



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

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 4<sup>th</sup> 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

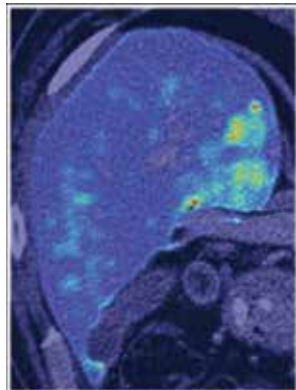
**Distributed by** Beijing Bureau for Distribution of Newspapers and Journals  
**Domestic** All Local Post Offices in China  
**Overseas** China International Book Trading Corporation  
(P.O.Box 399, Beijing 100048, P.R.China)  
**Printed by** Beijing Kexin Printing Co., Ltd.

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

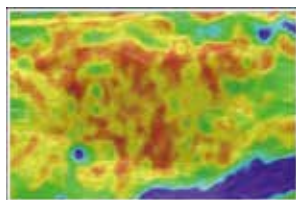
国外发行代号 M1406  
国内邮发代号 82-831  
国内定价 60.00元



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

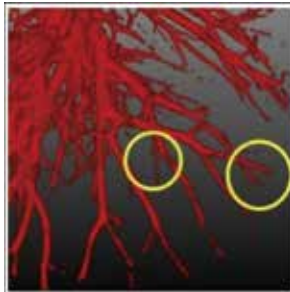
基于图像的介入器械前端变形建模与受力估计

任龙飞, 孟偲, 刘博, 白相志, 周付根 ..... 1001

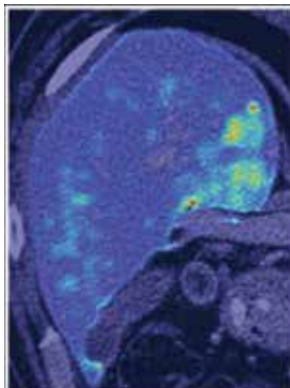


# CONTENTS

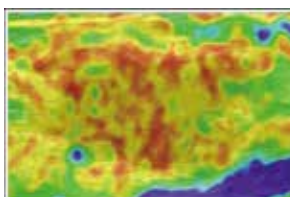
## JOURNAL OF IMAGE AND GRAPHICS



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)03-0923-12

论文引用格式: Zhao C Q, Wang H H, Zhao J J, Ji L W, Wang Q D, Li H Z and Zhao Z J. 2022. Cerebral stroke detection algorithm for visual Transformer and multi-feature fusion. Journal of Image and Graphics, 27(03): 0923-0934 (赵琛琦, 王华虎, 赵涓涓, 冀伦文, 王麒达, 李慧芝, 赵紫娟. 2022. 视觉 Transformer 与多特征融合的脑卒中检测算法. 中国图象图形学报, 27(03): 0923-0934) [DOI:10.11834/jig.210745]

# 视觉 Transformer 与多特征融合的脑卒中检测算法

赵琛琦<sup>1</sup>, 王华虎<sup>2</sup>, 赵涓涓<sup>1\*</sup>, 冀伦文<sup>3</sup>, 王麒达<sup>1</sup>, 李慧芝<sup>4</sup>, 赵紫娟<sup>1</sup>

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

**摘要:** **目的** 急性缺血性卒中是最常见的脑卒中类型, 具有发病率高、死亡率高和致残率高的特点。患者发病前症状不明显、发病急骤以及溶栓治疗时间窗窄等问题导致其成为临床上的高危疾病。中医望诊可以在疾病发展早期, 通过观察患者形、色、气和神的变化, 对患者病情进行诊断和预测, 达到“治未病”的目的, 与人工智能技术相结合, 可以解决缺乏客观和定量评价标准的问题。因此, 通过中医望诊中的脸部和手部图像, 充分利用两种图像的颜色、纹理等特征以及二者之间的关系特征, 本文提出一种基于序列自注意力网络的急性缺血性卒中辅助诊断方法。**方法** 对脸部和手部图像进行山根和大鱼际处的感兴趣区域提取。采用 YCbCr 颜色空间和灰度共生矩阵, 提取区域图像的颜色和纹理特征, 将颜色特征和纹理特征进行融合并将其与原图像特征相结合, 得到的特征图序列化地输入到 Transformer 模型中, 进一步学习高层次的空间特征和注意力特征。将模型输出结果输入到多层感知机中, 从而实现急性缺血性卒中的检测。**结果** 在收集的急性缺血性卒中患者数据集上进行实验, 结果表明, 提出的基于序列自注意力网络的方法取得了 83.57% 的准确率, 获得较高性能, 在速度和便携性上具有很大的优势。**结论** 该方法采用端到端的学习方式, 能够有效解决目前临床诊断因医疗资源的差异而受到影响的问题, 对于初步判断患者疾病具有指导性的作用, 为诊断急性缺血性卒中提供了一种新思路和新方法。

**关键词:** 中医望诊; 特征提取; 特征融合; 端到端; Transformer

## Cerebral stroke detection algorithm for visual Transformer and multi-feature fusion

Zhao Chenqi<sup>1</sup>, Wang Huahu<sup>2</sup>, Zhao Juanjuan<sup>1\*</sup>, Ji Lunwen<sup>3</sup>, Wang Qida<sup>1</sup>, Li Huizhi<sup>4</sup>, Zhao Zijuan<sup>1</sup>

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

**Abstract:** **Objective** Cerebral ischemic stroke is the most common type of cerebral stroke, which is characterized by high morbidity, mortality and disability. The lack of obvious symptoms before the onset of the disease, the rapid onset of the disease, and the narrow time window for thrombolytic therapy have led to it being a high-risk disease in clinical practice. Although initial progress has been made in cerebral stroke prevention and treatment, it remains a significant cause of disa-

