



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

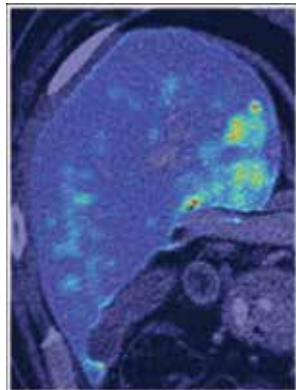
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

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

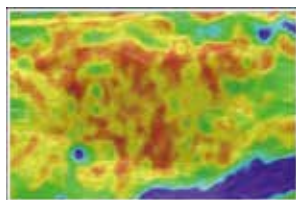
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



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, 黄轲, 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

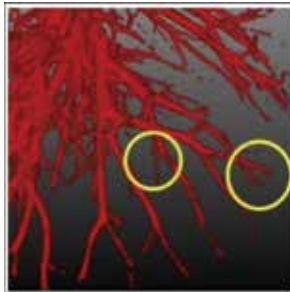
肖照林, 杨志林, 刘欢, 金海燕 0988

基于图像的介入器械前端变形建模与受力估计

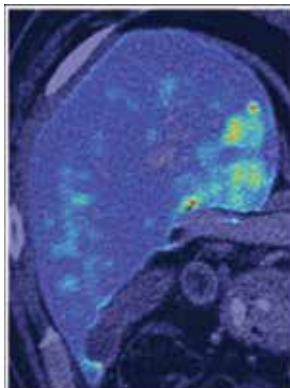
任龙飞, 孟偲, 刘博, 白相志, 周付根 1001

CONTENTS

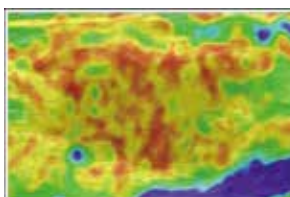
JOURNAL OF IMAGE AND GRAPHICS



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: TP399 文献标识码: A 文章编号: 1006-8961(2022)03-0935-13

论文引用格式: Wang Q D, Ji L W, Qiang Y, Wang H H, Zhao C Q, Li H Z and Zhao Z J. 2022. Multi-branch deep feature fusion method for traditional Chinese medicine intervened human cerebral stroke aided diagnosis. Journal of Image and Graphics, 27(03):0935-0947 (王麒达, 冀伦文, 强彦, 王华虎, 赵琛琦, 李慧芝, 赵紫娟. 2022. 多分支深度特征融合的中医脑卒中辅助诊断. 中国图象图形学报, 27(03):0935-0947) [DOI: 10.11834/jig.210744]

多分支深度特征融合的中医脑卒中辅助诊断

王麒达¹, 冀伦文², 强彦^{1*}, 王华虎³, 赵琛琦¹, 李慧芝⁴, 赵紫娟¹

1. 太原理工大学信息与计算机学院, 晋中 030600; 2. 太原理工大学期刊中心, 太原 030024;
3. 北京大学光华管理学院, 北京 100871; 4. 山西慧虎健康科技有限公司, 太原 030032

摘要: **目的** 脑卒中发病征兆不明显, 发病速度快且致死率高, 目前医学领域的主要诊疗手段仍是针对脑卒中病时和病后, 但在病前预测方面缺少有效办法。中医对于治未病等效果显著, 其中望诊更是中医诊疗的重要方式。本文结合中医望诊, 基于面部与手部图像提出了一种多分支深度特征融合的中医脑卒中辅助诊断方法。**方法** 针对不同部位图像, 分别构建两部位的双分支特征提取模块, 将面部和手部的重点区域作为主体分支提取主要特征。根据中医望诊在面部与手部诊疗的特点, 进一步将眉心的纹理特征和掌心的颜色特征作为辅助信息提取辅助特征; 在此基础上提出信息交互模块 (information interaction module, IIM), 将主要特征与辅助特征进行有效信息交互, 从而辅助主体分支提取更有区别性的信息; 最终将两部位的特征进行融合降维用于脑卒中辅助诊断。**结果** 本文将采集的 3 011 例面部和手部图像数据进行筛查扩充后作为实验数据集, 并在不同评价指标下与当前主流的分类模型进行对比。实验结果表明, 本文方法在准确性上达到了 83.36%, 相比 ResNet-34、DenseNet121、VGG16 (Visual Geometry Group 16-layer net) 和 InceptionV3 等其他主流分类模型性能提高了 3%~7%; 在特异性和敏感性方面分别为 82.47% 和 85.10%, 其效果优于对比方法。**结论** 本文方法能够有效结合中医望诊的诊疗经验并通过传统面部和手部图像实现对脑卒中的有效预测, 为中医望诊在脑卒中方面的客观化和便捷化发展提供了帮助。

关键词: 中医望诊; 脑卒中; 图像识别; 特征提取; 特征融合; 卷积神经网络

Multi-branch deep feature fusion method for traditional Chinese medicine intervened human cerebral stroke aided diagnosis

Wang Qida¹, Ji Lunwen², Qiang Yan^{1*}, Wang Huahu³, Zhao Chenqi¹, Li Huizhi⁴, Zhao Zijuan¹

1. College of Information and Computer, Taiyuan University of Technology, Jinzhong 030600, China; 2. The Journal Center, Taiyuan University of Technology, Taiyuan 030024, China; 3. Guanghua School of Management, Peking University, Beijing 100871, China;
4. Shanxi Huihu Health Science and Technology Company with Limited Liability, Taiyuan 030032, China

Abstract: **Objective** Stroke is a severe human cerebrovascular disease that causes brain tissue damage due to sudden rup-

收稿日期: 2021-08-20; 修回日期: 2021-11-23; 预印本日期: 2021-11-30

* 通信作者: 强彦 qiangyan@tyut.edu.cn

基金项目: 国家自然科学基金项目 (61872261); 虚拟现实技术与系统国家重点实验室开放课题项目 (VRLAB2020B06); 细胞生理学教育部重点实验室开放基金资助项目 (KLMEC/SXMU-201905); 山西省重点研发计划项目 (201903D321034)

Supported by: National Natural Science Foundation of China (61872261); Open Project of State Key Laboratory of Virtual Reality Technology and System (VRLAB2020B06); Open Foundation Project Key Laboratory of Cell Physiology of Ministry of Education (KLMEC/SXMU-201905); Key Research and Development Project of Shanxi Province (201903D321034)

ture of blood vessels or vascular obstruction originated blood flow inefficiency. The incidence of ischemic stroke is high frequency, high recurrence rate and high fatality rate. Traditional Chinese medicine (TCM) has its priority of stroke. In the four-diagnosis-inspection, listening, asking and feeling the pulse- (in Chinese) of Chinese medicine, the prime step of inspection shows that information extraction is one of the essential factors of TCM. But, TCM has its constraints of medical standardization and manual factor issues. A deep learning technology is benefit to further recognize the constraints of TCM.

