/%				
	0	1	2	3	4
0	98.75	1.25	0	0	0
1	2.1	97.90	0	0	0
2	1.72	9.87	88.41	0	0
3	0.05	2.74	7.21	89.91	0.09
4	1.76	3.21	5.36	8.68	80.99

注:加粗字体表示各列最优结果。

综上,利用 InceptionV3 网络和 EfficientNet 进行

分类,通过分析实验结果可以得出:EfficientNet 完成 DR 的分类任务比较好,但是准确率不太明显。加入注意力机制后,在分类时能对细粒度图像中的特征进行更多的关注,对网络的特征提取起到了积极的辅助作用,尤其是 EfficientNet 加入注意力机制后,kappa 值达到 0.84,高于 2015 年 Kaggle 网站(<https://www.kaggle.com/c/diabetic-retinopathy-detection/>)举办的 DR 分类比赛中取得的 0.829,分类性能得到提升。表明该五分类模型测试的结果与医生标注的数据集的结果具有很强的一致性,网络的分类性能大幅提高,模型的分类效果可以辅助临床医生进行检测,模型参数的利用率达到 99.4%,减少了训练过程中多余参数带来的计算负担,增加了模型的训练能力。

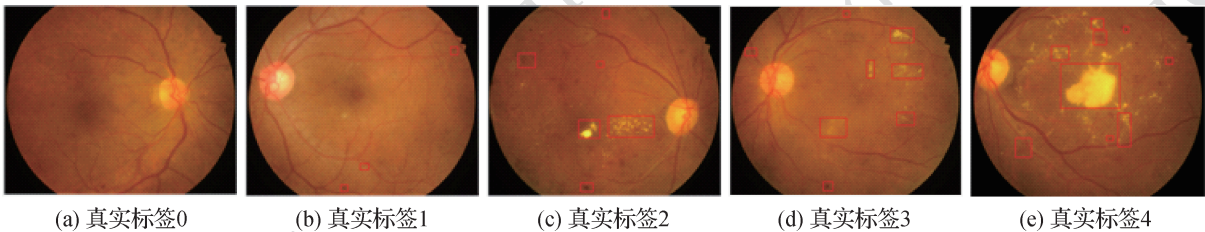


图 6 加入注意力机制后 EfficientNet 分类效果图

Fig. 6 EfficientNet classification effect maps after adding attention mechanism

((a) actual label 0; (b) actual label 1; (c) actual label 2; (d) actual label 3; (e) actual label 4)

4 结 论

本文提出一种融合注意力机制的 EfficientNet 分类算法 (A-EfficientNet) 用于糖尿病视网膜病变 (DR) 检测, 在 Kaggle 竞赛数据上验证了模型的分类效果。由于数据类别之间的不平衡性会导致机器在学习过程中出现过拟合的情况, 因此, 采用数据增强的方法对数据的类别进行了均衡处理。针对 DR 类别之间差异性较小的问题, 在原始的视网膜图像上进行一系列预处理, 使眼底中的血管特征更加明显; 利用 EfficientNet 将深度、宽度、像素值等指标进行有机结合, 通过综合优化, 避免了各项指标间的独立性, 增加了关联性; 接着在网络中加入注意力机制对细粒度图像进行特征提取, 使得网络可以更好地区分病变类型之间的差异。通过高效率网络 EfficientNet 与注意力机制的结合, 对糖尿病视网膜病变的五分类任务的精度得以提升。但在实验过程中, 图像的预处理、网络的特征提取和图像的五分类等阶段未形成完整的分类系统, 在后续的工作中将尝试将这 3 个阶段开发成完整的计算机辅助系统, 辅助医生进行临床检查。

参考文献 (References)