收稿日期: 2021-08-20; 修回日期: 2021-11-11; 预印本日期: 2021-11-18

\* 通信作者: 赵涓涓 zhaopian@tyut.edu.cn

**基金项目:** 国家自然科学基金项目(61972274); 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(61972274); Key Research and Development Project of Shanxi Province, China (201903D321034)

bility or death in adults. According to the survey, approximately 75% of stroke patients have varying degrees of functional impairment and loss of work, causing a heavy burden on families and society. With the accelerated aging and urbanization of society, the prevalence of unhealthy lifestyles among the population and the widespread exposure to cerebrovascular disease risk factors, in the disease burden of stroke has greatly increased, with a trend of rapid growth in low-income groups, marked gender and geographical differences and youthfulness. Therefore, effective ways to reduce disability and mortality rates should be developed. The early diagnosis of cerebral stroke is important. Many methods can be used to diagnose cerebral stroke in modern medicine, but the processes are relatively complex. In addition, some tests have certain drawbacks, and the presence of the disease is hard to detect in the early stages of illness, thus requiring advanced equipment and experienced clinicians. How to improve the accuracy of early diagnosis of cerebral stroke has become an important research hotspot for medical aid diagnosis. The characteristics and advantages of traditional Chinese medicine (TCM) are essential in the contemporary medical system of diseases, especially the inspection diagnosis of TCM, which is the most important in TCM diagnosis. Chinese medicine diagnosis is an objective and accurate empirical medicine, which has gradually formed and developed in long-term medical practice and clinically proven, with extremely rich connotations. Based on the basic principles of Chinese medicine diagnosis (the inspection diagnosis of TCM), and diagnosis can be improved by applying modern scientific knowledge and methods in practice. This method not only provides strong evidence for early diagnosis and treatment, but also has extremely important practical significance in saving medical resources, reducing the medical burden on patients and alleviating the harm caused by cerebral stroke disease.

**Method** First, feature extraction is performed on the images of the patient's face and hands. The color features are easily affected by light, and the chroma component in YCbCr color space is used to reduce the effect of luminance. The most important of the texture features are the features of texture length, depth and thickness in the images, and the gray level co-generation matrix (GLCM) was used to extract the image texture features effectively. Then, the higher-order spatial dimensional features further learned from the original image and the attentional features are learned from the different features by designing a reasonable dual Transformer joint classification model. Different transformer modules were cascaded, and multi-layer perception was used for image classification. This method not only considers color and texture features in the image, but also analyzes the spatial features of the image. Based on the differences arising from successive changes in color and texture between different regions in an image, this paper uses transformer to extract the attention features between different regions to improve the performance of the diagnostic model. In addition, the detection model is trained end-to-end. During the training process, the batch size is set to 4, the learning rate is set to  $1E-5$  and the maximum number of cycles is set to 100. The experiment uses NVIDIA TITAN XP GPU, and the data set was divided into five groups equally for five cross-validations. Finally, the average accuracy of all cross-validated results was taken as the final result of the experiment.

**Result** When detecting cerebral ischemic stroke, the models with color features (YCbCr) and texture features (GLCM) extracted separately achieved accuracies of 79.40% and 80.46% on the dataset, while the model with the fusion of color and texture features achieved an accuracy of 83.53% on the dataset, which was significantly better than the model without feature fusion. Color features and texture features can effectively improve the classification accuracy in classification by using a transformer model, and feature fusion can make the model further improve the detection accuracy. Under the premise of fusion of color and texture features, the accuracy of model classification using a transformer module has dropped by approximately 2%. This finding shows that features from different parts play different roles in the final detection, and the gaps between the same features from different parts can easily disappear in the process of feature fusion into one transformer module. The dual transformer joint classification model uses color, texture, spatial and attention features, and the combination of these features can effectively improve the performance of the model. In addition, the average accuracy of the proposed model on the dataset in this paper outperforms the experimental results of related classification models.

**Conclusion** In this paper, we proposed an end-to-end joint classification detection method based on the dual Transformer module. High-quality data were acquired using YCbCr color space and GLCM to accelerate the convergence process of the model. In addition, we extracted feature information from the patient's face and hand images. More importantly, the model learning capability was enhanced, and the model performance was improved using a self-attentive mechanism to learn the association between features and assign weights. The proposed model has a good diagnostic effect, and the automatic assisted diagnosis reduced the influence of subjective factors, which is valuable in

the study of cerebral ischemic stroke auxiliary diagnosis, provides a reference for clinicians to make decisions on cerebral ischemic stroke disease diagnosis and provides a new method for patients to conduct effective self-screening.

**Key words:** inspection diagnosis of traditional Chinese medicine; feature extraction; feature fusion; end-to-end; Transformer

## 0 引言

在全球致死疾病排名中,脑卒中(脑中风)已成为第二大致死因素,严重影响人们的身体健康。急性缺血性卒中(cerebral ischemic stroke, CIS)(Boling 和 Keinath, 2018)是最常见的卒中疾病类型,约占卒中中 2/3 (Strong 等, 2007)。尽管脑血管疾病始终是中医药发展战略中的首要议题,但关于中医药治疗 CIS 的流行病学资料并不多,设计良好的临床中医药治疗急性缺血性卒中的研究也很少。急性缺血性卒中对人类健康造成了严重的威胁,在全球研究人群中约 17% 的人深受其困扰 (Seshadri 和 Wolf, 2007)。急性缺血性卒中,会消耗大量的医疗资源、经济资源,对医疗体系产生极大的负担 (Virani 等, 2021)。随着人口老龄化的加剧,全国范围内患病人数增加,患病造成的死亡或残疾会对个人、家庭以及社会产生不良的影响。临床研究表明,发病时间分为急性、亚急性和慢性 3 种情况,急性发病时间短,如果患者能够在发病时间内得到准确的诊断和治疗,那么急性缺血性卒中所带来的影响可以得到有效改善,这不仅有利于患者的恢复,而且会大大降低疾病带来的并发症的影响 (Hsieh 等, 2021)。目前,脑中风临床的主要诊断手段是磁共振成像 (magnetic resonance imaging, MRI) 和计算机断层扫描 (computed tomography, CT), 病变区域中 CT 和 MRI 图像上有较为明显的区别,通过对较暗区域和较亮区域进行判断得出结论。尽管 MRI 和 CT 是临床医生常用的诊断手段,但急性缺血性卒中患者的诊断往往需要经验丰富的临床医生完成,不同医院医疗资源的差异将会影响诊断结果,漏诊和误诊率也会增加。为了解决上述问题,必须为急性缺血性卒中提供辅助诊断方法,降低漏诊和误诊率,并为急性缺血性卒中的早期发现提供新方法和新思路。西方医学未满足脑血管疾病控制和诊断的需求,临床医生开始考虑使用中医 (traditional Chinese medicine, TCM) 对脑血管疾病进行预防与诊断,该领域

