



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象学报 中国图形学报

2022
03
VOL.27

ISSN1006-8961
CN11-3758/TB



医学图像及临床应用



第27卷第3期（总第311期）
2022年3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 中国科学院空天信息创新研究院
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 吴一戎
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@aircas.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 Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
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@aircas.ac.cn
Telephone 010-58887035
Website www.cjig.cn

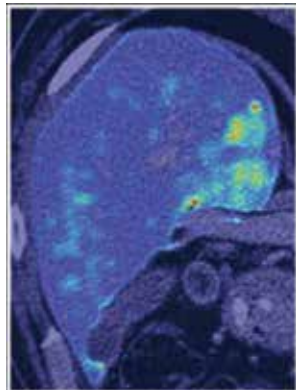
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

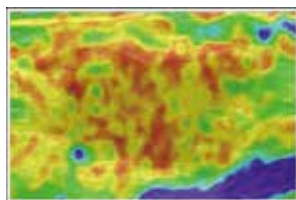
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



慢阻肺患者CT图像中肺内血管分割及定量分析(第0762页)



基于特征选择与残差融合的肝肿瘤分割模型(第0838页)



面向乳腺超声计算机辅助诊断的两阶段深度迁移学习(第0898页)

专刊封面展示了上海科技大学生物医学工程学院和上海联影智能医疗科技有限公司合作的脑影像智能分析方向研究成果。

《中国图象图形学报》医学图像及临床应用专刊简介

沈定刚, 刘天明, 周涛, 夏勇, 白相志, 王乾, 李刚, 徐寿平, 刘昊, 韩向娣 0653

综述

中国医学影像人工智能20年回顾和展望

蒋希, 袁奕宣, 王雅萍, 肖振祥, 朱美芦, 陈泽华, 刘天明, 沈定刚 0655

迁移学习在医学图像分类中的研究进展

黎英, 宋佩华 0672

生成对抗式网络及其医学影像应用研究综述

张颖麟, 胡衍, 东田理沙, 刘江 0687

基于深度学习的心脏磁共振影像超分辨率前沿进展

李书林, 冯朝路, 于鲲, 刘鑫, 江鑫, 赵大哲 0704

CT图像肺及肺病变区域分割方法综述

冯龙锋, 陈英, 周滔辉, 胡菲, 易珍 0722

计算机断层扫描图像

结构图注意力网络的新冠肺炎轻重症诊断

刘彦北, 李赫南, 张长青, 肖志涛, 张芳, 陈英, 高耀宗, 石峰, 单飞, 沈定刚 0750

慢阻肺患者CT图像中肺内血管分割及定量分析

赵宏, 李璋, 张杰华, 王琨, 孙家兴, 廖玺铭, 阎昱升, 钟正, 张鑫, 孙健, 于起峰, 葛俊辉 0762

基础信息保持和细节强化的胸部CT图像增强

张余, 张顺利, 白相志, 张利 0774

面向小样本股骨骨折分型的多视角注意力融合方法

张亚东, 汪玲, 兰海, 翟禹樵, 程洪 0784

三维多尺度嵌套U结构CT影像肺结节检测

许正奎, 张少敏, 支力佳, 周涛 0797

融合多尺度残差和注意力机制的特发性肺纤维化进展预测

陈舞, 孙军梅, 李秀梅 0812

COVID-19肺部CT图像多尺度编解码分割

陆倩杰, 柏正尧, 樊圣澜, 周雪, 许祝 0827

基于特征选择与残差融合的肝肿瘤分割模型

乔伟晨, 黄冕, 刘利军, 黄青松 0838

磁共振图像

面向多模态MRI脑胶质瘤区域三维分割与生存期预测的级联U-Net网络

余力, 刘霄雪, 闫朝阳, 李建瑞, 张志强, 黄轶枢, 徐军 0850

心脏动态MRI图像分割的时空多尺度网络

徐佳陈, 肖志勇 0862

融合跨阶段深度学习的脑肿瘤MRI图像分割

夏峰, 邵海见, 邓星 0873

融入平滑组稀疏化的脑部MRI图像分类

黄帅辉, 王金凤 0885

超声图像

面向乳腺超声计算机辅助诊断的两阶段深度迁移学习

贡荣麟, 施俊, 周玮珺, 汪程 0898

结合注意力机制的乳腺双模态超声分类网络

赵绪, 龚勋, 樊琳, 罗俊 0911

中医图像

视觉Transformer与多特征融合的脑卒中检测算法

赵琛琦, 王华虎, 赵涓涓, 冀伦文, 王麒达, 李慧芝, 赵紫娟 0923

多分支深度特征融合的中医脑卒中辅助诊断

王麒达, 冀伦文, 强彦, 王华虎, 赵琛琦, 李慧芝, 赵紫娟 0935

研究应用

多区域融合注意力网络模型下的核性白内障分类

章晓庆, 肖尊杰, Risa Higashita, 陈婉, 胡衍, 袁进, 刘江 0948

面向高度近视条纹损伤的深监督特征聚合网络

谭晓, 刁逸超, 陈新建, 石霏, 樊莹, 谢嘉受, 朱伟芳 0961

结合目标检测与匹配修正的手腕骨兴趣区域提取

毛科技, 汪敏豪, 陈立建, 陆伟, 武坤秀, 陈庆章, 赵小敏 0973

面向旷场实验视频分类的特征拼接矩阵学习方法

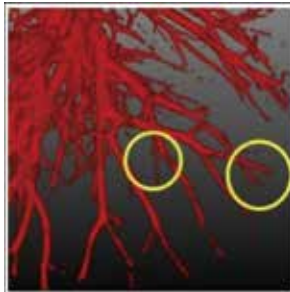
肖照林, 杨志林, 刘欢, 金海燕 0988

基于图像的介入器械前端变形建模与受力估计

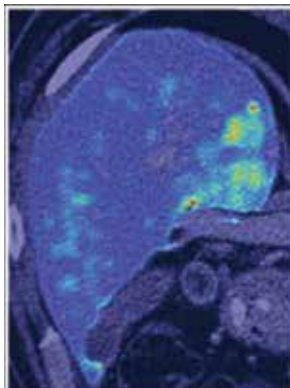
任龙飞, 孟偲, 刘博, 白相志, 周付根 1001

CONTENTS

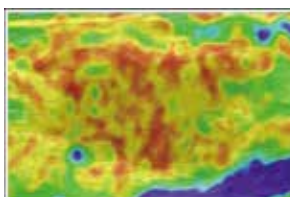
JOURNAL OF IMAGE AND GRAPHICS



Segmentation and quantitative analysis of intrapulmonary vasculature in CT images from COPD patients(P0762)



Feature selection and residual fusion segmentation network for liver tumor(P0838)



Two-stage deep transfer learning for human breast ultrasound computer-aided diagnosis(P0898)

Review

- A 20-year retrospect and prospect of medical imaging artificial intelligence in China
Jiang Xi, Yuan Yixuan, Wang Yaping, Xiao Zhenxiang, Zhu Meilu, Chen Zehua, Liu Tianming, Shen Dinggang 0655
- Review of transfer learning in medical image classification
Li Ying, Song Peihua 0672
- A review of generative adversarial networks and the application in medical image
Zhang Yinglin, Hu Yan, Risa Higashita, Liu Jiang 0687
- Critical review of human cardiac magnetic resonance image super resolution reconstruction based on deep learning method
Li Shulin, Feng Chaolu, Yu Kun, Liu Xin, Jiang Xin, Zhao Dazhe 0704
- Review of human lung and lung lesion regions segmentation methods based on CT images
Feng Longfeng, Chen Ying, Zhou Taohui, Hu Fei, Yi Zhen 0722

Computerd Tomography Image

- Diagnosis of COVID-19 by using structural attention graph neural network
Liu Yanbei, Li Henan, Zhang Changqing, Xiao Zhitao, Zhang Fang, Wei Ying, Gao Yaozong, Shi Feng, Shan Fei, Shen Dinggang 0750
- Segmentation and quantitative analysis of intrapulmonary vasculature in CT images from COPD patients
Zhao Hong, Li Zhang, Zhang Jiehua, Wang Kun, Sun Jiaxing, Liao Ximing, Yan Yusheng, Zhong Zheng, Zhang Xin, Sun Jian, Yu Qifeng, Ge Junhui 0762
- Human chest CT image enhancement based on basic information preservation and detail enhancement
Zhang Yu, Zhang Shunli, Bai Xiangzhi, Zhang Li 0774
- Multi-view attention fusion method for few-shot femoral fracture classification
Zhang Yadong, Wang Ling, Lan Hai, Zhai Yuqiao, Cheng Hong 0784
- Detection of pulmonary nodules in three-dimensional multiscale nested U-structure computed tomography images
Xu Zhengxi, Zhang Shaomin, Zhi Lijia, Zhou Tao 0797
- Multi-scale residual and attention mechanism fusion based prediction for the progression of idiopathic pulmonary fibrosis
Chen Wu, Sun Junmei, Li Xiumei 0812
- Multiscale codec network based CT image segmentation for human lung disease derived of COVID-19
Lu Qianjie, Bai Zhengyao, Fan Shenglan, Zhou Xue, Xu Zhu 0827
- Feature selection and residual fusion segmentation network for liver tumor
Qiao Weichen, Huang Mian, Liu Lijun, Huang Qingsong 0838