- Casanova R, Saldana S, Chew E Y, Danis R P, Greven C M and Ambrosius W T. 2014. Application of random forests methods to diabetic retinopathy classification analyses. *PLoS One*, 9(6): e98587 [DOI: 10.1371/journal.pone.0098587]
- Decenci re E, Zhang X W, Cazuguel G, Lay B, Cochener B, Trone C, Gain P, Ord  ez-Varela J R, Massin P, Erginay A, Charton B and Klein J C. 2014. Feedback on a publicly distributed image database: the Messidor database. *Image Analysis and Stereology*, 33(3): 231-234 [DOI: 10.5566/ias.1155]
- Doshi D, Shenoy A, Sidhpura D and Gharpure P. 2016. Diabetic retinopathy detection using deep convolutional neural networks//Proceedings of 2016 International Conference on Computing, Analytics and Security Trends (CAST). Pune, India: IEEE: 261-266 [DOI: 10.1109/CAST.2016.7914977]
- Flaxman S R, Bourne R R A, Resnikoff S, Ackland P, Braithwaite T, Cicinelli M V, Das A, Jonas J B, Keeffe J, Kempen J H, Leasher J, Limburg H, Naidoo K, Pesudovs K, Silvester A, Stevens G A, Tahhan N, Wong T Y and Zheng Y F. 2017. Global causes of blindness and distance vision impairment 1990-2020: a systematic review and meta-analysis. *Lancet Global Health*, 5(12): e1221-e1234 [DOI: 10.1016/S2214-109X(17)30393-5]
- Fleiss J L, Cohen J and Everitt B S. 1969. Large sample standard errors of kappa and weighted kappa. *Psychological Bulletin*, 72(5): 323-327 [DOI: 10.1037/h0028106]
- French Ministry. 2014. Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology [EB/OL]. [2019-11-11] <http://www.adcis.net/en/DownloadThirdParty/Messidor.html>
- Gargeya R and Leng T. 2017. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology*, 124(7): 962-969 [DOI: 10.1016/j.ophtha.2017.02.008]
- Graham B. 2015. Kaggle Diabetic Retinopathy Detection Competition Report. Coventry University of Warwick
- Gulshan V, Peng L, Coram M, Stumpe M C, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson P C, Mega J L and Webster D R. 2016. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama*, 316(22): 2402-2410 [DOI: 10.1001/jama.2016.17216]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Huang Y P, Cheng Y L, Bapna A, Firat O, Chen M X, Chen D H and Lee H J, Ngiam J, Le Q V, Wu Y H, Chen Z F. 2018. GPipe: efficient training of giant neural networks using pipeline parallelism [EB/OL]. [2019-11-11]. <https://arxiv.org/pdf/1811.06965.pdf>
- Kaggle. 2015. Diabetic retinopathy detection [EB/OL]. [2019-11-11]. <https://www.kaggle.com/c/diabetic-retinopathy-detection/>
- Kingma D P and Ba J. 2014. Adam: a method for stochastic optimization [EB/OL]. [2019-11-11]. <https://arxiv.org/pdf/1412.6980.pdf>
- Krizhevsky A, Sutskever I and Hinton G E. 2017. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6): 84-90 [DOI: 10.1145/3065386]
- Laud K, Shabto U and Tello C. 2017. Diabetic Retinopathy//Principles of Diabetes Mellitus. 2010: 331-346 [DOI: 10.1007/978-3-319-18741-9_2-1]
- Li Q, Bai Z Y and Liu Y F. 2018. Automated classification of diabetic retinal images by using deep learning method. *Journal of Image and Graphics*, 23(10): 1594-1603 (李琼, 柏正尧, 刘莹芳. 2018. 糖尿病性视网膜图像的深度学习分类方法. *中国图象图形学报*, 23(10): 1594-1603) [DOI: 10.11834/jig.170683]
- Li X G, Pang T T, Xiong B, Liu W X, Liang P and Wang T F. 2017. Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification//Proceedings of the 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). Shanghai, China: IEEE: 1-11 [DOI: 10.1109/CISP-BMEI.2017.8301998]

- Pan S J and Yang Q. 2010. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10): 1345-1359 [DOI: 10.1109/TKDE.2009.191]
- Pang H and Wang C. 2017. Deep learning model for diabetic retinopathy detection. *Journal of Software*, 28(11): 3018-3029 (庞浩, 王枫. 2017. 用于糖尿病视网膜病变检测的深度学习模型. *软件学报*, 28(11): 3018-3029) [DOI: 10.13328/j.cnki.jos.005332]
- Pratt H, Coenen F, Broadbent D M, Harding S P and Zheng Y L. 2016. Convolutional neural networks for diabetic retinopathy. *Procedia Computer Science*, 90: 200-205 [DOI: 10.1016/j.procs.2016.07.014]
- Rakhlina. 2017. Diabetic retinopathy detection through integration of deep learning classification framework [EB/OL]. [2019-11-11]. <https://www.biorxiv.org/content/10.1101/225508v1>
- Shahin E M, Taha T E, Al-Nuaimy W, El Rabaie S, Zahran O F and El-Samie F E A. 2012. Automated detection of diabetic retinopathy in blurred digital fundus images//*Proceedings of the 8th International Computer Engineering Conference (ICE-NCO)*. Cairo, Egypt: IEEE: 20-25 [DOI: 10.1109/ICENCO.2012.6487084]
- Srivastava N, Hinton G, Krizhevsky A, Sutskever I and Salakhutdinov R. 2014. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1): 1929-1958
- Tan M X and Le Q V. 2019. EfficientNet: rethinking model scaling for convolutional neural networks [EB/OL]. [2019-11-09]. <https://arxiv.org/pdf/1905.11946v3.pdf>
- Voets M, Møllersen K and Bongo L A. 2018. Reproduction study using public data of: development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *PLoS One*, 14(6): e0217541 [DOI: 10.1371/journal.pone.0217541]
- Wang Y, Wang G A, Fan W G and Li J X. 2018. A deep learning based pipeline for image grading of diabetic retinopathy//*Proceedings of 2018 International Conference on Smart Health*. Wuhan, China: Springer: 240-248 [DOI: 10.1007/978-3-030-03649-2_24]
- Zagoruyko S and Komodakis N. 2016. Wide residual networks//*Proceedings of the British Machine Vision Conference*. York: BMVA Press: #87 [DOI: 10.5244/C.30.87]
- Zeng X L, Chen H Q, Luo Y and Ye W B. 2019. Automated diabetic retinopathy detection based on binocular siamese-like convolutional neural network. *IEEE Access*, 7:30744-30753 [DOI: 10.1109/ACCESS.2019.2903171]
- Zhou K, Gu Z W, Liu W, Luo W X, Cheng J, Gao S H and Liu J. 2018. Multi-cell multi-task convolutional neural networks for diabetic retinopathy grading//*Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Honolulu, HI, USA: IEEE: 2724-2727 [DOI: 10.1109/EMBC.2018.8512828]

作者简介



张子振, 1992年生, 男, 硕士研究生, 主要研究方向为基于深度学习的医学图像处理。

E-mail: 639521600@qq.com



刘明, 通信作者, 男, 副教授, 主要研究方向为模式识别与信号处理。

E-mail: gleer@126.com

朱德江, 男, 硕士研究生, 主要研究方向为基于深度学习的图像处理。E-mail: 976544244@qq.com