的一些基础和临床研究越来越受到脑血管学界的关注 (Zhang 等, 2013; 高佳明 等, 2019)。缺乏客观和定量评价标准的中医在大多数西方国家被认为是复杂的、补充性的代替医学 (Hao 等, 2015), 然而在中国的西医、中医和中西医结合 3 种检查方式中,超过 71.2% 的患者更倾向于中西医结合,这为中医成为辅助诊断的方式提供了群体基础。中医望诊对急性缺血性卒中的早期诊断有重要的临床意义,《难经》中“望而知之谓之神”的思想,足以可见望诊在中医诊断中的重要地位 (张超 等, 2020)。望诊是医生通过观察患者面色、形态、局部情况和舌头等的变化来判断患者病情。人体外部表征与身体内部变化极其相关 (高利 等, 2007)。面部望诊和手部望诊对初步判断患者疾病具有指导性的作用。《灵枢》将面部划分为不同区域,不同的区域对应着不同的人体部位,这基本组成了人体结构的全局信息图。因此可以从面象的变化来判断脏腑的病症和疾病的性质,了解气血类疾病的情况 (Zhang 等, 2014)。大鱼际诊断法为《黄帝内经》中最早记载的关于手部作为局部望诊的方法,可用于急性缺血性卒中的诊断手段,因为它更为便捷、迅速、经济且易于操作,所以有利于患者早期自我检查以及临床早期筛查工作,便于对高危人群进行早发现、早诊断和早治疗 (郜亚茹 等, 2020)。中医望诊诊断脑血管疾病时通常是在自然光线下,通过观察患者的山根 (鼻根处) 褶皱以及面色,患者手掌外侧部位的隆起部位的皮肤颜色、皮肤纹理和肌肉形态等进行诊断。《运气学导论》中记载着很多通过观察患者面色并且结合脉象来判断患者发生脑血管意外可能性的案例。因此,中医望诊在脑血管疾病的诊断过程中发挥着重要作用,有利于早期脑血管疾病的筛查以及症状的鉴别,避免临床中误诊和漏诊情况的发生。

基于中医望诊的可能性,国内外研究者对智能中医展开了大量研究, Lin (2020) 根据中医面部的一些视觉特征可以为疾病的诊断提供有价值的线索,提出一种基于卷积神经网络 (convolutional neural network, CNN) 的面向客观中医面部诊断的面部肤



色自动分类模型。针对急性缺血性卒中的特点,建立回归分析模型分析中医脑血管疾病并发病状(刘泽文,2015)。部分研究者(席恋,2016;罗晓舟等,2017)使用机器学习方法对大量急性缺血性卒中患者的临床资料进行分析,同时结合中医相关知识,证明了中医智能诊断在中风诊断领域中的重要地位。研究者通过使用机器学习的模型,对患者数据进行线性学习,但中医诊断是从大脑进行的,具有一定的层次性、复杂性和非线性。大多数机器学习算法没有从大脑的认知角度考虑诊断的层次性,而疾病和综合征之间的关系不是一对一的,而是多对多的,这一事实使得机器进行综合征预测变得非常困难。深度神经网络在理论上比浅层模型具有更高的表征能力。以缺血性脑卒中疾病数据为研究内容,结合多种特征提出长短期记忆网络(long short-term memory, LSTM)模型,为缺血性脑卒中的临床辅助诊断提供一种方法,从而为医师诊断提供了参考(骆铁姝等,2020)。结合注意力机制的LSTM模型对比不使用注意力机制的模型,注意力机制方法在不同数据集上均能显著提升识别性能(余晨阳等,2019)。Karthik和Menaka(2018)从医学图像领域出发,利用端到端的自动框架识别脑卒中的损伤区域。颜色特征和纹理特征融合的方式可有效解决运动目标识别的问题(武明虎等,2017),还可以减少不同光照对颜色特征的影响。

尽管深度学习网络取得了可喜的成绩,但是深

度学习与中医领域相结合的研究还较少,如何基于简单的手部和脸部图像,将中医望诊与人工智能技术相结合成为一个重要的研究点。针对以上问题,通过建立基于脸部和手掌图像的深度模型,能够较好地辅助中医进行脑血管疾病的预测。尽管很多研究者对脑血管疾病已经进行了相关的研究工作,但是基于中医望诊结合深度学习来进行脑血管疾病的相关研究很少。

本文的主要贡献有:

1) 提出一种以中医望诊为基础的序列自注意力网络模型辅助诊断脑卒中的方法。该方法以中医理论为基础,将中医诊断中的面部和手部望诊转化为颜色特征和纹理特征进行脑卒中的智能检测。