Magnetic Resonance Image

- Brain glioma tumor segmentation and survival prediction from multi-modality MRIs via cascaded U-Net
Yu Li, Liu Xiaoxue, Yan Chaoyang, Li Jianrui, Zhang Zhiqiang, Huang Yunzhi, Xu Jun 0850
- A spatio-temporal multi-scale network for cardiac dynamic MRI image segmentation
Xu Jiachen, Xiao Zhiyong 0862
- Cross-stage deep-learning-based MRI fused images of human brain tumor segmentation
Xia Feng, Shao Haijian, Deng Xing 0873
- Brain MRI image classification incorporating smooth group sparseness
Huang Shuaihui, Wang Jinfeng 0885

Ultrasound Image

- Two-stage deep transfer learning for human breast ultrasound computer-aided diagnosis
Gong Ronglin, Shi Jun, Zhou Weijun, Wang Cheng 0898
- Attention-based networks of human breast bimodal ultrasound imaging classification
Zhao Xu, Gong Xun, Fan Lin, Luo Jun 0911

Traditional Chinese Medical Image

- Cerebral stroke detection algorithm for visual Transformer and multi-feature fusion
Zhao Chenqi, Wang Huahu, Zhao Juanjuan, Ji Lunwen, Wang Qida, Li Huizhi, Zhao Zijuan 0923
- Multi-branch deep feature fusion method for traditional Chinese medicine intervened Human cerebral stroke aided diagnosis
Wang Qida, Ji Lunwen, Qiang Yan, Wang Huahu, Zhao Chenqi, Li Huizhi, Zhao Zijuan 0935

Research and Application

- Nuclear cataract classification based on multi-region fusion attention network model
Zhang Xiaoqing, Xiao Zunjie, Risa Higashita, Chen Wan, Hu Yan, Yuan Jin, Liu Jiang 0948
- Deep-supervision and feature-aggregation network for linear lesion segmentation of high myopia
Tan Xiao, Diao Yichao, Chen Xinjian, Shi Fei, Fan Ying, Xie Jiamin, Zhu Weifang 0959
- Hand-wrist region of interest extraction based on object detection and matching correction
Mao Keji, Wang Minhao, Chen Lijian, Lu Wei, Wu Kunxiu, Chen Qingzhang, Zhao Xiaomin 0973
- A spliced feature matrices learning method for open field test video classification
Xiao Zhaolin, Yang Zhilin, Liu Huan, Jin Haiyan 0988
- Image-based deformation modeling and force estimation of the front end of vascular interventional surgical apparatus
Ren Longfei, Meng Cai, Liu Bo, Bai Xiangzhi, Zhou Fugen 1001

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2022)03-0948-13

论文引用格式: Zhang X Q, Xiao Z J, Risa Higashita, Chen W, Hu Y, Yuan J and Liu J. 2022. Nuclear cataract classification based on multi-region fusion attention network model. Journal of Image and Graphics, 27(03): 0948-0960 (章晓庆, 肖尊杰, 东田理沙, 陈婉, 胡衍, 袁进, 刘江. 2022. 多区域融合注意力网络模型下的核性白内障分类. 中国图象图形学报, 27(03): 0948-0960) [DOI: 10.11834/jig.210735]

多区域融合注意力网络模型下的核性白内障分类

章晓庆¹, 肖尊杰¹, 东田理沙^{1,2}, 陈婉³, 胡衍¹, 袁进³, 刘江^{1,4,5*}

1. 南方科技大学计算机科学与工程系, 深圳 518055; 2. TOMEY 株式会社, 名古屋 451-0051, 日本;
3. 中山大学中山眼科中心, 广州 510060; 4. 中国科学院宁波材料技术与工程研究所慈溪生物医学工程研究所, 宁波 315201;
5. 广东省类脑智能计算重点实验室, 深圳 518055

摘要: 目的 核性白内障是主要致盲和导致视觉损害的眼科疾病, 早期干预和白内障手术可以有效改善患者的视力和生活质量。眼前节光学相干断层成像图像 (anterior segment optical coherence tomography, AS-OCT) 能够非接触、客观和快速地获取白内障混浊信息。临床研究已经发现在 AS-OCT 图像中核性白内障严重程度与核性区域像素特征, 如均值存在强相关性和高可重复性。但目前基于 AS-OCT 图像的自动核性白内障分类工作较少且分类结果还有较大提升空间。为此, 本文提出一种新颖的多区域融合注意力网络 (multi-region fusion attention network, MRA-Net) 对 AS-OCT 图像中的核性白内障严重程度进行精准分类。**方法** 在提出的多区域融合注意力模型中, 本文设计了一个多区域融合注意力模块 (multi-region fusion attention, MRA), 对不同核性区域特征表示进行融合来增强分类结果; 另外, 本文验证了以人和眼为单位的 AS-OCT 图像数据集拆分方式对核性白内障分类结果的影响。**结果** 在一个自建的 AS-OCT 图像数据集上结果表明, 本文模型的总体分类准确率为 87.78%, 对比方法至少提高了 1%。在 10 种分类算法上的结果表明: 以眼为单位的 AS-OCT 数据集优于以人为单位的 AS-OCT 数据集的分类结果, F1 和 Kappa 评价指标分别最大提升了 4.03% 和 8%。**结论** 本文模型考虑了特征图不同区域特征分布的差异性, 使核性白内障分类更加准确; 不同数据集拆分方式的结果表明, 考虑到同一个人两只眼的核性白内障严重程度相似, 建议白内障的 AS-OCT 图像数据集拆分以人为单位。

关键词: 核性白内障分类; 眼前节光学相干断层成像图像 (AS-OCT); 多区域融合注意力模块; 深度学习; 核性区域

Nuclear cataract classification based on multi-region fusion attention network model

Zhang Xiaoqing¹, Xiao Zunjie¹, Risa Higashita^{1,2}, Chen Wan³, Hu Yan¹, Yuan Jin³, Liu Jiang^{1,4,5*}

1. Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China;
2. TOMEY Corporation, Nagoya 451-0051, Japan; 3. Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou 510060, China;
4. Cixi Institute of Biomedical Engineering, Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Sciences, Ningbo 315201, China; 5. Guangdong Provincial Key Laboratory of Brain-inspired Intelligent Computation, Shenzhen 518055, China

Abstract: **Objective** Cataracts are the primary inducement for human blindness and vision impairment. Early intervention

收稿日期: 2021-08-20; 修回日期: 2021-11-19; 预印本日期: 2021-11-26

* 通信作者: 刘江 liuj@sustech.edu.cn

基金项目: 国家自然科学基金项目 (8210072776); 广东省普通高校重点领域专项基金项目 (202DZX3043); 广东省重点实验室项目 (2020B121201001)

Supported by: National Natural Science Foundation of China (8210072776); Key Areas of Guangdong Province Colleges and Universities Special Funding (202DZX3043); Guangdong Provincial Key Laboratory (2020B121201001)

and cataract surgery can effectively improve the vision and life quality of cataract patients. Anterior segment optical coherence tomography (AS-OCT) image can capture cataract opacity information through a non-contact, objective, and fast manner. Compared with other ophthalmic images like fundus images, AS-OCT images are capable of capturing the clear nucleus region, which is very significant for nuclear cataract (NC) diagnosis. Clinical studies have identified that a strong opacity correlation relationship and high repeatability between average density value of the nucleus region and NC severity levels in AS-OCT images. Moreover, the clinical works also have suggested that the correlation relationships between different nucleus regions and NC severity levels. These original research works provide the clinical reference for automatic AS-OCT image-based NC classification. However, automatic NC classification based on AS-OCT images has been rarely studied, and there is much improvement room for NC classification performance on AS-OCT images. **Method** Motivated by the clinical research of NC, this paper proposes an efficient multi-region fusion attention network (MRA-Net) model by infusing clinical prior knowledge, aiming to classify nuclear cataract severity levels on AS-OCT images accurately. In the MRA-Net, we construct a multi-region fusion attention (MRA) block to fuse feature representation information from different nucleus regions to enhance the overall classification performance, in which we not only adopt the summation operation to fuse different region information but also apply the softmax function to focus on salient channel and suppress redundant channels. In respect of the residual connection can alleviate the gradient vanishing issue, the MRA block is plugged into a cluster of Residual-MRA modules to demonstrate MRA-Net. Moreover, we also test the impacts of two different dataset splitting methods on NC classification results: participant-based splitting method and eye-based splitting method, which is easily ignored by previous works. In the training, this paper resizes the original AS-OCT images into 224×224 pixels as the network inputs and set batch size to 16. Stochastic gradient descent (SGD) optimizer is used as the optimizer with default settings and we set training epochs to 100. **Result** Our research analysis demonstrates that the proposed MRA-Net achieves 87.78% accuracy and obtains 1% improvement than squeeze and excitation network (SENet) based on a clinical AS-OCT image dataset. We also conduct comparable experiments to verify that the summation operation works better the concatenation on the MRA block by using ResNet as the backbone network. The results of two dataset splitting methods also that ten classification methods like MRA-Net and SENet obtain better classification results on the eye-based dataset than the participant-based dataset, e. g., the highest improvements on F1 and Kappa are 4.03% and 8% correspondingly. **Conclusion** Our MRA-Net considers the difference of feature distribution in different regions in a feature map and incorporates the clinical priors into network architecture design. MRA-Net obtains surpassing classification performance and outperforms advanced methods. The classification results of two dataset splitting methods on AS-OCT image dataset also indicated that given the similar nuclear cataract severity in the two eyes of the same participant. Thus, the AS-OCT image dataset is suggested to be split based on the participant level rather than the eye level, which ensures that each participant falls into the same training or testing datasets. Overall, our MRA-Net has the potential as a computer-aided diagnosis tool to assist clinicians in diagnosing cataract.