Method Our research illustrates a dual-branch cross-attention feature fusion model (DCFFM) based on facial images and hand images. It can assists in predicting stroke disease well. The overall model is divided into three parts: facial feature extraction module, hand feature extraction module and feature fusion module. For the facial feature extraction module, we construct the subject branch and auxiliary information branch to extract facial features. In accordance with the guidance of Chinese medicine doctors, we pre-process the facial image and use the key diagnostic area of the stroke in the face as the input of the main branch. In addition, we also integrate the knowledge of inspection of TCM to cut out the image of the area around the eyebrows, and use the Sobel filter to extract the gradient image as the input of the auxiliary information branch. For the hand feature extraction module, this demonstration adopts the same double-branch structure to cut out the palm area as the input of the main branch. In order to more stably and accurately reflect the pathological condition of the hand and reflect the small changes in the characteristics of the hand, we convert the palm image from the RGB color space to the HSV color space and transfer it to the differentiated auxiliary information branch. The proposed branches are respectively input to their respective convolution blocks, and the depth characteristics of the input data are extracted based on the convolution operation. Max pooling is used on the feature maps and batch normalization is used to prevent the model from over-fitting. In addition, we use two loss functions to constrain the training of the two feature extraction modules, and use the total loss to constrain the entire model. Between the two branches of each feature extraction module, we built an information interaction module (IIM) for further information interaction amongst the branches to reveal the model extract distinctive features. It assigns a certain weight to the feature map of auxiliary information and then interconnected with the feature of the subject branch. We use 1×1 convolution fusion to reduce dimensionality. Under no special operations circumstances, the IIM can be trained in an end-to-end manipulation. For the feature fusion module, multiple convolutional layers are used for overall fusion dimensionality reduction to generate the prediction result via the multi-branch deep feature fusion based on fusing the depth features of the facial feature extraction module and the hand feature extraction module. **Result** In order to aid model training and improve the stability and robustness of the model, our demonstration screens and extends the collected 3 011 face and hand image data. We remove some scrambled images with scars and conduct the data extension by horizontal flipping. We remove some images of peeling, disability, background clutter and implement random option for the remaining images to expand the data by horizontal flipping. It is determined that 3 964 images of positive and negative samples are involved as the experimental data set. Our multiple sets of comparative experiments and ablation experiments have been facilitated based on a variety of evaluation indicators to verify the performance of the model, such as accuracy, specificity, sensitivity, and F1-score. First, we compare the overall performance of the proposed method with the current mainstream classification algorithms. Experimental results show that the accuracy of the method proposed reaches 83.36%, which is 3%—7% higher than the performance of other mainstream classification models. Based on ten-fold cross-validation, the specificity and sensitivity reached 82.47% and 85.10% respectively. The illustrated sensitivity shows a relatively large advantage, indicating that the method in this paper has a better performance for the detection of true positives. Next, we still verified the impact of facial feature extraction module, hand feature extraction module and IIM on the performance of the model. This analyzed results show that feature extraction of face data and hand data can effectively improve the performance of the model simultaneously. In addition, the IIM has targeted the sensitivity and specificity of the model. **Conclusion** Our method can use human facial features and human hand data to assist in stroke prediction, and has good stability and robustness. Meanwhile, the demonstrated IIM also promotes the information interaction between multi-branch tasks.

Key words: inspection diagnosis of traditional Chinese medicine; stroke; image identification; feature extraction; feature fusion; convolutional neural network

0 引言

脑卒中是一种由于脑部血管突然破裂或因血管阻塞导致血液不能流入大脑而引起脑组织损伤的一种急性脑血管疾病。其中缺血性卒中(又称脑梗)发病率最高(Zhou等, 2019), 占脑卒中发病率的70%以上(Feigin等, 2015)。近年来, 虽然诊疗技术不断发展, 但由于其发病速度快、复发率高且致死率高, 仍是各国的主要死亡原因之一。由于临床缺乏有效的诊疗手段, 即使进行针对性治疗, 患者也会出现一系列并发症, 无法有效恢复。目前普遍认为预防是最好的措施, 因此, 如何在发病前及时检测, 提醒患者预防与保养是医学领域急需解决的重大问题。然而, 临床检测存在如下弊端: 脑卒中患者发病前征兆不明显, 尽管患者会在发病前出现一侧肢体无力或麻木、说话不清和头晕头痛等症状, 但这些并非典型临床症状, 患者往往不会重视。

不同类型的脑卒中治疗方式也存在差异, 磁共振成像(magnetic resonance imaging, MRI)和计算机断层扫描(computed tomography, CT)是临床筛查诊断的有效方式(Zhang等, 2005)。近年来, 许多研究专注于脑部的相关病症。Cho等人(2019)提出级联深度学习模型以辅助真的颅内出血; Kuo等人(2019)利用4 396幅CT图训练了一个全卷积神经网络; Chilamkurthy等人(2018)开发并验证了一套深度学习算法以自动检测颅内出血。这些研究针对的患者不是已经发病, 就是出现一些严重的临床表现, 即使治疗也极易复发。因此, 如何针对性有效预防脑卒中显得极为重要。

中医经过长期不断发展, 对脑卒中已经有了较为完整的认识。其中中医四诊望、闻、问、切中将“望诊”排在首位, 不仅是中医诊病中提取患者信息的首选方式, 更是突出了望诊在中医诊病中的重要地位。《黄帝内经·灵枢·本脏》中有“有诸于内, 必行于外”、“视其外应, 以知其内脏, 则知其病矣”等记载, 《灵枢·九宫八风论》中也有“三虚而偏中于邪风, 称之为击仆偏枯”等关于脑卒中的记载, 因此无论是脑卒中哪一种类型, 人体的内脏都会与身体的各外部部位产生一种隐形的映射关系, 通过外部特征的变化间接了解身体内脏的变化情况。因此许多脑卒中患者更加倾向于由传统中医医师通过便

捷的望诊进行初步诊断, 但传统中医进行望诊需要诊断医师具备多年的行医经验, 并且不同医师在手诊标准上会有些许差异, 出现医生数量少、患者数量多、患者等候时间长以及医师诊断压力大等问题。许多研究也验证了人体外部表征与体内各器官间的变化有着不可或缺的关系。Zhang等人(2017)提出一个基于几何特征的自动识别和分析舌型的分类方法。Li等人(2007)通过CIELAB(CIE $L^*a^*b^*$)颜色空间提出一种基于中医理论的舌色分类数值方法。Kim等人(2008)提出一种利用面部图像周围颜色分布的心脏病诊断系统。Li等人(2019)提出多模态的学习方法来利用不同体表特征的相关性, 以此辅助检测糖尿病。Su等人(2011)对肺癌患者舌象进行定性和定量分析后对肺癌患者进行分类。除此之外, Wang等人(2017a)对脉诊进行相应研究, 通过提取脉冲特征辅助糖尿病诊断。Wang等人(2014)实现了多通道脉冲信号采集的复合系统, 能够分析病人的健康状况并且性能良好。Yan等人(2014)描述了一种新颖的呼吸分析系统, 通过分析一些指标浓度实现糖尿病患者的检测。Lin等人(2020)开发了一种基于人脸照片的CAD(coronary artery disease)深度学习算法。Zhang等人(2021)提出了双堆栈网络, 通过分析患者面部图像区分慢性肾病和糖尿病。韩学杰等人(2018)从医学角度验证了动脉血管病与手部等身体部位有着不可忽视的联系, 因此, 如何通过深度学习技术将中医望诊进一步客观化, 进而有效辅助中医进行脑卒中预测成为研究的重中之重。

本文基于面部图像与手部图像提出一种双分支交叉注意特征融合模型(dual-branch cross-attention feature fusion model, DCFFM), 能够很好地辅助传统中医进行脑卒中疾病的预测。本文模型分别构建面部与手部的深度特征提取任务模块, 在各模块中构建主体分支与辅助分支。除此之外, 在模块的双分支间构建了信息交互模块, 进一步帮助模型提取具有区别性的特征。最终, 将各模块间的特征进行特征融合以实现预测。

本文的主要贡献有: 1) 提出一种基于面部图像与手部图像的深度特征融合辅助中医预测脑卒中的方法, 首次尝试将面诊与手诊相结合进行脑卒中病症预测。2) 设计了一种多分支深度特征融合学习模型, 将两个部位重点诊断区域作为辅助信息, 并构建

一种新的信息交互模块 (information interaction module, IIM), 以一种新颖的方式结合不同分支的特征分量, 使模型更有针对性地进行训练。3) 通过实验验证了本文方法的可行性。结果表明, 相比于传统的分类方法, 本文模型在准确率上可以达到 83% 左右。

1 方法

本文模型的整体架构如图 1 所示, 整体分为面

部特征提取模块、手部特征提取模块和特征融合模块 3 部分。首先, 根据中医医师的指导, 对采集的面部图像和手部图像进行预处理, 截取出脑卒中在面部和手部的重点诊断区域, 并依据中医望诊经验进行对应图像处理作为辅助信息分支的输入, 将原始图像作为主体分支的输入。接着, 通过模型训练进行两个分支的深度特征提取与融合, 并通过信息交互模块实现分支间的信息共享。最后, 将两部位的深度特征进行整体融合降维, 输出多分支深度特征融合下的预测结果。