2) 将脸部(山根)和手部(大鱼际)感兴趣区域的颜色和纹理特征进行融合,结合原图像序列化地输入注意力模型中,从而学习出不同特征图像对模型性能的差异性,以便模型提取到更有判别能力的注意力特征,再将两个Transformer模块得到的向量使用多层感知机进行分类,使模型更有针对性地进行训练。

3) 在收集的脑卒中患者数据集上的实验结果表明,本文模型可以取得较好的准确率。

## 1 本文方法

图1给出了本文端到端的Transformer的中医望诊图像诊断急性缺血性卒中的模型,总体方案流

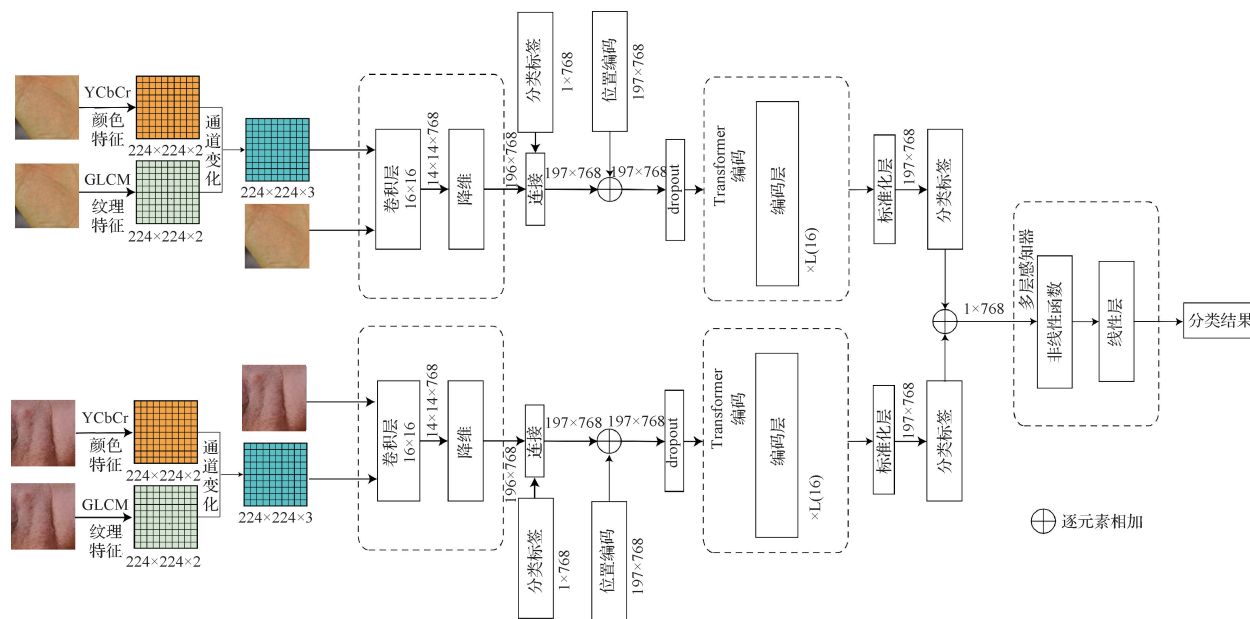


图1 中医望诊图像诊断急性缺血性卒中的模型示意图

Fig. 1 Schematic diagram of a model for cerebral stroke with inspection diagnosis of traditional Chinese medicine (TCM) images



程如下:

1) 肤色特征提取。在实际的肤色检测中, 肤色会受到采集设备和光照条件的影响, 采用 YCbCr 颜色空间可以将亮度信息和颜色信息完全分类, 使肤色特征不受光线的影响。

2) 皮肤纹理特征提取。皮肤纹理特征不仅可以作为肤色的辅助手段, 而且也是预测脑血管疾病不可缺少的特征。对脸部山根区域和手部大鱼际区域的图像, 利用灰度共生矩阵 (gray level co-generation matrix, GLCM) 提取纹理特征。

3) 图像特征融合。将提取的 YCbCr 颜色空间的特征和纹理特征的角二矩阵、对比度、逆差分矩和熵特征进行连接, 达到利用不同尺度特征图的目的。

4) 图像特征分类。面部和手部图像之间的联系也对辅助诊断起到一定的作用, 采用 Transformer 模型可以进行序列化的输入, 输出图像的分类结果。

综上所述, 本文方法依照中医望诊理论, 考虑肤色和纹理特征, 并将不同位置的特征通过序列 Transformer 模型提取区域注意力特征, 从而提高模型分类准确性。

### 1.1 肤色特征提取

中医望诊中将面色大致分为: 青、红、黄、白、黑, 传统的中医望诊中由医师观察患者的面色, 对患者进行诊断。因此, 诊断结果往往受到医师医术水平的限制, 并且容易受到光线的影响, 缺少客观化的标准。对人脸面色使用计算机技术进行定性和定量的分析, 可以弥补中医的不足。研究发现影响肤色判断的主要因素是亮度, 消除亮度影响之后的肤色信息, 肤色特征将更为准确。色彩空间的选择将直接影响特征提取的结果, 为了减少亮度的影响, 使用线性 YCbCr 颜色空间模型。

YCbCr 是一种编码的非线性 RGB 信号, 通常被欧洲电视演播室用于图像压缩工作。颜色由亮度 (由非线性 RGB 计算出的亮度) 表示, 由 RGB 颜色空间进行加权转化而来。YCbCr 是数字视频领域常用的彩色空间, 这种表示方式可以很容易地去除一些多余的颜色信息, YCbCr 中的亮度和色度分量可以分离显示。在 YCbCr 颜色空间中有两个元素, 亮度信息存储为单个分量 (Y), 色度信息存储为两个色度分量 (Cb 和 Cr)。Cb 表示蓝色分量与参考值之间的差值。Cr 表示红色分量与参考值之间的差值。YCbCr 值可以从 RGB 色彩空间中转化得到