Key words: nuclear cataract classification; anterior segment optical coherence tomography (AS-OCT) image; multi-region fusion attention block; deep learning; nucleus region

0 引言

年龄相关性白内障(age related cataract)是全世界排名首位的致盲性疾病和导致视觉损伤的主要原因(Gali 等, 2019), 其发病率随着年龄增长而显著增加。在中国, 随着社会人口的老龄化, 到 2030 年白内障患病人数估计达到 3 亿。白内障的临床表现为眼晶状体结构出现混浊症状。依据症状出现位置, 白内障从内到外可分为核性白内障(nuclear cat-

aract, NC)、皮质性白内障(cortical cataract, CC)和后囊性白内障(posterior subcapsular cataract, PSC)。早期干预和白内障手术可以有效改善病人的视力和生活质量, 同时也能降低失明率和减轻社会负担。

核性白内障是一种最常见的年龄相关性白内障, 临床症状为晶状体的核性区域逐渐变硬且色泽逐渐加深。在临床上, 眼科医生基于白内障临床诊断标准通常采用裂隙灯影像(slit lamp image)诊断核性白内障的严重程度。其中, 晶状体混浊分类系统(lens opacities classification system III, LOCS III)

(Chylack 等, 1993) 是一个建立在裂隙灯图像上且广泛认可的白内障临床诊断标准。根据核性白内障的混浊度发展阶段和临床实际诊断需求 (Ozgokce 等, 2019), 核性白内障严重程度在 LOCS III 分级系统下可分为轻度 (mild) 和严重 (severe) 两个发展阶段。轻度阶段指白内障患者的晶状体核性区域出现混浊症状但不明显, 核性白内障级别对应 1 级和 2 级, 这类患者可以使用点滴药物缓解白内障发展进程; 严重阶段指白内障患者的晶状体核性区域出现明显混浊症状, 核性白内障级别对应 3 级及以上级别, 这类患者需要接受临床跟踪或白内障手术。

但是, 临床上采用裂隙灯图像和眼底图像诊断白内障的方式与医生临床经验和知识相关, 具有较强的主观性, 容易出现误诊。为了提高白内障诊断的精准性, 研究学者在多种眼科影像模式下提出了不少机器学习和深度学习用于白内障自动分类和筛查。Li 等人 (2007) 和 Huang 等人 (2009) 基于裂隙灯图像提出一个核性白内障自动诊断系统, 由晶状体核性区域检测、像素特征提取和分级预测算法 3 部分组成, 在一个临床裂隙灯图像数据集的结果表明, 该自动诊断系统取得了 0.36 的平均误差。Xu 等人 (2013) 同样在临床裂隙灯图像数据集进行核性白内障分类, 使用群稀疏回归算法 (group sparsity regression, GSR) 取得了 83.4% 准确率, 并进一步通过语义相似度算法 (Xu 等, 2016) 提高分类结果。Caixinha 等人 (2016) 在超声图像上利用支持向

量机 (support vector machine, SVM) 进行软硬白内障分类并取得了不错的分类效果, 但所用的超声图像数据集是来自动物且样本量较小。Cao 等人 (2020b) 在眼底图像上通过改进 Haar 小波变换算法实现白内障精准筛查, 取得了 85.98% 准确率。除了传统图像处理方法和机器学习方法以外, 研究学者也采用深度学习方法进行白内障分类。Gao 等人 (2015) 基于裂隙灯图像提出卷积神经网络 (convolutional neural network, CNN) 和循环神经网络 (recurrent neural network, RNN) 混合模型用于核性白内障分类并取得了 82.5% 准确率。Xu 等人 (2019) 利用 Faster RCNN 网络框架实现端对端的核性区域自动检测和核性白内障分级并取得 83.5% 准确率。中山大学研究团队 (Long 等, 2017) 提出一个基于深度学习的先天性白内障筛查平台并取得了不错的筛查结果。Xu 等人 (2020) 在眼底图像下提出一个总体局部混合卷积神经网络模型, 通过集成融合不同部分混浊病理信息, 白内障筛查结果优于已发表文献的工作。李建强等人 (2018) 同样在眼底图像下提出深度神经网络用于白内障筛查并取得不错结果。

眼前节光学相干断层成像技术 (anterior segment optical coherence tomography, AS-OCT) 是一种新型 OCT 成像技术, 如图 1(a) 所示, 具有速度快、易使用、用户友好、非接触、高分辨率和定量化测量等特点。近年来, 眼科医生已经广泛使用 AS-OCT 图

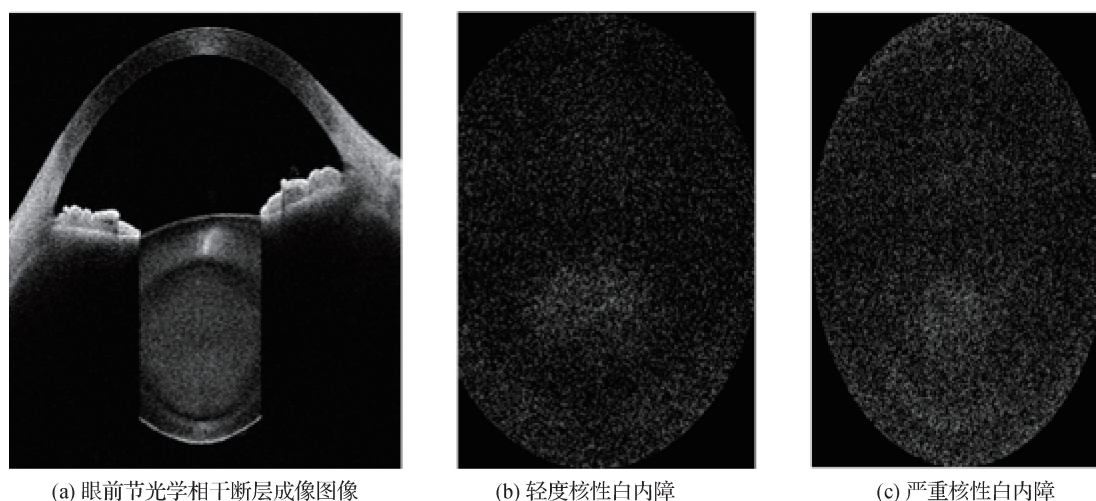


图 1 基于 AS-OCT 图像的两核性白内障严重程度分级

Fig. 1 Two nuclear cataract severity levels based on AS-OCT images

((a) anterior segment optical coherence tomography (AS-OCT) image; (b) mild level; (c) severe level)

像诊断眼病,如角膜疾病和闭角青光眼。Keller 等人(2018)和 Dos Santos 等人(2019)在 AS-OCT 图像上利用深度分割网络模型对角膜结构进行分割,并得到不错的分割结果。Fu 等人(2018,2019)在 AS-OCT 图像上利用深度神经网络对青光眼开闭角进行筛查,在一个临床 AS-OCT 数据集上取得了 92.5% 准确率。在白内障诊断方面,AS-OCT 图像能清楚地获取包括核性区域、皮质性区域和囊性区域的完整晶状体结构,可以较好地辅助眼科医生诊断不同类型白内障。而其他眼科图像如裂隙灯图像仅能获取核性区域且不清楚,眼底图像只能获取白内障的混浊度信息,不能获取核性区域。在临床基础研究中,已经有不少研究学者基于 AS-OCT 图像研究晶状体区域的像素特征与核性白内障严重程度之间的相关关系。Wong 等人(2009)在 AS-OCT 图像下使用线性拟合方法计算核性区域的像素均值特征与核性白内障严重程度之间相关性,临床统计结果显示两者存在较好的相关性和高可重复性,这也表明 AS-OCT 图像能够客观地描述核性白内障严重程度和能够用于辅助医生精准诊断白内障。De Castro 等人(2018)、Wang 等人(2021)和 Makhotkina 等人(2018)同样在 AS-OCT 图像下分析核性区域的像素均值特征与核性白内障严重程度之间的相关性,并得到了类似的临床统计结果,但 Makhotkina 等人(2018)在核性区域下半部分的相关性弱于核性区域上半部分,表明核性白内障严重程度与白内障的混浊位置分布有关,可作为临床先验知识注入到算法设计中,以此来提高白内障分类结果。