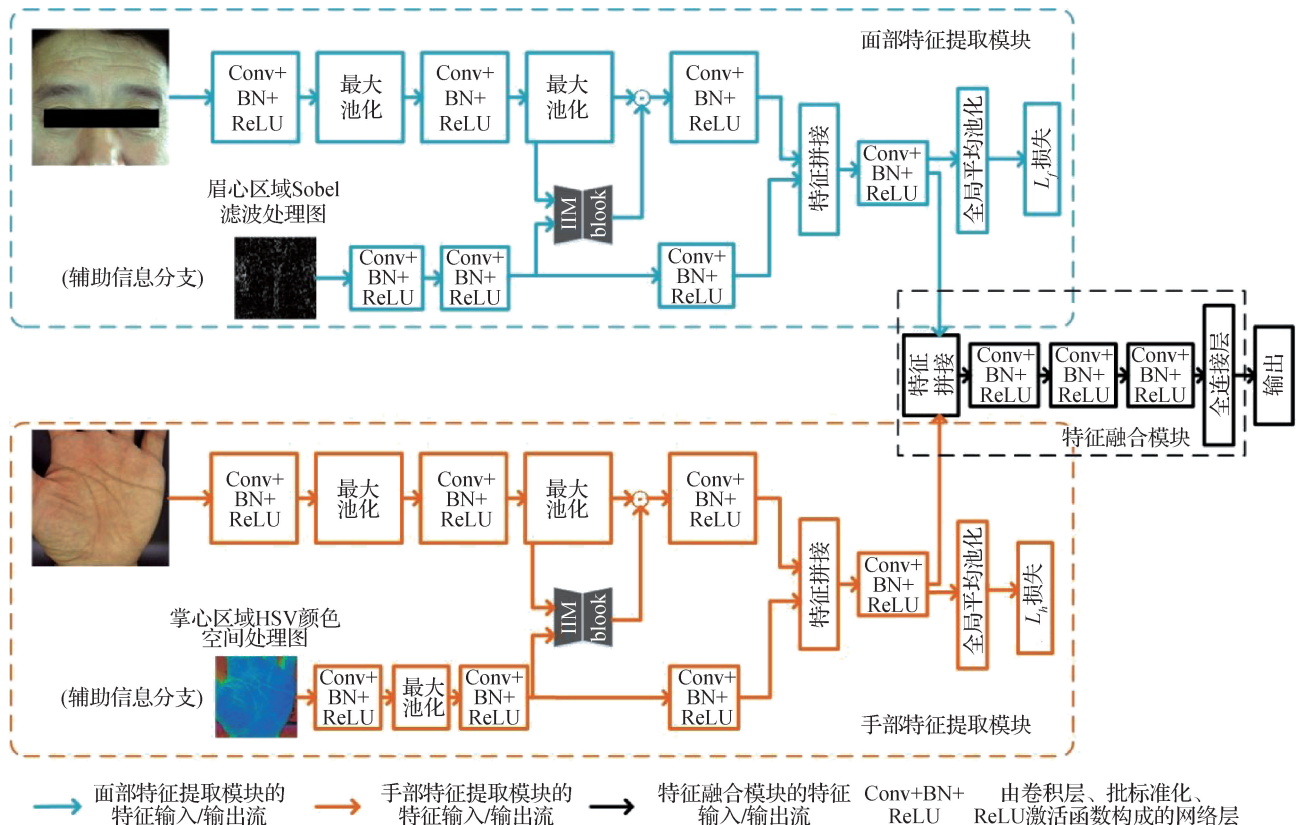


图 1 DCFFM 模型整体架构图
Fig. 1 Overall network structure of DCFFM

1.1 数据预处理

本文收集由专业设备拍摄的面部与手部高清 RGB 图像。模型的输入需尽可能只包含脸部和手部区域以学习正确的映射关系。基于此, 在专业中医医师指导下, 设计了一种确定图像尺寸大小方法, 以确保在平衡网络参数数量的同时尽可能多地获得有用信息。

对于手部图像, 根据中医医师传统手诊经验, 脑卒中的手诊重点诊断区域集中在掌丘、大小鱼际等掌心区域, 如图 2 所示, 将中指与手掌交界点作为

Y_{h1} , 手掌与手腕交界点作为 Y_{h2} , 大拇指与手掌交界点作为 X_{h1} , 手掌掌心靠小拇指一侧的边界处作为 X_{h2} , 该点平行对应于大拇指与手掌交界点 X_{h1} , 则 ΔX_h 和 ΔY_h 为

$$\begin{cases} \Delta X_h = |X_{h1} - X_{h2}| \\ \Delta Y_h = |Y_{h1} - Y_{h2}| \end{cases} \quad (1)$$

最后, 选择两值中较大的一项得到 P_{size} , 将 P_{size} 作为掌心区域图像的高和宽并进行截取。具体为

$$P_{size} = \max(\Delta X_h, \Delta Y_h) \quad (2)$$

对于面部图像的预处理如图 3 所示, 首先将额

上与发际线交界处的中心点标记为 Y_{f1} , 将鼻尖标记为 Y_{f2} , 则 ΔY_f 为

$$\Delta Y_f = |Y_{f1} - Y_{f2}| \quad (3)$$

接着, 按照 ΔY_f 的大小截取面部的图像。在中医面诊中, 对于脑卒中的诊断重点区域主要集中于眉心等周边区域。因此, 为了进一步促进网络提取

有效特征, 进一步对面部图像进行截取, 从而得到眉心辅助区域图像。将面部两眉毛靠中心方向的边界点分别标记为 X_{f1} 和 X_{f2} , 则 ΔX_f 为

$$\Delta X_f = |X_{f1} - X_{f2}| \quad (4)$$

最后, 将 ΔX_f 作为眉心辅助区域图像的宽和高进行图像截取, 得到脸部图像和眉心辅助区域图像。

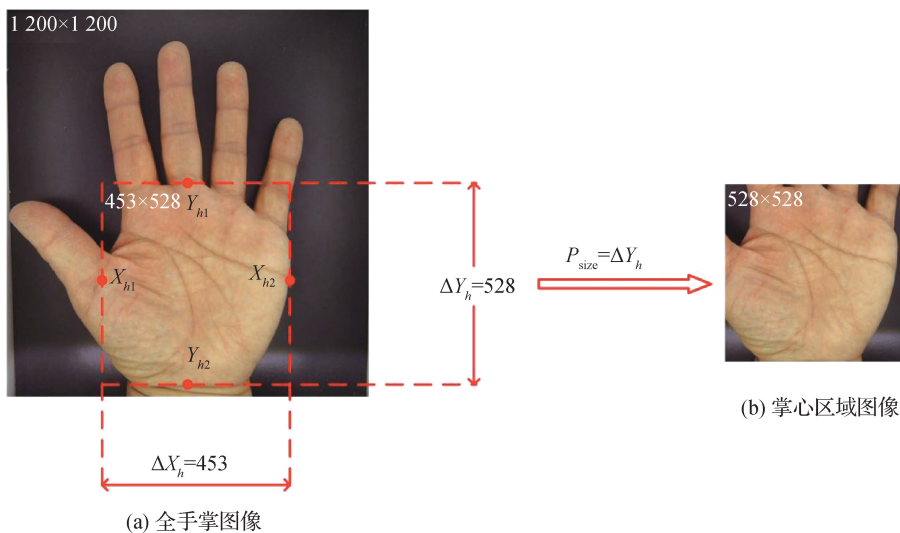


图2 手部图像预处理

Fig. 2 Hand image preprocessing ((a) full palm image; (b) palm area image)