(Shaik 等, 2015), 计算为

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.279 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

### 1.2 皮肤纹理特征提取

中医望诊中对脑血管疾病的观察主要集中于脸部的山根处, 皱褶长度、条数与疾病程度有很大关联。纹理特征的提取成为脑血管疾病辅助诊断系统的关键步骤, 不同于颜色和亮度等特征, 纹理特征反映的是物体本身的视觉特征, 并且不易受到环境的干扰。由于纹理特征具有多样性和复杂性, 采用灰度共生矩阵的方法对图像灰度信息进行不同灰度信息频次的统计。可以较为精准地预测和反映山根处灰度图像中的方向、变化幅度和相邻间隔等多种特征。

纹理是空间上的灰度频繁出现而形成的, 在一定距离内两个像素之间会存在灰度关系, 灰度共生矩阵便是分析纹理之间存在的灰度关系。灰度共生矩阵可以从方向、相邻间隔和变化幅度等综合信息中分析图像的局部模式以及它们的排列规则灰度。灰度共生矩阵是在一幅图像中存在两个灰度为  $i$  和  $j$  的像素点, 像素点之间的距离为  $d$ , 两个灰度的像素点在图像区域内同时发生的概率。灰度共生矩阵用  $P_d(i, j; \theta)$  ( $i, j = 0, 1, \dots, L-1$ ) 表示, 其中,  $i, j$  表示两像素灰度;  $L$  为图像的灰度级数;  $d$  是像素之间的位置关系, 像素在  $X$  方向和  $Y$  方向上的距离分别为  $\Delta x$  和  $\Delta y$ , 表示为  $\delta(\Delta x, \Delta y)$ 。根据定义, 两个像素点在  $0^\circ, 45^\circ, 90^\circ, 135^\circ$  这 4 个方向上位置关系如图 2 所示。在对灰度共生矩阵的角二阶矩、逆差分矩、对比度以及熵进行特征提取的过程中主要是对纵向纹理特征进行提取, 本文中只采用一个方向位置关系进行特征提取。

1) 角二矩阵。反映图像区域内灰度分布的均匀性, 能量大小只反映纹理粗细程度, 能量值越大纹理越粗。计算为

$$C_1 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (P_d(i, j; \theta))^2 \quad (2)$$

2) 对比度。反映图像的清晰度, 对比度越大, 纹理深度越深, 图像越清晰; 反之纹理深度越浅, 图像越模糊。计算为

$$C_2 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i-j)^2 P_d(i, j; \theta) \quad (3)$$

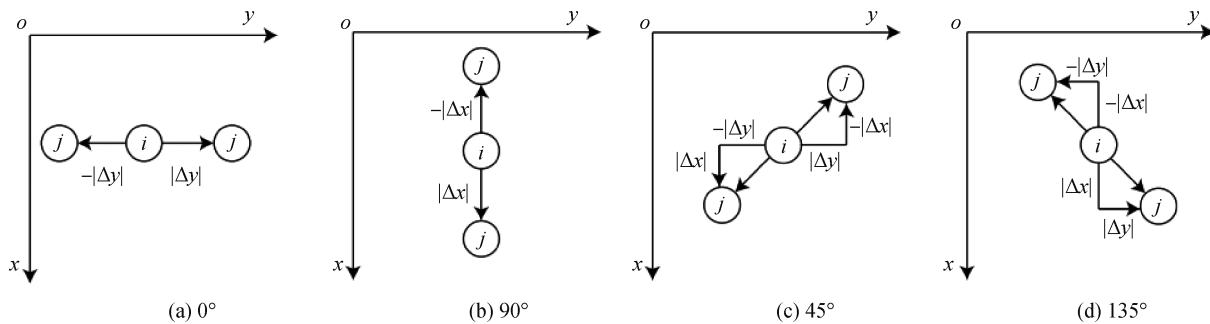


图2 4种常用方向上的位置关系

Fig. 2 Position relationships in four common directions((a)0°; (b)90°; (c)45°; (d)135°)

3) 逆差分矩。反映图像分布平滑度的度量,式中  $(i-j)^2$  越小,表示纹理越细;反之表示纹理越粗。计算为

$$C_3 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \left[ \frac{P_d(i, j; \theta)}{1 + (i - j)^2} \right] \quad (4)$$

4) 熵。表示图像内容的随机性,主要反映图像的信息量和复杂度。熵值越大,纹理越复杂;反之纹理越简单。计算为

$$C_4 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P_d(i, j; \theta) \lg P_d(i, j; \theta) \quad (5)$$

在进行灰度共生矩阵处理图像边界信息时,根据处理的窗口大小和步长大小,扩展源图像的边界,使源图像扩大维数,并且将添加的边界像素值也参加计算,最终使原来的边界元素也能成为中心元素

进行计算,使最终处理结果仍保持原始图像尺寸。

通过灰度共生矩阵对脸部和手部图像进行纹理特征提取,能够使得计算机“认识”图像并且更容易对图像进行定量分析。

### 1.3 颜色特征和纹理特征融合

皮肤颜色容易受到光照影响,在采用的 YCbCr 颜色空间中, Y 分量表示亮度, CbCr 表示色度,这两个分量已经可以表达面色。为了消除亮度分量变化带来的影响,颜色特征选择色度的两个分量作为主要颜色特征。纹理特征采用灰度共生矩阵提取 4 个特征向量进行特征表示,为了使得颜色特征和纹理特征较为方便地融合,纹理特征的提取采用滑动窗口的方式使其与颜色特征维度相一致。纹理特征提取过程如图 3 所示。