总体而言,以上临床研究为 AS-OCT 图像下核性白内障自动分类提供了临床支持和依据。Cao 等人(2020a)借助深度卷积网络模型对 AS-OCT 图像的晶状体结构进行分割,得到晶状体的核性区域、皮质性区域和囊性区域,可用于辅助医生诊断不同类型白内障。Wang 等人(2021)基于 AS-OCT 图像进行软硬核分类并取得不错的分类效果。Zhang 等人(2020)使用深度卷积神经网络 GraNet 实现 AS-OCT 图像下核性白内障分类,但仅取得低于 60% 的准确率。章晓庆等人(2021)从 AS-OCT 图像晶状体核性区域中提取了 18 个像素特征,并使用皮尔逊相关系数方法和随机森林算法分别进行特征重要性分析和分类,取得了 75.53% 准确率。章晓庆等人(2021)研究中的 AS-OCT 图像数据集拆分是以眼睛为单

位,但临床研究工作的 AS-OCT 数据集拆分以个体为单位,目前尚未有研究工作对比这两种数据集拆分方式对白内障的分类结果影响。以上基于 AS-OCT 图像的自动核性白内障分类研究工作也表明分类算法在核性区域的分类结果优于整个晶状体区域,主要是因为晶状体不仅与核性白内障有关,还与皮质性白内障和后囊性白内障有关。

从现有的研究工作可以看出,AS-OCT 图像下的核性白内障分类临床研究较多,但基于 AS-OCT 图像的核性白内障自动分类研究工作还比较少且尚未有研究工作将临床先验知识引入到算法设计中来提升分类结果。为此,本文基于 AS-OCT 图像提出了一个多区域融合注意力网络模型(multi-region fusion attention network, MRA-Net)对核性白内障严重程度进行自动分类,用于辅助医生精准地诊断核性白内障。与挤压激励注意力机制(squeeze-and-excitation attention, SE)模型(Hu 等,2018)不同,在提出的 MRA-Net 模型中,本文引入了一个多区域融合注意力模块(multi-region fusion attention block, MRA)对特征图不同区域的特征表示信息进行融合来提升分类性能。为了验证基于人和眼睛的数据集拆分方法对核性白内障分类结果的影响,本文采用了 10 种分类算法进行较为全面的验证。

本文主要贡献如下:1)提出了一个多区域融合注意力网络模型,自动学习和融合不同区域的特征表示信息,以此来提高核性白内障的分类结果。2)在一个以人为单位的自建 AS-OCT 图像数据集进行验证,本文模型取得了 0.877 8 的准确率,优于先进的分类算法。3)验证了以人和眼为单位两个 AS-OCT 图像数据集对分类性能的影响,结果表明 10 种分类算法在以眼睛为单位的 AS-OCT 图像数据集上的分类结果优于以人为单位的 AS-OCT 图像数据集。但考虑到同一个人的两只眼的核性白内障严重程度相似和鉴于临床白内障研究规范,本文建议基于 AS-OCT 图像的白内障数据集拆分以人为单位。

1 AS-OCT 图像数据集

本文使用的 AS-OCT 图像数据集来自一个本地社康中心,数据采集设备是日本 TOMEY 公司的 AS-OCT 仪器 CASIA2。数据集包含 371 名受试者和 520 只眼睛,其中,左眼 269 只,右眼 251 只;所有受

试者平均年龄为 61.30 ± 18.65 。所有受试者知情数据用途。

因为没有基于 AS-OCT 图像的白内障分类临床诊断标准,受试者的 AS-OCT 图像的核性白内障级别是通过裂隙灯图像映射得到的。每一名受试者都拍摄了 AS-OCT 图像和裂隙灯图像,并由 3 名经验丰富的眼科医生基于 LOCS III 分级系统标记了每个受试者的裂隙灯图像的白内障严重程度,从而得到 AS-OCT 图像下的核性白内障标记并保证标记的可靠性。

核性白内障只与晶状体核性区域有关,本文利用晶状体结构分割算法(Cao 等,2020a)得到核性区域(nucleus region)或感兴趣区域(region of interest, ROI)。在 LOCS III 分级系统下,白内障患者的核性白内障等级是 1 级或 2 级,临床上表现是有症状但不明显,即轻度核性白内障;白内障患者的核性白内障等级 ≥ 3 级时,临床表现为核性区域具有明显的混浊,即严重核性白内障。两种严重程度核性白内障的核性区域 AS-OCT 图像如图 1 所示。在 AS-OCT 图像数据集中,受试者的每只眼采集 20 幅 AS-OCT 图像,本文在眼科医生帮助下剔除了眼皮遮挡的晶状体不完整的图像,本文可用的 AS-OCT 图像数量为 9 551 幅。

为了验证以人和眼睛为单位的两种数据集对分类算法结果的影响,本文将 AS-OCT 图像数据集按照人和眼睛为单位得到两种不同的数据集:个体数据集(participant dataset)和眼睛数据集(eye dataset)。在个体数据集中,一个受试者的左右眼图像不是全部在训练集中,就是全部在测试集中。在眼睛数据集中,每只眼的 AS-OCT 图像不是全部在训练集中,就是全部在测试集中。表 1 为个体数据集和眼睛数据集的训练集和测试集的核性白内障级别分布。在

表 1 基于个体数据集和眼睛数据集的
核性白内障严重程度标记分布

Table 1 The distribution of nuclear cataract severity levels based on participant dataset and eye dataset

数据集分类	个体数据集		眼睛数据集	
	轻度	严重	轻度	严重
训练集	2 308	4 379	2 305	4 430
测试集	762	2 102	765	2 051
总计	3 070	6 481	3 070	6 481

实验中,本文从训练集中随机抽取 20% 数据为验证集。图 2 为用于 AS-OCT 图像下核性白内障自动分类的多区域融合注意力网络框架。

2 多区域融合注意力机制模型

2.1 多区域融合注意力模块

临床研究表明,AS-OCT 图像上的不同核性区域的混浊度分布不同,并与核性白内障严重程度存在相关性(de Castro 等,2018; Makhotkina 等,2018; Chen 等,2019; Wang 等,2021)。受临床研究发现的启发,本文提出一个问题,即是否可以对不同核性区域的特征表示信息进行融合来提高核性白内障的分类结果。为此,本文设计了一个多区域融合注意力模块(MRA)来动态地学习和融合不同核性区域特征表示信息并突出不同区域特征表示信息的差异性。多区域融合注意力机制模块由多区域融合结构和全连接层(fully-connected layer, FC)两部分组成。图 3 给出了一个区域融合注意力机制模块示例图,融合区域为特征图的上半部分(U_t)和下半部分(U_b),代表核性区域上下两部分。在多区域融合结构中,本文首先将每个特征图分为上下两部分,然后采用全局平均池化(global average pooling, GAP)对特征图上下两部分分别进行压缩操作,得到两个部分的全局特征表示值,其定义为

$$z_t = \frac{1}{W \times H/2} \sum_{i=1}^{H/2} \sum_{j=1}^W U_t(i, j) \quad (1)$$

$$z_b = \frac{1}{W \times H/2} \sum_{i=1}^{H/2} \sum_{j=1}^W U_b(i, j) \quad (2)$$