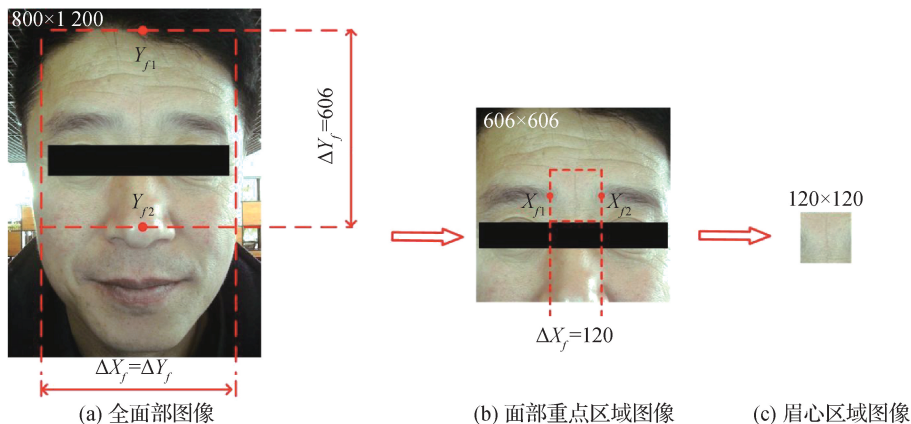


图3 面部图像预处理

Fig. 3 Facial image preprocessing ((a) facial image; (b) face key area image; (c) image of brow center area)

手部图像选取 P_{size} 后, 若是由 ΔY_h 确定 P_{size} , 则在水平方向上以 X_{h1} 与 X_{h2} 的中点作为截取的出发点, 向水平左右方向同步进行截取, 直到截取宽度与 ΔY_h 相同; 若是由 ΔX_h 确定 P_{size} , 则在垂直方向上以 Y_{h2} 为出发点, 垂直向上进行截取, 直到截取高度与 ΔX_h 相同。对于面部区域截取, 本文采取与掌心图像截取相同的方式。对于眉心区域, 在确定好 ΔX_f 后, 在垂直方向上以 X_{f1} 与 X_{f2} 的中点为出发点, 向垂

直上下两个方向同时进行截取, 直到截取高度等于 ΔX_f 。

为了尽可能统一掌心区域、面部区域以及眉心区域的尺寸, 且保证重点诊断区域均包含在截取后的图像中, 本文统计了所有手掌图像、面部图像、眉心图像按数据预处理后对应的尺寸, 如表1所示, 最终发现大多数掌心和面部图像尺寸都小于530, 眉心图像的尺寸都小于132, 因此, 本文将模型输入的

掌心图像和面部图像的尺寸统一为 530,眉心图像的尺寸统一为 132,从而尽可能减少图像尺寸不统一对模型性能造成的影响。

表 1 截取图像数量统计

Table 1 Statistics of the number of captured images

不同部位图像	不同图像尺寸范围下的图像数量				
	≤80	≤132	≤364	≤440	≤530
掌心区域	—	—	24	689	765
面部区域	—	—	—	427	1 064
眉心区域	245	1 246	—	—	—

注:“—”表示无此范围内的数据。

1.2 面部和手部图像的主体分支与辅助信息分支

如图 1 所示,本文提出的模型有面部特征提取、手部特征提取和特征融合 3 个模块。其中面部特征提取模块和手部特征提取模块都构建了双分支,即主体分支与辅助信息分支。

主体分支中,将面部和手部的重点区域图像作为分支的输入,包含最多的信息。由于专业拍摄的图像分辨率(1 200 × 1 200 像素)高于一般图像,如果将完整图像作为网络的输入,则需要很深的神经网络进行降维。然而当网络中使用过多卷积层时,就会产生大量参数,导致模型过拟合问题。如果不想使用过多卷积层,就需要使用几次下采样来减小特征图的大小以便符合网络的尺寸要求,从而学习更深的特征用于预测,但是多次下采样又可能会丢失很多有用的特征。除此之外,原图中包含很多无关信息,在中医望诊中,脑卒中在面部与手部的诊断也主要集中在额头、眉心、掌丘和鱼际等部位,医生通过观察这些部位是否有异常现象(眉心有竖纹、两眉间凹陷发红、掌丘隆起肥厚和大小鱼际红肿等症状)来提前预测患者是否有脑卒中的风险。因此,主体分支的输入应尽可能避免其他信息的干扰,从而学习正确的映射关系。

辅助信息分支中,针对面部图像与手部图像在中医望诊中的区别,对辅助信息分支的输入进行不同的处理。对于面部图像,眉心部位的纹路状况对于脑卒中的诊断尤为重要。因此,对截取的眉心区域图像做进一步处理,为了避免图像中一些细小纹路的噪声干扰,本文针对眉心区域的辅助分支,采用 Laplace、Canny 和 Sobel 这 3 种不同的滤波器得到眉

心区域的纹路图像。Laplace 滤波器对边缘信息较为敏感,某些噪声边缘也会提取出来,因此并不适合眉心区域的纹路提取。Canny 滤波器产生的边缘信息没有强弱之分并且检测算法较为复杂,作为辅助分支,本文希望辅助分支在提取有效特征的同时也注重检测效率。而 Sobel 滤波器的检测效率在实际应用中比 Canny 滤波器高,并且滤波后的图像边缘有强弱之分,具有良好的抗噪性,因此 Sobel 滤波器是本文方法的最优选择。本文方法首先对图像进行基于中值滤波的图像平滑,之后采用 Sobel 滤波器(Sobel 和 Feldman,1973),通过图像水平和垂直两个方向的梯度来逼近图像中每一个像素的梯度,从而得到梯度图像。因为对于方向复杂的眉心纹路,单靠水平或垂直方向的纹路不足以有效提取眉心区域的纹理特征,而将横向与纵向梯度进行结合的梯度图像更有助于本文模型完成任务。具体计算为

$$G = \sqrt{G_h^2 + G_v^2} \tag{5}$$

式中, G_h 为图像的水平梯度, G_v 为图像的垂直梯度。对于手部图像,手掌的掌丘和鱼际等部位的颜色、形态状况对于脑卒中的诊断具有重要参考意义。传统的手部图像是基于 RGB 颜色空间,易受光照、遮挡和阴影等情况的影响,而 HSV (hue, saturation, value)颜色空间由色相(hue)、饱和度(saturation)和明度(value)3 部分组成,Cheng 等人(2001)证明 HSV 颜色空间可以很好地降低这些因素的影响,鲁晓篝等人(2019)也利用此颜色空间进行舌诊客观化的研究。因此,本文将 RGB 颜色空间的掌形图像转化为 HSV 颜色空间图像作为辅助信息分支的输入,从而更稳定准确地反映手部生理病理状况,可以体现出手部特征的微小变化。图 1 左侧展示的即是处理后的眉心梯度图像以及掌心 HSV 颜色空间图像。

对于中医望诊在面部与手部的相关诊断知识,眉心部分的纹路特征以及掌心部位的形态特征是传统中医医师进行望诊的重点。因此在面部特征提取模块的眉心辅助分支中采用 Sobel 等滤波器进行实验而不采用转化颜色空间等方法,这样可以更好地关注眉心的纹路特征,从而更好地辅助面部特征提取。对于手部特征提取模块,本文同样只转化不同颜色空间进行实验,因为掌心纹路复杂且手掌掌纹

非常容易干扰模型提取有效特征。

值得注意的是,虽然本文在每个模块中只使用了两个分支,但仍然可以引入很多新的分支。最终,主体分支输入数据从 $X_m \in \mathbf{R}^{H,W,C}$ 转变为 $X_m \in \mathbf{R}^{S,H,W,C}$, 辅助分支输入数据从 $X_a \in \mathbf{R}^{H,W,C}$ 转变为 $X_a \in \mathbf{R}^{S,H,W,C}$, 其中 S 为分支个数, H 和 W 分别为分支的高度和宽度, C 为通道数。将上述分支分别输入到各自的卷积块中,通过卷积操作提取输入数据的深度特征,并对特征映射进行最大池化(maxpooling)操作,对特征映射进行下采样。这一部分中,使用 BN(batch normalization)(Ioffe 和 Szegedy, 2015)防止模型过拟合。BN 通过引入一对可学习参数 γ 和 β 确保模型能够根据数据的真实分布尽可能地进行学习。接着将主体分支和辅助分支的特征映射输入到信息交互模块中融合成为一个新的注意特征映射,并将其与主体分支的输出进行哈达玛积。然后将两分支的特征映射进行拼接,并通过 3×3 的卷积进一步对融合特征进行提取,最终使用全局平均池化(global average pooling, GAP)(Lin 等, 2013)取代卷积神经网络中传统的全连接层,为每个类别生成相应的特征图。

1.3 信息交互模块(IIM)