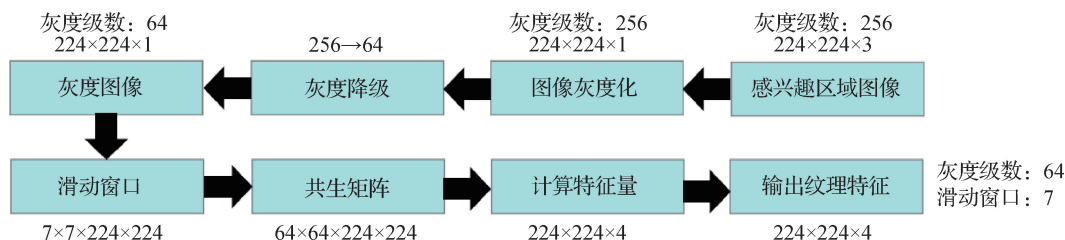


图3 目标区域纹理特征提取

Fig. 3 Target area texture feature extraction

### 1.4 模块设计

面部图像和手部图像的分类都采用 Transformer 模块,模型主要由 3 个部分组成(如图 4),分别为嵌入层、编码层和多层感知器分类层。其中,  $P_i$  表示位置编码,  $E_i$  表示融合特征编码,  $t_i$  表示图像线性投影编码,  $T_i$  表示 Transformer 模块的输出结果, CLS 为分类向量。

在标准的 Transformer 模块中,输入的是 2 维矩阵,而图像数据是 3 维矩阵,并不符合标准。因此进行

嵌入时需要对图像数据进行变化。将输入模型的图像 ( $224 \times 224$  像素)按照  $16 \times 16$  的大小进行划分得到 196 块。此时划分之后的数据为  $[16, 16, 3]$ ,通过映射可以得到长度为 768 的向量,从而满足 Transformer 模块输入要求。在处理得到的向量中加入专门用于分类的可训练参数,与之前划分完的块拼接在一起得到  $[197, 768]$  的 2 维矩阵。位置编码采用随机初始化通过训练一个位置编码参数来学习记录图片的位置信息。

在编码层中首先研究标准的自注意力机制的计

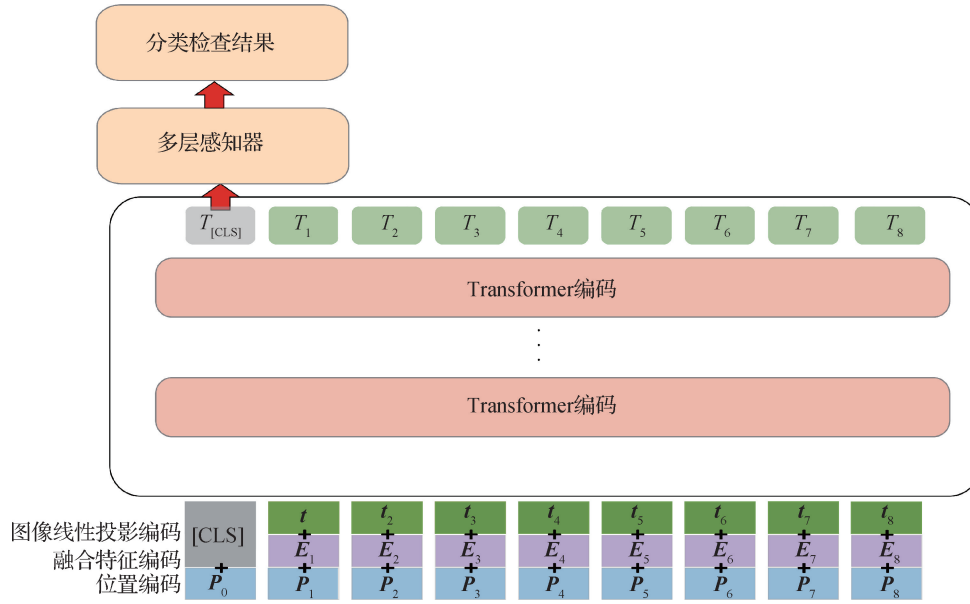


图4 模型基本模块

Fig. 4 Basic model module

算(Cordonnier 等, 2020), 假设  $X \in \mathbf{R}^{T \times D_{in}}$  和  $Y \in \mathbf{R}^{T' \times D_{in}}$  分别为两个输入矩阵, 注意力层  $D_{in}$  输入  $D_{out}$  输出, 维度映射如下

$$f_{\text{attention}}(Q, K, V) = f_{\text{softmax}}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (6)$$

式中,  $Q = XW_Q$ ,  $K = XW_K$ ,  $V = XW_V$  注意力层由 Query 矩阵  $W_Q \in \mathbf{R}^{D_{in} \times D_k}$ , Key 矩阵  $W_K \in \mathbf{R}^{D_{in} \times D_k}$  和 Value 矩阵  $W_V \in \mathbf{R}^{D_{in} \times D_{out}}$  构成。当  $X \in \mathbf{R}^{T \times D_{in}}$  和  $Y \in \mathbf{R}^{T' \times D_{in}}$  为同一个矩阵时, 称之为自注意力模块。在式(1)计算中需要附加偏置, 此时 Key 和 Query 的投影的计算为  $K = XW_K + \mathbf{1}_{T \times 1}b_K$  和  $Q = YW_Q + \mathbf{1}_{T' \times 1}b_Q$  其中  $\mathbf{1}_{a \times b}$  是一个维度为  $a$  和  $b$  的全 1 矩阵。在引入偏置后对  $QK^T$  展开, 即

$$\begin{aligned} QK^T &= (XW_Q + \mathbf{1}_{T \times 1}b_Q^T)(YW_K + \mathbf{1}_{T' \times 1}b_K^T) = \\ &\underbrace{XW_Q W_K^T Y^T}_{\text{context}} + \underbrace{\mathbf{1}_{T \times 1}b_Q^T W_K^T Y}_{\text{content}} + \\ &XW_Q b_K \mathbf{1}_{1 \times T} + \mathbf{1}_{T \times T}b_Q^T b_K \end{aligned} \quad (7)$$