式中, H 和 W 表示特征图的高和宽, z_t 和 z_b 分别表示特征图上半部分(top half)和下半部分(bottom half)的特征表示信息。

紧随压缩后,采用相加(summation)操作进行区域融合,对特征图上下两部分的全局特征表示值进行相加,其定义为

$$z = z_t + z_b \quad (3)$$

式中, z 是区域融合后的全局特征表示值。

全连接层结构包含两个全连接层。第 1 个全连接层的作用是构建不同特征图之间的依赖关系。第 2 个全连接层的作用是激励,赋予特征图不同的严重程度,计算为

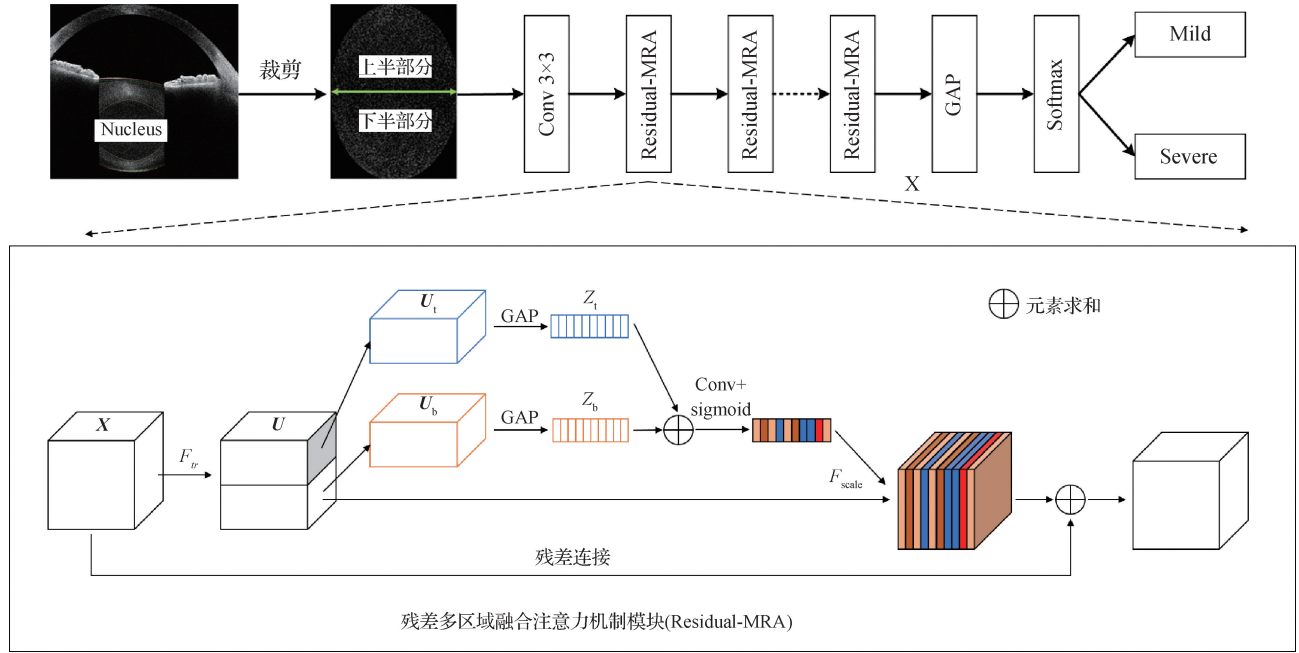


图 2 用于核性白内障自动分类的多区域融合注意力网络框架

Fig. 2 Overall framework of multi-region fusion attention network for automatic nuclear cataract classification

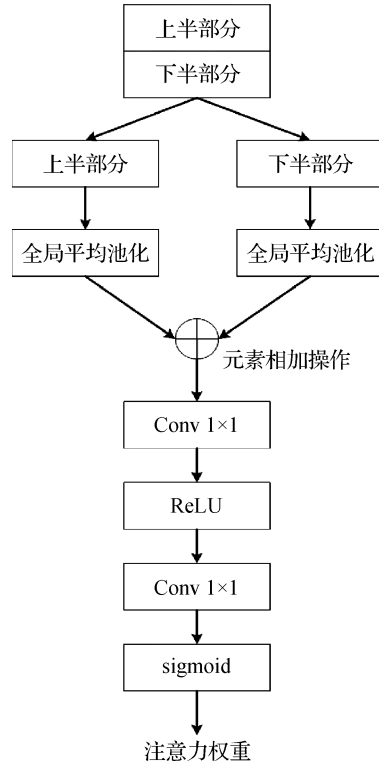


图 3 多区域融合注意力模块结构

Fig. 3 The structure of multi-region fusion attention block

$$s = F_{ex}(g(z, \mathbf{W})) = \sigma(\mathbf{w}_2 \delta(\mathbf{w}_1 z)) \quad (4)$$

式中, F_{ex} 表示激励操作, $\mathbf{w}_1$ 和 $\mathbf{w}_2$ 分别表示两个全连接层的权重矩阵; δ 和 σ 表示 ReLU 和 sigmoid 激

活函数; s 是该操作的输出, 可以刻画特征图的重要程度, 与每个特征图相乘, 得到加权以后的特征图的特征表示信息。

为了验证相加操作的区域融合效果, 本文设计了拼接 (concatenation) 操作。拼接是指将特征图上下部分当做两个不同独立的特征图, 具体拼接方式为

$$\mathbf{V} = \text{Concate}(\mathbf{U}_t, \mathbf{U}_b) \quad (5)$$

式中, Concate 是拼接操作, $\mathbf{V}$ 是拼接以后特征值集合, 拼接操作会导致通道数量加倍。

2.2 网络模型

本文提出的多区域融合注意力机制模型 (MRA-Net) 的总体框架如图 2 所示。该模型以残差卷积神经网络 (residual convolutional neural network, ResNet) 模型为主干网络框架 (backbone network)。ResNet 在很多分类任务中取得了不错的预测结果, 且不少先进的注意力机制方法以 ResNet 为主干网络框架, 可用于验证本文提出的 MRA 模块的有效性。本文将多区域融合注意力模块与残差连接模块融合, 形成多残差多区域融合注意力模块 (residual-MRA), 不仅能自动地学习和调整特征图不同部分的特征表示信息, 还能缓解网络退化问题。图 2 同时给出了一个多区域融合残差模块示例图, 由一个

残差和一个多区域融合注意力机制模块构成。与现有的 CNN 模型一样, MRA-Net 采用全局池化层 (GAP) 替代全连接层, 与全连接层相比, GAP 能取得相近的分类结果, 并且是无参的。分类器是 softmax 函数, 其已经广泛在深度神经网络模型中作为分类器, 输出不同严重程度的核性白内障预测结果。选择通用的交叉熵损失函数 (cross entropy loss function) 为损失函数。

3 实验结果与分析

3.1 评价标准

本文采用准确率 (accuracy, ACC)、宏观敏感度 (macro sensitivity, Sen)、宏观精确率 (macro precision, PR)、宏观 F1 指标 (macro F1 measure, F1) 和 Kappa 系数等评价指标来评估方法的性能。准确率是用来衡量算法预测正确的样本数目占总样本数的比例, 敏感度是指算法预测正确的正例样本的比例; 精确率是算法预测为正例的样本中实际为正例的比例; F1 指标是精确率和敏感度的调和均值。具体计算为

$$ACC = \frac{TP + TN}{TP + FN + TN + FP} \quad (6)$$

$$Sen = \frac{TP}{TP + FN} \quad (7)$$

$$PR = \frac{TP}{TP + FP} \quad (8)$$

$$F_1 = \frac{2 \times Sen \times PR}{Sen + PR} \quad (9)$$

式中, TP、TN、FP、FN 分别代表真阳性 (true positive)、真阴性 (true negative)、假阳性 (false positive) 和假阴性 (false negative)。在核性白内障严重程度的分类任务中, 真阳性表示严重核性白内障样本准确预测为严重核性白内障, 如果错误预测为轻度核性白内障, 则称为假阴性; 真阴性是指轻度核性白内障样本准确预测为轻度核性白内障, 如果错误预测为严重核性白内障, 则称为假阳性。

3.2 实验设置

本文使用 Pytorch 深度学习平台、Python、scikit-learn 和 OpenCV 实现本文模型和基准模型。深度学习模型使用随机梯度下降优化器 (stochastic gradient descent, SGD) 作为优化器并采用默认参数设置, 训练周期为 100 epochs, 批量大小 (batch size) 为

16, 初始学习率 (learning rate) 为 0.1, 每 20 epochs 降低 10 倍, 第 70 epoch 时将学习率固定为 0.000 25。实验硬件环境配置为 1 个 Nvidia Titan GPU, 10 GB DDR 内存, 操作系统为 Ubuntu。

3.3 基准方法

本文模型与以下基准方法进行对比:

1) 机器学习方法。参考章晓庆等人 (2021) 的方法, 本文也基于 AS-OCT 图像从晶状体核性区域提取 18 个像素特征并使用经典的机器学习方法识别不同核性白内障严重程度, 如支持向量机 (SVM)、决策树 (decision tree, DT)、随机森林 (random forest, RF)、逻辑回归 (logistic regression, LR)、GradientBoost 和 Adaboost 等。

2) 挤压激励网络模型 (squeeze-and-excitation network, SENet) (Hu 等, 2018)。SENet 使用一种通道注意力机制来动态调整不同通道的权重, 以此提高卷积神经网络的分类性能。

3) 卷积块注意力组件模型 (convolutional block attention module, CBAM) (Woo 等, 2018)。CBAM 使用通道和空域注意力机制来自适应地调整通道和特征图权重。

4) 其他先进的卷积神经网络模型。包括 ResNet、GraNet、GoogLeNet、VGGNet (Visual Geometry Group Network)、AlexNet、EfficientNet (Tan 和 Le, 2019) 和 SKNet (Li 等, 2019) (Selective Kernel network) 等。所有卷积神经网络模型都采用全局池化层 (GAP) 替代全连接层, 采用现有的主流卷积神经网络模型架构设计。