信息交互模块(IIM)的整体结构如图4所示。在这一部分中,为了使辅助分支更好地指导主体分支进行特征提取,本文设计了一种新的方法让不同任务分支之间进行信息交互。在辅助分支与主体分支间有许多有利于其他任务分支的信息,仅仅依靠深层的特征映射拼接无法很好地利用这些信息。因此,构建信息交互模块,使辅助信息帮助主体任务

更有效地捕捉对特定任务有用的信息,而不是允许它们在没有任何信息交换的情况下自由拟合。首先将辅助信息特征映射赋予一定的权重,之后将两个分支具有相同空间大小的特征连接起来,最后对融合特征进行降维,实现跨通道信息交互。如图4所示, IIM 的输入主要分为主体分支特征映射和辅助信息分支特征映射两部分。主体分支特征映射表示为 $F_f \in \mathbf{X}_m^{1,H,W,C}$ 和 $F_h \in \mathbf{X}_m^{1,H,W,C}$, 其中 F_f 为面部主体特征映射, F_h 为手部主体特征映射; 辅助信息分支特征映射表示为 $F_{fa} \in \mathbf{X}_a^{1,H,W,C}$ 和 $F_{ha} \in \mathbf{X}_a^{1,H,W,C}$, 其中 F_{fa} 为眉心辅助特征映射, F_{ha} 为掌心辅助特征映射。对于权重 λ , 将面部分支 IIM 中的 λ 表示为 λ_1 , 将手部分支 IIM 中的 λ 表示为 λ_2 , 通过实验验证, 本文将 λ_1 设置为 0.7, λ_2 设置为 0.4。最终 IIM 的输出可以定义为

$$\begin{cases} IIM_f = (T_{1 \times 1}(F_f \oplus \lambda_1 F_{fa}) + 1) \odot F_f \\ IIM_h = (T_{1 \times 1}(F_h \oplus \lambda_2 F_{ha}) + 1) \odot F_h \end{cases} \quad (6)$$

式中, $T_{1 \times 1}$ 表示使用 BN、卷积核为 1×1 、stride = 1 的卷积层; $\oplus$ 表示连接运算; $\odot$ 表示哈达玛积; “+1”操作是受 Wang 等人(2017b)和 He 等人(2016)方法启发引出的剩余身份映射,避免可能由连续的逐层乘法引起的梯度爆炸或消失,以帮助网络学习更鲁棒的注意力映射。值得注意的是,使用 1×1 的卷积层是因为 1×1 的卷积核能够仅根据通道执行计算,而不是融合不同空间位置的特征或改变特征的空间大小,在实现降维的同时也实现了通道间的信息交互。通过在 1×1 卷积权重 W 上反向传播特定任务损耗和 L2 权重衰减损耗,在没有任何特殊操作的情况下,能以端到端的方式训练 IIM。

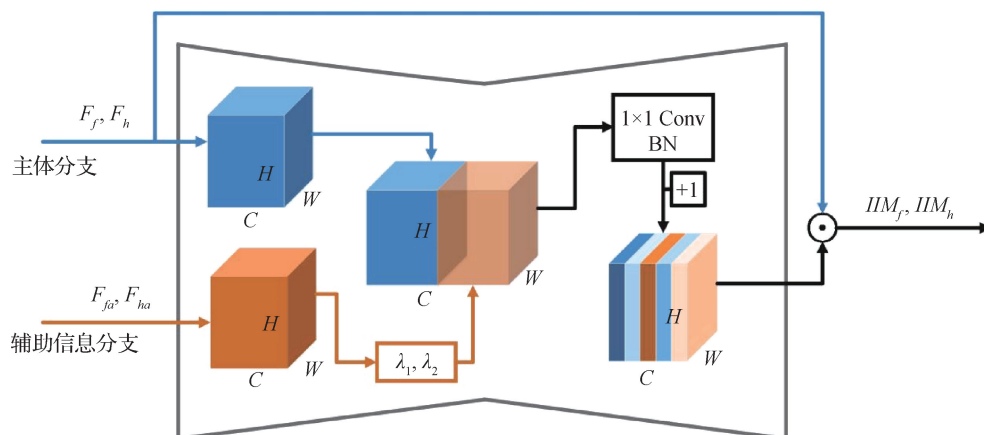


图4 信息交互模块(IIM)

Fig. 4 Information interaction module(IIM)

1.4 损失函数

为了更好地约束面部特征提取模块和手部特征提取模块的训练,在两个模块的最后一层均采用 GAP 代替卷积神经网络中传统的全连接层,并且针对两个模块分别设计了各自的损失函数。由于两个部位的特征提取模块均为分类任务,因此面部特征提取模块和手部特征提取模块的损失函数 L_f 和 L_h 均使用交叉熵损失。具体为

$$L_f = -[y_f \times \log(p_f) + (1 - y_f) \times \log(1 - p_f)] \quad (7)$$

$$L_h = -[y_h \times \log(p_h) + (1 - y_h) \times \log(1 - p_h)] \quad (8)$$

式中, p_f 是面部特征提取模块最后经过 softmax 函数之后的值, p_h 是手部特征提取模块最后经过 softmax 函数之后的值, $y_f \in \{0,1\}$ 是患者面部真实的标签, $y_h \in \{0,1\}$ 患者手部真实的标签。

最终的总损失函数 L_{total} 为

$$L_{\text{total}}(\sigma_f, \sigma_h) = \frac{1}{2\sigma_f^2} L_f + \frac{1}{2\sigma_h^2} L_h + \log \sigma_f^2 + \log \sigma_h^2 \quad (9)$$

式中, σ_f 和 σ_h 均为可学习的观测噪声参数,该值与分布的不确定性有关。

总损失函数受 Cipolla 等人(2018)方法影响,采用权重不确定的方法设置任务权重,该方法可以统一多任务各损失的尺度且不再受额外权重超参数的影响。

2 实验与结果分析

2.1 数据集

本文使用的实验数据集来自合作中医药公司的脑卒中患者病例,收集了包括 3 011 例脑卒中患者的脸部与手部高分辨率图像,并由具有 20 年丰富中医诊疗经验的医师进行标签标注。本文按照如下规则进行数据筛查:1) 手部具有脱皮现象的图像($n = 17$);2) 手掌区域具有伤疤的图像($n = 5$);3) 面部具有明显伤疤的图像($n = 12$);4) 无法完全露出眉心额头等区域的图像($n = 8$)。筛查过后的数据集包含 2 969 例脑卒中患者的脸部与手部图像,其中正样本 995 例,负样本 1 974 例。考虑到正负样本数量不平衡,不利于模型训练,对正样本图像采取水平翻转的方式进行数据扩充。采取这种方法是希望尽可能靠近真实患者的手部情况,而水平翻转可以

很好地满足这样的需求,其他的数据扩充方式(裁剪、旋转和局部变形等)都可能或多或少地损失或改变一些手部及面部特征。通过增加数据量的方式,可以更好地让模型学习到数据背后的隐含规律,有效提高模型的泛化能力。最终本文所用数据集包含 3 964 例脑卒中患者的脸部与手部图像,其中正样本 1 990 例,负样本 1 974 例。

2.2 实现细节

实验均在一个工作站进行,工作站配备 Ubuntu 18.04 LTS 系统,1 个 2.90 GHz Intel(R) Xeon(R) W-2102 CPU 和 1 个 NVIDIA TITAN XP GPU。网络模型使用 python3.6 搭建,通过 pytorch 在工作站展开实验。实验采用 10 折交叉验证评估模型,将数据集分为 10 组,每折使用任意 9 组作为训练集,剩下的 1 组作为测试集,依次迭代。优化器采用 mini-batch Adam,其中初始学习率设置为 0.01, batchsize 设置为 32,模型迭代次数为 500。为了获得更好的模型性能,本文采用的学习率变化方法为

$$l(s) = \begin{cases} l(0) & 0 \leq s \leq 100 \\ l(0) \times 0.5 & 100 < s \leq 150 \\ l(0) \times 0.1 & 150 < s \leq 250 \\ l(0) \times 0.01 & 250 < s \leq 400 \\ l(0) \times 0.001 & 400 < s \leq 500 \end{cases} \quad (10)$$