可以发现式(7)后两项对整个计算的结果是没有贡献的, 本文计算得到的 Attention 是一个  $T \times T$  的矩阵, 而  $XW_Q b_K$  和  $\mathbf{1}_{T \times 1}b_Q^T b_K$  得到的都是  $T \times 1$  的矩阵, 最后通过  $\mathbf{1}_{1 \times T}$  而扩充成  $T \times T$  的矩阵, 因此每一行中列的内容是相同的, 最后两项对 Attention 计算得到的是一个常量, 所以最后两项可以舍去。此外, 式(7)中的第 1 项计算了 Query 和 Key, 从不同序列图像中捕获信息(context)。而第 2 项只包含了 Key

行内信息, 捕获了同一序列图像的信息(content)。

通常情况下, 注意力机制将相同的层融合, 组成  $N_h$  头的多头注意力机制, 即

$$\begin{aligned} f_{\text{MultiHead}}(X, Y) &= \text{concat}_{i \in [N_h]} [H^{(i)}] W^0 \\ H^{(i)} &= f_{\text{attention}}(XW_Q^{(i)}, YW_K^{(i)}, YW_V^{(i)}) \end{aligned} \quad (8)$$

式中, 参数矩阵  $W_Q^{(i)}, W_K^{(i)} \in \mathbf{R}^{D_{in} \times d_k}$  和  $W_V^{(i)} \in \mathbf{R}^{D_{in} \times d_{out}}$  在每个注意力机制模块中得到,  $i$  表示不同的注意力模块, 以及额外的参数矩阵  $W^0 \in \mathbf{R}^{N_h d_{out} \times D_{out}}$  使得  $N_h$  头的输出空间为  $\mathbf{R}^{D_{out}}$ ,  $N_h$  表示注意力头的数量。同时定义  $D_k = N_h \times d_k$  为总的列空间维度, 而  $d_k$  为每个注意力头中列的维度。图 5 是编码块和多层感知器的基本结构。

## 2 实验及分析

### 2.1 实验室参数设置

本文分类模型在训练时, batch 的大小设置为 4, 学习率设置为  $1E-5$ , 最大循环次数设置为 100, 实验采用 NVIDIA TITAN XP GPU。实验测试采用交叉验证的方法, 将数据集平均分成 5 组, 进行 5 次交叉验证, 取所有交叉验证的结果的平均准确率作为实验的最终结果。

### 2.2 数据集

实验采用的数据集为急性缺血性卒中患者病例



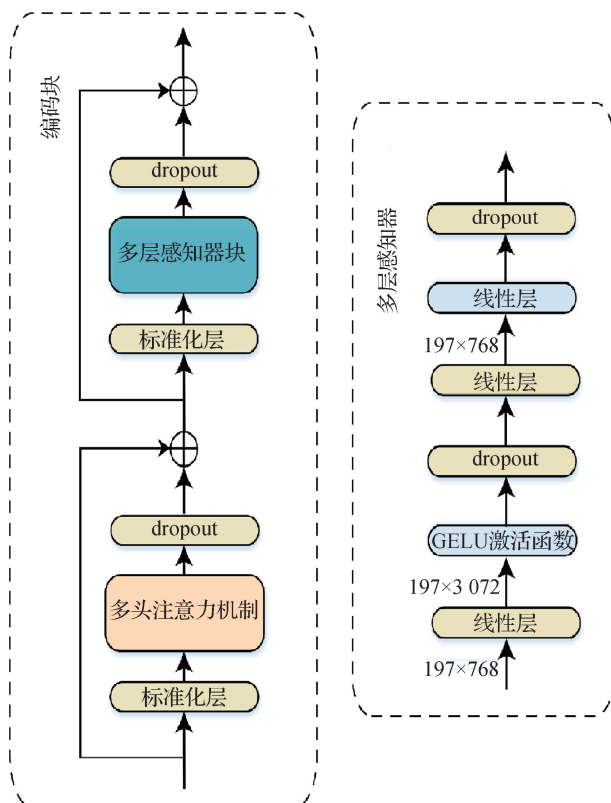


图5 编码块和多层感知器(MLP)的基本结构

Fig. 5 The basic structure of the encoder and MLP(multi-layer perception) layers

集,包括1 037名急性缺血性卒中患者和1 974名正常人面部和手部相对应的高分辨率图像。实验前,对数据集进行筛选,手部大鱼际处皮肤有脱皮、伤疤等现象以及图像光度过于明亮或黑暗的图像不参与实验。数据集有995例正样本和1 974负样本数据,正负样本数据不平衡,进行数据扩充是非常有必要的。

### 2.3 模型评价指标

使用准确率(acc)、特异性(spe)、敏感性(sen)和F1-分数(F1)等评估性能指标作为模型评价指标,计算为

$$f_{acc} = \frac{TP + TN}{TP + FP + TN + FN} \quad (9)$$

$$f_{sen} = \frac{TP}{TP + FN} \quad (10)$$

$$f_{spe} = \frac{TN}{TN + FP} \quad (11)$$

$$F1 = \frac{2TP}{2TP + FN + FP} \quad (12)$$

式中,TP为真阳性,TN为真阴性,FP为假阳性,FN

为假阴性。

### 2.4 数据预处理

由于实验数据集中正负样本不平衡可能会给实验带来负面影响,因此需要对数据集中的正样本进行数据扩充,提高模型鲁棒性,避免过拟合。数据增强的典型方法有翻转、旋转、缩放、随机拆件或补零、色彩抖动和加噪声。为避免数据扩充过程中对纹理特征产生影响,尤其眉心区域的纹理特征中纵向纹理特征对最终结果起到关键作用,水平翻转在增强数据的同时可以最大程度保留纵向纹理特征不对模型性能产生影响,本文只采用对图像进行水平翻转的方式。