3.4 与不同注意力机制模型对比

表 2 为本文的 MRA 模块、SE 模块和 CBAM 模块在以人为单位的 AS-OCT 图像数据集上的准确率实验结果对比。结果分析可见, 本文提出的 MRA 模块都优于 ResNet 和先进的注意力机制方法。本文 MRA 模块在准确率上比 ResNet 和 SE 分别提高了约 1.2% 和 1.0%。MRA 模块的设计中引入了临床先验知识, 考虑了特征图不同部分特征的信息差异性。而 SE 模块只考虑整个特征图的特征表示信息, 没有考虑特征图不同部分的特征表示信息。CBAM 模型虽然使用了通道和空域两种注意力机制, 类似于 SE 模块, 但只能学习整体特征图和单个像素的特征表示信息, 而不能有效地学习不同特征图的区域信息。表 2 的实验结果也表明, 不同核

表2 不同注意力机制在个体 AS-OCT

图像数据集的实验结果

Table 2 Experimental results of different attention on
subject's AS-OCT image dataset

方法	ResNet18	ResNet34
Baseline	0.865 9	0.863 9
SE	0.868 7	0.864 8
CBAM	0.869 4	0.865 9
MRA	0.877 8	0.875 0

注:加粗字体表示各列最优结果。

性区域的特征表示信息融合到模型设计中能够提高核性白内障的分类结果。ResNet18 及以 ResNet18 为主干框架的注意力机制网络的分类效果

优于 ResNet34 及以 ResNet34 为主干框架的注意力机制网络。一个可能的原因是 ResNet34 比 ResNet18 具有更多网络层数和更多学习参数,需要更多数据去训练,但本文所能获取的数据总量有限。

图4为基于 MRA 模块在个体 AS-OCT 图像数据集上的两种不同区域融合操作(相加与拼接)的分类结果对比,分别用 MRA-A 和 MRA-B 表示,主干框架是 ResNet18 和 ResNet34。实验结果分析可知,两种融合操作取得相近的分类效果,但基于相加操作的 MRA-A 模块的性能总体上优于基于拼接操作的 MRA-B 模块的性能,表明相加操作能较好地融合不同区域的特征表示信息并验证不同区域特征表示信息的差异性。

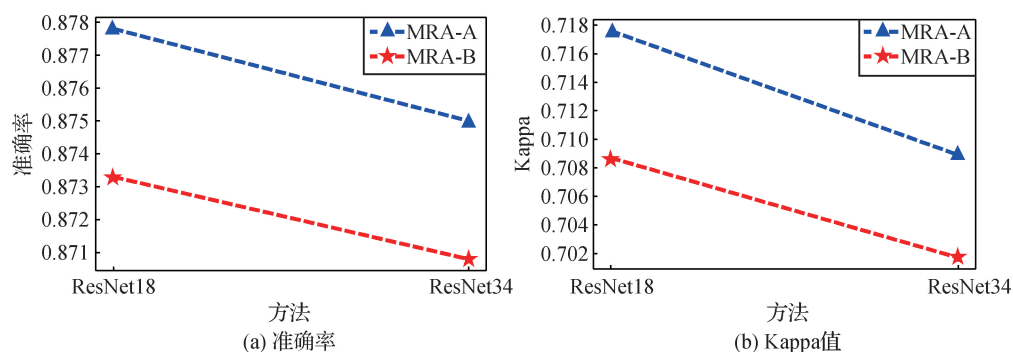


图4 基于 MRA 模块的两种融合操作性能对比

Fig. 4 Performance comparison of two fusion operations based on MRA block((a) accuracy;(b) Kappa)

3.5 与基准方法结果对比

本文 MRA-Net 模型与基准方法在个体 AS-OCT 图像数据集上的分类结果对比如表3所示。

从表3实验结果分析可见,本文模型分类结果优于基准方法,MRA-A-Net-18 的总体分类性能最好,准确率、F1 和 Kappa 值分别为 0.877 8、0.857 5 和 0.717 6。Kappa 和 F1 值比 CBAM 和 SENet 分别提升了约 1.5% 和 1%。与先进的 CNN 模型如 EfficientNet、ResNet、GoogLeNet 和 SKNet 相比,MRA-Net 在 5 个评价指标上都有较大的性能提升,其中 Kappa 评价指标提高了 2% 以上,这是因为本文模型既能利用注意力机制融合特征图的不同区域特征表示信息,又能通过残差连接结构缓解模型退化问题。本文模型在 5 个评价指标上比 SVM 分别提高了 2.72%、3.02%、2.83%、2.86% 和 5.91%,验证了

本文模型的有效性。表3的实验结果也表明深度学习方法在个体 AS-OCT 图像上分类结果总体上优于机器学习方法,因为深度学习方法具有很多参数且能从图像中自动学习和提取有关病理的信息,而机器学习方法主要依赖于提取的特征。SVM 的分类结果优于 SKNet,说明通过提取有效特征也可以取得不错的核性白内障分类结果。表3同时显示了基于 AS-OCT 图像在晶状体区域和核性区域的结果对比。结果表明,GraNet 在核性区域的分类结果比在晶状体区域提升了 28%,说明核性白内障病理信息主要在核性区域,其他区域如皮质性区域对核性白内障分类结果有干扰。本文的机器学习方法分类也好于 RF(章晓庆等,2021)的性能,因为本文使用了更多 AS-OCT 图像数据,也证明数据量是影响分类结果的一个重要因素。

表 3 本文模型与基准模型在个体 AS-OCT 图像数据集上的实验结果对比
Table 3 Classification result comparison of MRA-Nets and baselines based on subject's AS-OCT dataset

方法	ACC	F1	PR	Sen	Kappa
SVM	0.850 6	0.827 3	0.810 2	0.868 9	0.658 5
RF	0.836 9	0.809 0	0.793 4	0.842 5	0.621 3
LR	0.834 2	0.815 0	0.802 4	0.875 3	0.638 7
DT	0.823 0	0.797 9	0.783 7	0.823 9	0.598 0
AdaBoost	0.821 2	0.797 1	0.783 3	0.844 8	0.601 0
GradientBoost	0.838 7	0.812 5	0.796 5	0.850 0	0.628 9
RF(章晓庆 等,2021)	0.755 3	0.744 0	0.751 6	0.708 5	—
GraNet(Zhang 等,2020)	0.578 5	—	—	—	—
AlexNet	0.865 6	0.834 9	0.823 7	0.849 9	0.670 3
VGGNet	0.867 7	0.849 4	0.825 7	0.859 7	0.679 8
ResNet18	0.865 9	0.837 5	0.823 7	0.858 0	0.675 9
GraNet	0.858 9	0.836 2	0.818 5	0.876 3	0.675 7
GoogLeNet	0.868 4	0.846 3	0.828 1	0.884 8	0.695 4
EfficientNet(Tan 和 Le,2019)	0.856 5	0.819 8	0.814 2	0.826 1	0.639 6
SKNet(Li 等,2019)	0.848 1	0.814 4	0.803 0	0.830 4	0.629 5
SENet-18(Hu 等,2018)	0.868 7	0.847 6	0.829 0	0.889 2	0.698 4
CBAM(Woo 等,2018)	0.869 4	0.849 5	0.831 0	0.895 1	0.702 7
MRF-Net-A-34(本文)	0.875 0	0.852 3	0.834 9	0.889 8	0.708 9
MRF-Net-A-18(本文)	0.877 8	0.857 5	0.838 5	0.897 5	0.717 6

注:加粗字体表示各列最优结果,“—”表示缺失数据。

3.6 不同数据集实验结果对比

章晓庆等人(2021)对比分析了 RF 方法在以眼睛和图像为单位的 AS-OCT 图像数据集上的分类结果。结果表明,RF 在以图像为单位的数据集上的结果明显优于在以眼睛为单位的数据集上的结果。因为以图像为单位的 AS-OCT 图像数据集中训练集和测试集都有同一只眼睛的图像数据,会出现数据冗余问题,从而导致算法在以图像为单位的数据集上的分类结果很好。不同于章晓庆等人(2021)的方法,本文通过个体数据集和眼睛数据集验证了以人和眼睛为单位的 AS-OCT 图像数据对分类算法的结果影响。表 4 为表 3 中 10 种性能较好的机器学习和深度学习方法基于眼睛数据集的分类结果,其中深度学习和机器学习方法各 5 种。与表 3 相比,10 种分类算法在眼睛数据集上