式中, s 为迭代次数,初始学习率 $l(0)$ 为 0.01。在训练过程中,总损失最小的模型保存到验证集,以便在测试集上进行验证。为了避免该数据集的轻微不平衡和过度拟合问题,通过准确性(accuracy)、敏感性(sensitivity)、特异性(specificity)和 F1-score 来评估性能。它们的数学定义为

$$f_{\text{Acc}} = \frac{T_p + T_n}{T_p + F_p + T_n + F_n} \times 100\% \quad (11)$$

$$f_{\text{Sen}} = \frac{T_p}{T_p + F_n} \times 100\% \quad (12)$$

$$f_{\text{Spe}} = \frac{T_n}{T_n + F_p} \times 100\% \quad (13)$$

$$F1 = \frac{2T_p}{2T_p + F_n + F_p} \times 100\% \quad (14)$$

式中, T_p 为正确分类的正例数目, F_n 为错误分类为负例的正例数目, T_n 为正确分类的负例数目, F_p 为错误分类为正例的负例数目。

2.3 实验结果及分析

2.3.1 分类性能对比实验

由于现有方法中未发现类似研究的方法,为了验证本文方法有效性,实验中将当前几种主流的分类网络应用于本文任务中,将面部提取模块与手部提取模块分别替换为其他分类网络,在其他条件不变的情况下进行训练并统计这些方法的性能指标。

图 5 显示了不同方法在准确性上的差异,表 2 显示了本文方法 DCFFM 与其他几种分类网络在准确性、敏感性、特异性和 F1-score 上的表现。其中,DCFFM 的准确性、灵敏性、特异性和 F1-score 分别为 83.36%、85.10%、82.47% 和 80.25%,在各指标上的表现均好于其他方法。其中敏感性显示出比较大的优势,说明本文方法对真阳性的检测有较好表现。

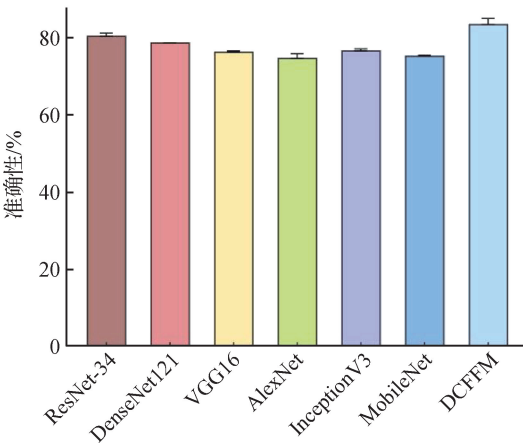


图 5 不同方法的准确性柱状图对比
Fig. 5 Comparison of histogram of accuracy among different methods

表 2 不同方法的各评价指标对比
Table 2 Comparison of results of each evaluation index among different methods

方法	准确性	敏感性	特异性	F1-score
ResNet-34	80.02 ± 2.35	83.49 ± 2.23	80.18 ± 5.34	78.36 ± 1.23
DenseNet121	78.63 ± 0.36	81.42 ± 3.46	78.11 ± 1.40	74.12 ± 0.49
VGG16	76.26 ± 1.24	79.81 ± 2.42	76.49 ± 3.32	73.94 ± 1.22
AlexNet	74.01 ± 2.15	76.09 ± 2.20	73.29 ± 2.08	72.33 ± 1.87
InceptionV3	76.28 ± 2.27	78.25 ± 3.31	75.38 ± 1.17	73.97 ± 4.11
MobileNet	75.37 ± 4.13	77.39 ± 3.67	74.33 ± 3.47	73.86 ± 2.92
本文 DCFFM	83.36 ± 2.11	85.10 ± 0.82	82.47 ± 4.16	80.25 ± 1.57

注:加粗字体表示各列最优结果。

为了进一步验证本文方法的稳定性和鲁棒性,对所有方法均在采集的数据集上进行 10 折交叉验证。图 6 为各方法敏感性和特异性指标对比。其中,图 6(a)为 10 折交叉验证下,不同方法 10 次 fold 的敏感性指标比较。DCFFM 方法相比其他方法在敏感性上有明显提升,其中 fold9 为 85.13%,相比较第二表现好的 ResNet-34 提升了 1.85%。除此之外,在 fold2、fold5、fold7 的敏感性都远高于其余方法,进一步说明本文方法在检测真阳性上的突出性能。图 6(b)为 10 折交叉验证下,不同方法 10 次 fold 的特异性指标比较。DCFFM 方法除了在 fold6 上比 ResNet-34 低 0.75%,在其他 fold 上均高于 ResNet-34 以及其他方法。除此之外,在 fold2 上 DCFFM 的特异性为 82.54%,明显优于其他方法。

整体分析可以证明,本文所提方法相比其他方法具有更好的稳定性和鲁棒性。

2.3.2 辅助信息分支对模型性能的影响

为了量化本文方法中辅助信息分支的有效性,去掉辅助信息分支,仅用主体分支作为面部和手部的特征提取网络并进行最终预测。实验结果如表 3 所示,去掉面部和手部的辅助信息分支后,准确性、敏感性、特异性和 F1-score 分别为 81.62%、82.96%、79.03% 和 78.47%。只有主体分支进行特征提取时,整体表现在各指标上都有明显下降,其中准确率和特异性都下降 2% 以上。实验结果表明,当加入辅助信息分支时,不同部位的针对性特征会更好辅助主体分支,充当一个“向导”去帮助主体分支学习更具有区别性的特征,从而提升模型性能。

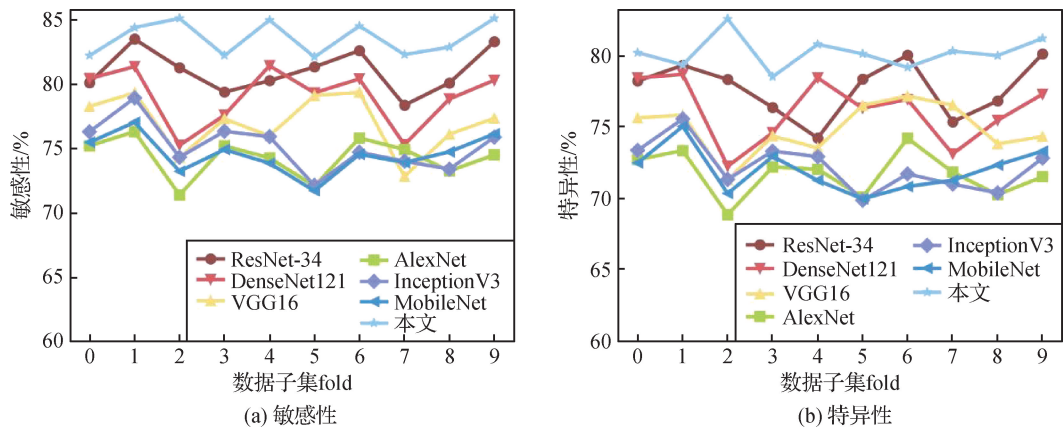


图6 不同方法在10折交叉验证下的性能比较

Fig. 6 Performance comparison of different methods under ten-fold cross-validation ((a) sensitivity; (b) specificity)

表3 辅助信息分支对模型性能的影响

Table 3 The impact of auxiliary information branch on model performance

方法	准确性	敏感性	特异性	F1-score
本文(无辅助分支)	81.02 ± 2.02	82.96 ± 1.33	79.03 ± 4.30	78.47 ± 1.12
本文	83.21 ± 2.41	84.65 ± 1.03	82.63 ± 3.16	80.10 ± 0.57

注:加粗字体表示各列最优结果。

由于面部特征提取分支与手部特征提取分支没有直接联系,因此对不同分支内的IIM中 λ_1 与 λ_2 的取值,分别选择不同取值单独进行实验。作为辅助信息分支中的权重,选取0.1,0.2,⋯,0.9共10个取值分别进行实验验证,结果如表4和表5所示。可以看出,当 λ_1 取值0.7、 λ_2 取值0.4时,模型的各项性能指标最好。

表4 IIM中 λ_1 不同取值对模型性能的影响

Table 4 Influence of different values of λ_1 in IIM on model performance

λ_1	准确性	敏感性	特异性	F-score
0.1	82.2	84.31	81.08	80.27
0.2	83.63	84.19	81.5	80.71
0.3	82.26	84.77	80.81	79.62
0.4	82.72	85.01	80.26	79.24
0.5	82.9	83.44	80.1	79.33
0.6	81.29	83.06	79.4	78.09
0.7	84.01	85.13	82.25	80.31
0.8	80.28	82.03	78.35	77.29
0.9	79.31	81.76	77.3	76.51

注:加粗字体表示各列最优结果。

表5 IIM中 λ_2 不同取值对模型性能的影响

Table 5 Influence of different values of λ_2 in IIM on model performance

λ_2	准确性	敏感性	特异性	F-score
0.1	81.28	83.11	79.74	79.1
0.2	83.1	84.52	82.48	80.15
0.3	83.05	84.83	82.37	80.17
0.4	83.36	85.1	82.47	80.25
0.5	83.3	85.05	81.05	80.77
0.6	82.1	83.42	80.11	79.08
0.7	82.72	83.8	81.49	80.2
0.8	81.33	83.06	79.53	77.18
0.9	80.38	82.19	78.72	76.08

注:加粗字体表示各列最优结果。

对于不同方法下的辅助信息分支,本文增加了相应的对比试验,验证了面部特征提取模块与手部特征提取模块中不同技术处理的辅助信息分支对于整体模型性能的影响。但值得注意的是,实验并未考虑在一个辅助信息分支中采用另一个辅助分支中

的方法,因为对于中医望诊在面部与手部的相关诊断知识,眉心部分的纹路特征以及掌心部位的形态特征是传统中医医师进行望诊的重点。因此在面部特征提取模块的眉心辅助分支中采用 Sobel 等滤波器进行实验而不采用转化颜色空间等方法,这样可以更好地关注于眉心的纹路特征,从而更好地辅助面部特征提取,实验结果如表 6 所示,可以看出,面部特征提取模块中采用 Sobel 滤波器作为辅助信息分支的处理方法是最优选择。对于手部特征提取模块,只转化不同颜色空间进行实验,因为掌心纹路复杂且手掌掌纹非常容易干扰模型提取有效特征,实验结果如表 7 所示,可以看出,手部特征提取模块中采用 HSV 颜色空间作为辅助信息分支的处理方法是最优选择。

表 6 面部特征提取模块中不同滤波器对模型性能的影响
Table 6 The influence of different filters in facial feature extraction module on model performance

/%				
滤波器	准确性	敏感性	特异性	F1-score
Sobel	83.12	85.32	81.77	80.93
Canny	82.33	85.01	81.62	80.1
Laplace	82.47	84.91	80.72	80.29

注:加粗字体表示各列最优结果。

表 8 信息交互模块(IIM)对模型性能的影响
Table 8 The impact of IIM on model performance

/%				
方法	准确性	敏感性	特异性	F1-score
本文(无 IIM)	82.41 ± 0.33	83.29 ± .61	79.96 ± 2.59	79.52 ± 2.11
本文	83.10 ± 1.38	85.02 ± 0.43	82.33 ± 1.26	80.41 ± 1.02

注:加粗字体表示各列最优结果。

2.3.4 面部图像与手部图像对于模型性能的影响
本文针对面部图像与手部图像双特征模块进行了消融实验对比,以此证明两个部位的同时输入可以实现相互促进,而不是相互限制。实验将面部特征提取模块与手部特征提取模块分别去掉,在仅保留一种部位的特征提取模块下进行实验,观察单部位特征提取模块与双部位特征提取模块的性能变化,实验结果如表 9 所示。当去掉手部特征提取模块仅依靠面部数据(DCFFM-face)预测脑卒中时,准确性、敏感性、特异性和 F1-score

表 7 手部特征提取模块中不同颜色空间
对模型性能的影响
Table 7 The effect of different color spaces in the hand feature extraction module on the performance of the model

/%				
颜色空间	准确性	敏感性	特异性	F1-score
HSV	83.41	86.02	82.51	81.03
YCbCr	82.3	84.82	81.27	80.11
Lab	78.33	81.7	77.39	76.24

注:加粗字体表示各列最优结果。

2.3.3 信息交互模块(IIM)对模型性能的影响
为了证明本文提出的信息交互模块(IIM)在 DCFFM 模型中的作用,实验验证了 DCFFM 分别在添加 IIM 与删除 IIM 时的整体性能表现。出于简化目的,实验在不改变其他分支的网络结构基础上,直接删除 IIM 后将辅助信息分支特征和主体分支特征进行拼接,实验结果如表 8 所示。在添加 IIM 后,整体性能依旧有一定提升,虽然准确性和 F1-score 指标提升较不明显,但在敏感性和特异性上均有 1% 的提升。这意味着 IIM 通过对辅助分支赋予一定的权重,可以使主体分支与辅助信息分支之间的特征信息以一种全新的方式进行进一步交互,进而帮助模型进行有效学习。

分别为 78.83%、81.28%、76.39% 和 77.04%。当去掉面部数据仅保留手部特征提取模块(DCFFM-hand)时,准确性、敏感性、特异性和 F1-score 分别为 80.63%、82.77%、79.04% 和 80.11%。当两个数据同时输入时,整体性能在准确率和敏感性上都有明显提升。通过实验也进一步说明仅依靠一种数据很难达到很好的效果,而同时将面部数据和手部数据进行分析并进行有效融合,通过训练不断进行信息交互能够有效提升模型性能。

表 9 单部位数据与多部位数据对模型性能的影响

Table 9 The influence of single-part data and multi-part data on model performance

方法	准确性	敏感性	特异性	F1-score
DCFFM-face	78.83	81.28	76.39	77.04
DCFFM-hand	80.63	82.77	79.04	80.11
本文	83.02	84.71	81.29	80.27

注:加粗字体表示各列最优结果。

3 结 论

鉴于对脑卒中的病前预测缺少有效方法,本文依据中医望诊在“治未病”方面的显著作用,结合中医望诊在脑卒中领域的相关诊疗知识,基于面部与手部图像数据提出一种多分支深度特征融合的中医脑卒中辅助诊断方法。构建面部与手部特征提取模块,并在各模块中采取双分支的结构用于提取不同部位的深度特征;将面部和手部的重点区域作为主体分支以提取主要特征信息,并在中医望诊知识指导下将眉心和掌心的不同特征作为辅助信息以进一步提取有效信息;构建信息交互模块(IIM)将主要特征与辅助特征进行有效信息交互;将两部位特征进行融合降维以实现脑卒中预测。多组实验证明,本文方法能够有效利用面部和手部数据进行脑卒中的辅助预测,提出的信息交互模块在一定程度上也促进了多分支任务的信息交互。本文方法较 Res-Net-34、DenseNet121、VGG16(Visual Geometry Group 16-layer net)和 InceptionV3 等主流分类算法在准确性、敏感性、特异性和 F1-score 指标上均有一定优势,能够有效辅助中医医师快捷准确地进行脑卒中诊断。

本文方法虽取得一定效果,但在整体准确性的提升以及如何引入更多符合中医诊疗知识的辅助信息等方面仍需进一步研究。基于这些问题,后续工作将集中于以下两点开展:1)结合更多部位的图像数据进行模型搭建,进一步提高准确率;2)考虑到实际应用需求,对网络模型进行进一步修改和优化。

参考文献 (References)

Cheng H D, Jiang X H, Sun Y and Wang J L. 2001. Color image seg-

mentation; advances and prospects. Pattern Recognition, 34(12): 2259-2281 [DOI: 10.1016/S0031-3203(00)00149-7]

Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau N G, Venugopal V K, Mahajan V, Rao P and Warier P. 2018. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. Lancet, 392(10162): 2388-2396 [DOI: 10.1016/S0140-6736(18)31645-3]