的分类效果总体上优于在个体数据集上的分类效果。

为了更好地比较 10 种分类算法在个体数据集和眼睛数据集的分类效果,图 5 给出了本文选择的 10 种方法在两种数据集上的 F1 和 Kappa 结果对比。从图 5 实验结果分析可见,在 5 个机器学习方法中,LR 在眼睛数据集上比在个体数据集上取得了最好的 F1 和 Kappa 结果提升,分别提高 2.1% 和 3.73%;在 5 个卷积神经网络模型中,SKNet 在眼睛数据集上比在个体数据集上取得了最好的分类结果提升,F1 和 Kappa 值分别提高了 4.03% 和 8%。一个可能的原因是一位白内障患者两只眼的白内障严重程度发展是接近或相同的。因为基于青光眼的眼科图像划分通常也基于人,这些研究工作考虑到一个人两只眼睛可能同时患有青光眼(Fu 等,2018,

表 4 基于眼睛 AS-OCT 数据集的 10 种分类算法的核性白内障分类结果
Table 4 Nuclear cataract classification results of ten methods on eye's AS-OCT dataset

方法	ACC	F1	PR	Sen	Kappa
SVM	0.854 1	0.834 0	0.817 7	0.878 5	0.672 4
RF	0.844 5	0.822 1	0.806 4	0.862 5	0.648 4
LR	0.859 0	0.836 6	0.820 0	0.871 6	0.676 0
AdaBoost	0.839 8	0.813 8	0.798 7	0.845 4	0.630 6
GradientBoost	0.845 9	0.825 3	0.807 7	0.863 5	0.651 1
ResNet18	0.878 6	0.842 0	0.853 4	0.832 6	0.684 3
SKNet(Li 等,2019)	0.882 5	0.854 7	0.847 8	0.862 8	0.709 5
SENet(Hu 等,2018)	0.876 1	0.846 1	0.840 6	0.852 2	0.692 2
CBAM(Woo 等,2018)	0.876 4	0.834 4	0.859 7	0.816 8	0.669 8
MRF-Net-34(本文)	0.886 4	0.862 1	0.850 3	0.877 7	0.724 6

注:加粗字体表示各列最优结果。

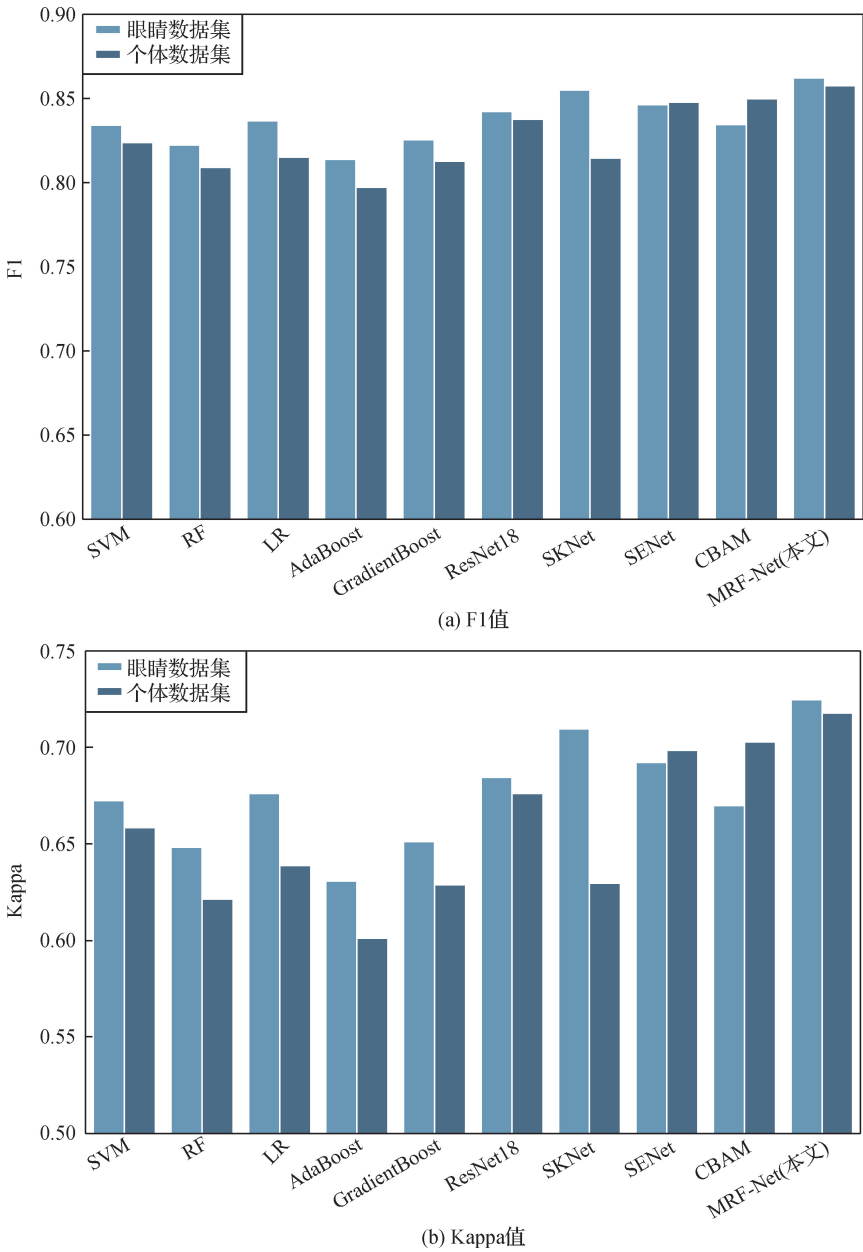


图 5 10 种分类算法在两种 AS-OCT 数据集上的分类效果对比

Fig. 5 Classification performance comparison of ten classification algorithms on two AS-OCT datasets((a) F1 value;(b) Kappa value)

2019),而白内障临床研究工作规范也是基于人为单位采集有关医学数据进行研究(Wong等,2009;Makhotkina等,2018;Chen等,2019),因此本文建议白内障的AS-OCT图像数据集拆分以人为单位。本文得到的实验结果可以为以后基于AS-OCT图像白内障数据集的划分提供参考。

本文模型在基于眼睛的AS-OCT数据集上的分类结果也优于对比方法,验证了本文模型具有较好的鲁棒性。在未来,还将采集更多的白内障患者数据来验证本文不同数据集划分方法的实验结果。

4 结 论

依据AS-OCT图像中不同核性区域的混浊分布差异性特点,本文将临床先验知识融入卷积神经网络模型设计,提出一种多区域融合注意力网络模型,通过对特征图不同区域的特征表示信息进行融合来实现核性白内障严重程度精准预测。在AS-OCT图像数据集上的结果表明,本文模型的准确率、敏感度和F1分别为0.877 8、0.897 5和0.857 5,优于对比方法。同时,本文验证了以人和眼睛为单位的两个AS-OCT图像的白内障数据集对分类结果的影响,并结合现有研究工作建议按人为单位拆分白内障数据集,与临床研究保持一致。

在下一步工作中,将进一步探索晶状体核性区域不同部分对分类效果的影响,并设计轻量级神经网络模型并用于辅助眼科医生基于AS-OCT图像诊断核性白内障。另外,本文目前使用的数据都来自白内障病人,下一步将收集正常受试者数据来验证算法的可靠性和适用性。

参考文献 (References)

- Caixinha M, Amaro J, Santos M, Perdigão F, Gomes M and Santos J. 2016. In-vivo automatic nuclear cataract detection and classification in an animal model by ultrasounds. *IEEE Transactions on Biomedical Engineering*, 63(11): 2326-2335 [DOI: 10.1109/TBME.2016.2527787]
- Cao G P, Zhao W, Higashita R, Liu J, Chen W, Yuan J, Zhang Y B and Yang M. 2020a. An efficient lens structures segmentation method on AS-OCT images//Proceedings of the 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Montréal, Canada: IEEE: 1646-1649 [DOI: 10.1109/EMBC44109.2020.9175944]
- Cao L, Li H Q, Zhang Y J, Zhang L and Xu L. 2020b. Hierarchical method for cataract grading based on retinal images using improved Haar wavelet. *Information Fusion*, 53: 196-208 [DOI: 10.1016/j.inffus.2019.06.022]
- Chen D, Li Z L, Huang J H, Yu L Q, Liu S J and Zhao Y E. 2019. Lens nuclear opacity quantitation with long-range swept-source optical coherence tomography: correlation to LOCS III and a Scheimpflug imaging-based grading system. *British Journal of Ophthalmology*, 103(8): 1048-1053 [DOI: 10.1136/bjophthalmol-2018-312661]
- Chylack L T Jr, Wolfe J K, Singer D M, Leske M C, Bullimore M A, Bailey I L, Friend J, McCarthy D and Wu S Y. 1993. The lens opacities classification system III. The longitudinal study of cataract study group. *Archives of Ophthalmology*, 111(6): 831-836 [DOI: 10.1001/archophth.1993.01090060119035]
- de Castro A, Benito A, Manzanera S, Mompeán J, Cañizares B, Martínez D, Marín J M, Grulkowski I and Artal P. 2018. Three-dimensional cataract crystalline lens imaging with swept-source optical coherence tomography. *Investigative Ophthalmology and Visual Science*, 59(2): 897-903 [DOI: 10.1167/iovs.17-23596]
- Dos Santos V A, Schmetterer L, Stegmann H, Pfister M, Messner A, Schmidinger G, Garhofer G and Werkmeister R M. 2019. CorneaNet: fast segmentation of cornea OCT scans of healthy and keratoconic eyes using deep learning. *Biomedical Optics Express*, 10(2): 622-641 [DOI: 10.1364/BOE.10.000622]
- Fu H Z, Baskaran M, Xu Y W, Lin S, Wong D W K, Liu J, Tun T A, Mahesh M, Perera S A and Aung T. 2019. A deep learning system for automated angle-closure detection in anterior segment optical coherence tomography images. *American Journal of Ophthalmology*, 203: 37-45 [DOI: 10.1016/j.ajo.2019.02.028]
- Fu H Z, Xu Y W, Lin S, Wong D W K, Mani B, Mahesh M, Aung T and Liu J. 2018. Multi-context deep network for angle-closure glaucoma screening in anterior segment OCT//Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention. Granada, Spain: Springer: 356-363 [DOI: 10.1007/978-3-030-00934-2_40]
- Gali H E, Sella R and Afshari N A. 2019. Cataract grading systems: a review of past and present. *Current Opinion in Ophthalmology*, 30(1): 13-18 [DOI: 10.1097/ICU.0000000000000542]
- Gao X T, Lin S and Wong T Y. 2015. Automatic feature learning to grade nuclear cataracts based on deep learning. *IEEE Transactions on Biomedical Engineering*, 62(11): 2693-2701 [DOI: 10.1109/TBME.2015.2444389]
- Hu J, Shen L and Sun G. 2018. Squeeze-and-excitation network//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 7132-7141 [DOI: 10.1109/CVPR.2018.00745]