Cho J, Park K S, Karki M, Lee E, Ko S, Kim J K, Lee D, Choe J, Son J, Kim M, Lee S, Lee J, Yoon C and Park S. 2019. Improving sensitivity on identification and delineation of intracranial hemorrhage lesion using cascaded deep learning models. Journal of Digital Imaging, 32(3): 450-461 [DOI: 10.1007/s10278-018-00172-1]

Cipolla R, Gal Y and Kendall A. 2018. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 7482-7491 [DOI: 10.1109/CVPR.2018.00781]

Feigin V L, Krishnamurthi R V, Parmar P, Norrving B, Mensah G A, Bennett D A, Barker-Collo S, Moran A E, Sacco R L, Truelsen T, Davis S, Pandian J D, Naghavi M, Forouzanfar M H, Nguyen G, Johnson C O, Vos T, Meretoja A, Murray C J L, Roth G A, GBD 2013 Writing Group and GBD 2013 Stroke Panel Experts Group. 2015. Update on the global burden of ischemic and hemorrhagic stroke in 1990-2013: the GBD 2013 study. Neuroepidemiology, 45(3): 161-176 [DOI: 10.1159/000441085]

Han X J, Gao Y R, Liu D S, Yu X and Ren C. 2018. Research on the theory and mechanism of three-dimensional observation diagnosis of coronary heart disease. Chinese Journal of Basic Medicine in Traditional Chinese Medicine, 24(4): 441-442, 456 (韩学杰, 郜亚茹, 刘大胜, 于潇, 任聪. 2018. 冠心病大鱼际三维望诊法的理论与机制探究. 中国中医基础医学杂志, 24(4): 441-442, 456)

He K M, Zhang X Y, Ren S Q and Jian S. 2016. Identity mappings in deep residual networks//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherlands: Springer: 630-645 [DOI: 10.1007/978-3-319-46493-0_38]

Ioffe S and Szegedy C. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift//Proceedings of the 32nd International Conference on International Conference on Machine Learning. Lille, France: ICML: 448-456

Kim B H, Lee S H, Cho D U and Oh S Y. 2008. A proposal of heart diseases diagnosis method using analysis of face color//Proceedings of 2008 International Conference on Advanced Language Processing and Web Information Technology. Dalian, China: IEEE: 220-225 [DOI: 10.1109/ALPIT.2008.27]

Kuo W C, Häne C, Mukherjee P, Malik J and Yuh E L. 2019. Expert-level detection of acute intracranial hemorrhage on head computed tomography using deep learning. Proceedings of the National Academy of Sciences of the United States of America, 116(45):

- 2737-22745 [DOI: 10.1073/pnas.1908021116]
- Li B C, Huang Q M, Lu Y, Chen S H, Liang R and Wang Z P. 2007. A method of classifying tongue colors for traditional Chinese medicine diagnosis based on the CIELAB color space//Proceedings of the 1st International Conference on Medical Biometrics. Hong Kong, China: Springer; 153-159 [DOI: 10.1007/978-3-540-77413-6_20]
- Li J X, Zhang B, Lu G M, You J and Zhang D. 2019. Body surface feature-based multi-modal learning for diabetes mellitus detection. Information Sciences. 472: 1-14 [DOI: 10.1016/j.ins.2018.09.010].
- Lin M, Chen Q and Yan S C. 2013. Network in network [EB/OL]. [2021-08-10]. <http://arxiv.org/pdf/1312.4400v3.pdf>
- Lin S, Li Z G, Fu B W, Chen S P, Li X, Wang Y, Wang X Y, Lv B, Xu B, Song X T, Zhang Y J, Cheng X, Huang W J, Pu J, Zhang Q, Xia Y L, Du B, Ji X Y and Zheng Z. 2020. Feasibility of using deep learning to detect coronary artery disease based on facial photo. European Heart Journal, 41(46): 4400-4411 [DOI: 10.1093/eurheartj/ehaa640]
- Lu X G, Ma K W, Wang Y Q and Hao Y M. 2019. Study on tongue color features of healthy people with "Joy" emotion based on objectification of tongue diagnosis. World Science and Technology-Modernization of Traditional Chinese Medicine, 21(3): 471-476 (鲁晓箐, 马凯雯, 王忆勤, 郝一鸣. 2019. 基于舌诊客观化的情志“喜”舌色特征研究. 世界科学技术-中医药现代化, 21(3): 471-476) [DOI: 10.11842/wst.2019.03.022]
- Sobel I and Feldman G. 1973. A 3×3 isotropic gradient operator for image processing. Pattern Classification and Scene Analysis, 271-272
- Su W, Xu Z Y, Wang Z Q and Xu J T. 2011. Objectified study on tongue images of patients with lung cancer of different syndromes. Chinese Journal of Integrative Medicine, 17(4): 272-276 [DOI: 10.1007/s11655-011-0702-6]
- Wang D M, Zhang D and Lu G M. 2017a. Generalized feature extraction for wrist pulse analysis: from 1-D time series to 2-D matrix. IEEE Journal of Biomedical and Health Informatics, 21(4): 978-985 [DOI: 10.1109/JBHI.2016.2628238]
- Wang F, Jiang M Q, Qian C, Yang S, Li C, Zhang H G, Wang X G and Tang X O. 2017b. Residual attention network for image classification//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE; 6450-6458 [DOI: 10.1109/CVPR.2017.683]
- Wang P, Zuo W M and Zhang D. 2014. A compound pressure signal acquisition system for multichannel wrist pulse signal analysis. IEEE Transactions on Instrumentation and Measurement, 63(6): 1556-1565 [DOI: 10.1109/TIM.2013.2267458]
- Yan K, Zhang D, Wu D R, Wei H and Lu G M. 2014. Design of a breath analysis system for diabetes screening and blood glucose level prediction. IEEE Transactions on Biomedical Engineering, 61(11): 2787-2795 [DOI: 10.1109/TBME.2014.2329753]
- Zhang D, Zhang H and Zhang B. 2017. Tongue Shape Classification by Geometric Features. Tongue Image Analysis. Springer, Singapore. 133-153. [DOI: 10.1007/978-981-10-2167-1_8]
- Zhang L, Samaras D, Tomasi D, Volkow N and Goldstein R. 2005. Machine learning for clinical diagnosis from functional magnetic resonance imaging//Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE; 1211-1217 [DOI: 10.1109/CVPR.2005.219]
- Zhang Q, Zhou J H, Zhang B and Wu E H. 2021. DsNet: dual stack network for detecting diabetes mellitus and chronic kidney disease. Information Sciences, 547: 945-962 [DOI: 10.1016/j.ins.2020.08.074]
- Zhou M G, Wang H D, Zeng X Y, Yin P, Zhu J, Chen W Q, Li X H, Wang L J, Wang L M, Liu Y N, Liu J M, Zhang M, Qi J L, Yu S C, Afshin A, Gakidou E, Glenn S, Krish V S, Miller-Petrie M K, Mountjoy-Venning W C, Mullany E C, Redford S B, Liu H Y, Naghavi M, Hay S I, Wang L H, Murray C J L and Liang X F. 2019. Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the global burden of disease study 2017. Lancet, 394(10204): 1145-1158 [DOI: 10.1016/S0140-6736(19)30427-1]

作者简介



王麒达, 1997年生, 男, 硕士研究生, 主要研究方向为医学图像处理和深度学习。

E-mail: wangqida0375@link.tyut.edu.cn



强彦, 通信作者, 男, 教授, 主要研究方向为人工智能、图像处理、云计算技术和计算机视觉。

E-mail: qiangyan@tyut.edu.cn

冀伦文, 男, 博士研究生, 主要研究方向为机器学习和计算机视觉。E-mail: jilunwen@tyut.edu.cn

王华虎, 男, 硕士研究生, 主要研究方向为中医学、人工智能和医学图像处理。E-mail: wanghuahu@stu.pku.edu.cn

赵琛琦, 男, 硕士研究生, 主要研究方向为医学图像处理、自然语言处理与深度学习。

E-mail: zhaochenqi0097@link.tyut.edu.cn

李慧芝, 女, 本科生, 主要研究方向为人工智能和深度学习。E-mail: 42585479@qq.com

赵紫娟, 女, 博士, 主要研究方向为中医大数据分析和深度学习。E-mail: zhaoziju0064@link.tyut.edu.cn