- Huang W, Li H Q, Chan K L, Lim J H, Liu J and Wong T Y. 2009. A computer-aided diagnosis system of nuclear cataract via ranking//Proceedings of the 12th International Conference on Medical Image Computing and Computer-Assisted Intervention. London, UK: Springer; 803-810 [DOI: 10.1007/978-3-642-04271-3_97]
- Keller B, Draelos M, Tang G, Farsiu S, Kuo A N, Hauser K and Izatt J A. 2018. Real-time corneal segmentation and 3D needle tracking in intrasurgical OCT. *Biomedical Optics Express*, 9(6): 2716-2732 [DOI: 10.1364/BOE.9.002716]
- Li H Q, Lim J H, Liu J and Wong T Y. 2007. Towards automatic grading of nuclear cataract//Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Lyon, France; IEEE; 4961-4964 [DOI: 10.1109/IEMBS.2007.4353454]
- Li J Q, Zhang L L, Zhang L, Yang J J and Wang Q. 2018. Cataract recognition and grading based on deep learning. *Academic Journal of Second Military Medical University*, 39(8): 878-885 (李建强, 张苓琳, 张莉, 杨吉江, 王青. 2018. 基于深度学习的白内障识别与分级. *第二军医大学学报*, 39(8): 878-885) [DOI: 10.16781/j.0258-879x.2018.08.0878]
- Li X, Wang W H, Hu X L and Yang J. 2019. Selective kernel networks//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA; IEEE; 510-519 [DOI: 10.1109/CVPR.2019.00060]
- Long E P, Lin H T, Liu Z Z, Wu X H, Wang L M, Jiang J W, An Y Y, Lin Z L, Li X Y, Chen J J, Li J, Cao Q Z, Wang D N, Liu X Y, Chen W R and Liu Y Z. 2017. An artificial intelligence platform for the multihospital collaborative management of congenital cataracts. *Nature Biomedical Engineering*, 1: #0024 [DOI: 10.1038/s41551-016-0024]
- Makhotkina N Y, Berendschot T T J M, Van Den Biggelaar F J H M, Weik A R H and Nuijts R M M A. 2018. Comparability of subjective and objective measurements of nuclear density in cataract patients. *Acta Ophthalmologica*, 96(4): 356-363 [DOI: 10.1111/aos.13694]
- Ozogkce M, Batur M, Alpaslan M, Yavuz A, Batur A, Seven E and Arslan H. 2019. A comparative evaluation of cataract classifications based on shear-wave elastography and B-mode ultrasound findings. *Journal of Ultrasound*, 22(4): 447-452 [DOI: 10.1007/s40477-019-00400-6]
- Tan M X and Le Q. 2019. Efficientnet: rethinking model scaling for convolutional neural networks//Proceedings of the 36th International Conference on Machine Learning. Long Beach, USA; PMLR; 6105-6114
- Wang W, Zhang J Q, Gu X X, Ruan X T, Chen X Y, Tan X H, Jin G M, Wang L H, He M G, Congdon N, Liu Z Z, Luo L X and Liu Y Z. 2021. Objective quantification of lens nuclear opacities using swept-source anterior segment optical coherence tomography. *British Journal of Ophthalmology* [DOI: 10.1136/bjophthalmol-2020-318334]
- Wong A L, Leung C K S, Weinreb R N, Cheng A K C, Cheung C Y L, Lam P T H, Pang C P and Lam D S C. 2009. Quantitative assessment of lens opacities with anterior segment optical coherence tomography. *British Journal of Ophthalmology*, 93(1): 61-65 [DOI: 10.1136/bjo.2008.137653]
- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer; 3-19 [DOI: 10.1007/978-3-030-01234-2_1]
- Xu C X, Zhu X J, He W W, Lu Y, He X X, Shang Z J, Wu J, Zhang K K, Zhang Y L, Rong X F, Zhao Z N, Cai L, Ding D Y and Li X R. 2019. Fully deep learning for slit-lamp photo based nuclear cataract grading//Proceedings of the 22nd International Conference on Medical Image Computing and Computer-Assisted Intervention. Shenzhen, China; Springer; 513-521 [DOI: https://doi.org/10.1007/978-3-030-32251-9_56]
- Xu X, Zhang L L, Li J Q, Guan Y and Zhang L. 2020. A hybrid global-local representation CNN model for automatic cataract grading. *IEEE Journal of Biomedical and Health Informatics*, 24(2): 556-567 [DOI: 10.1109/JBHI.2019.2914690]
- Xu Y W, Gao X T, Lin S, Wong D W K, Liu J, Xu D, Cheng C Y, Cheung C Y and Wong T Y. 2013. Automatic grading of nuclear cataracts from slit-lamp lens images using group sparsity regression//Proceedings of the 16th International Conference on Medical Image Computing and Computer-Assisted Intervention. Nagoya, Japan; Springer; 468-475 [DOI: 10.1007/978-3-642-40763-5_58]
- Xu Y W, Duan L X, Wong D W K, Wong T Y and Liu J. 2016. Semantic reconstruction-based nuclear cataract grading from slit-lamp lens images//Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention. Athens, Greece; Springer; 458-466 [DOI: 10.1007/978-3-319-46726-9_53]
- Zhang X Q, Fang J S, Xiao Z J, Chen B, Higashita R, Chen W, Yuan J and Liu J. 2021. Research on classification algorithms of nuclear cataract based on anterior segment coherence tomography image [J/OL]. *Computer Science*; 1-10 [2021-07-20] (章晓庆, 方建生, 肖尊杰, 陈滨, Higashita R, 陈婉, 袁进, 刘江. 2021. 基于眼前节相干光断层扫描成像的核性白内障分类算法[J/OL]. *计算机科学*; 1-10). <http://kns.cnki.net/kcms/detail/50.1075.7p.2021/104.1611.002.html>
- Zhang X Q, Xiao Z J, Higashita R, Chen W, Yuan J, Fang J S, Hu Y and Liu J. 2020. A novel deep learning method for nuclear cataract classification based on anterior segment optical coherence tomography images//Proceedings of 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Toronto, Canada; IEEE; 662-668 [DOI: 10.1109/SMC42975.2020.9283218]

作者简介



章晓庆, 1993 年生, 男, 博士研究生, 主要研究方向为深度学习和医学图像处理。

E-mail: 11930927@mail.sustech.edu.cn



刘江, 通信作者, 男, 教授, 主要研究方向为医学图像处理和人工智能。

E-mail: liuj@sustech.edu.cn

肖尊杰, 男, 硕士研究生, 主要研究方向为医学图像处理。

E-mail: 11930387@mail.sustech.edu.cn

东田理沙, 女, 正高级工程师, 主要研究方向为眼科医学图像处理。E-mail: lisahigashita@gmail.com

陈婉, 女, 讲师, 主要研究方向为眼科疾病诊断和眼科人工智能。E-mail: yeah-cw@126.com

胡衍, 女, 助理研究员, 主要研究方向为眼科疾病诊断和人工智能。E-mail: huy3@sustech.edu.cn

袁进, 男, 教授, 主要研究方向为眼科疾病诊断和眼科人工智能。E-mail: yuanjin@cornea@126.com