在数据扩充完成之后采用脸部关键点识别方法(陈雨薇,2019)对脸部山根处和手部大鱼际处进行感兴趣区域提取,大鱼际处利用关键点算法中识别到中指根部点、大拇指根部点和手腕处点,通过坐标构建矩阵。山根处以面部两眉毛靠中心方向的边界点坐标构建矩阵,分别得到224×224像素的感兴趣区域图像。

### 2.5 消融实验

本文模型主要由特征提取、Transformer和多分支分类3个模块组成,为验证模块的有效性,对所采用的模块进行组合实验,在收集的数据集上进行验证,并进行多次实验得出最终实验结果,如表1所示。

表1 不同特征的对比结果

Table 1 Comparative results of different features

| 模型                       | 准确率          | 敏感性          | 特异性          | F1-分数        |
|--------------------------|--------------|--------------|--------------|--------------|
| YCbCr-MuTransformer      | 79.40        | 83.63        | 78.74        | 78.53        |
| GLCM-MuTransformer       | 80.46        | 83.11        | 79.32        | 79.89        |
| YCbCr-GLCM-MuTransformer | <b>83.53</b> | <b>85.75</b> | <b>81.05</b> | <b>81.82</b> |
| YCbCr-GLCM-Transformer   | 82.41        | 83.35        | 80.24        | 81.09        |

注:加粗字体为每列最优值。

由表1消融实验结果可知,分别提取颜色特征(YCbCr)和纹理特征GLCM的模型在数据集上取得了79.40%和80.46%的检测急性缺血性卒中的准确率,而颜色特征和纹理特征融合之后的模型在数据集上取得了83.53%的准确率,明显优于特征融合前。由此可知,颜色特征和纹理特征在Transformer

模型分类中可以有效提高分类准确性,两个特征的融合能够进一步提升检测精度。在颜色特征和纹理特征融合的前提下,对脸部和手部图像仅使用一个 Transformer 模块与分别对脸部和手部图像使用 Transformer 模块 (MuTransformer) 进行比较,使用一个 Transformer 模块的模型分类准确率下降了 1% 左右。这说明不同部位的特征在最终的检测中起到不同的作用,将特征融合到一个 Transformer 模块之后,不同部位相同特征之间的差异很容易在特征融合过程消失。融合之后的颜色和纹理特征,Transformer 模块的空间特征以及注意力特征,这几种特征的结合能有效提升急性缺血性卒中模型的性能。

## 2.6 与其他模型对比

通过与其他方法进行对比,验证本文方法的有效性。表 2 列出了本文方法与其他方法在同一种数据集上的多特征对比结果和在不同数据集的单特征

对比结果。本文方法在数据集上获得 83.53% 的准确率,优于现在对比方法。如骆轶姝等人(2020)在数据集上采用多特征的不同长短期记忆(LSTM)模型,其中基于双向 LSTM 和 LSTM 组合的多特征模型 (multiple features long short-term memory, MBLSTM)取得了类似模型中最高的准确率,可以充分说明在同样的多特征模型中双向 LSTM 可以获取更多的空间特征。使用 VGG16 (Visual Geometry Group Network 16) 和 CNN 作为特征提取的模型在一定程度上比未进行特征提取的模型性能好,同时 VGG16 相比 CNN 层数更深、通道数更多,因此有更多的特征被提取出来。在实验中,特征提取方法对模型精度有着较大的影响,本文模型结合中医望诊的基本理论,能够有效提取相关特征,通过与深度模型特征相结合,可以最大化地利用不同特征,从而取得最好的模型性能。

表 2 不同方法的对比结果  
Table 2 Comparative results of different methods

|                      |                     |              |              |              | /%           |
|----------------------|---------------------|--------------|--------------|--------------|--------------|
| 方法                   |                     | 准确率          | 敏感性          | 特异性          | F1-分数        |
| 多特征                  | MLSTM               | 76.06        | 80.25        | 75.20        | 75.27        |
|                      | MBLSTM-LSTM         | 78.96        | 81.24        | 78.47        | 78.41        |
|                      | MBLSTM              | 77.33        | 80.30        | 76.98        | 76.04        |
|                      | CNN-LSTM            | 79.03        | 81.89        | 77.96        | 78.87        |
|                      | VGG16-LSTM          | 79.23        | 81.93        | 78.87        | 78.95        |
|                      | VGG16-BLSTM         | 80.63        | 82.77        | 79.25        | 79.15        |
| M-MuTransformer( 本文) |                     | <b>83.53</b> | <b>85.75</b> | <b>81.05</b> | <b>81.82</b> |
| 单特征                  | YCbCr-MuTransformer | 79.40        | 83.63        | 78.74        | 78.53        |
|                      | GLCM-MuTransformer  | 80.46        | 83.11        | 79.32        | 79.89        |
|                      | C-BLSTM             | 76.72        | 79.70        | 76.51        | 75.94        |
|                      | S-BLSTM             | 77.35        | 80.37        | 76.94        | 76.64        |
|                      | J-BLSTM             | 76.54        | 78.67        | 76.13        | 75.36        |

注:加粗字体为每列最优值。

为了验证中医望诊的基本理论在检测急性缺血性卒中模型中是否起到作用,在表 2 中列出了相关文献在采用不同特征进行检测分类的性能表现。表 2 中 C-BLSTM 代表超声特征模型,S-BLSTM 代表生化特征模型,J-BLSTM 代表基本信息特征模型,从表 2 可以看出,采用中医望诊对应的颜色特征或纹理

特征,在模型的性能上有约 3% 的提升,充分证明了中医望诊理论对应的颜色特征和纹理特征可以有效提升模型性能。

图 6(a)(b)给出了同类型方法中表现最好的模型的敏感性和特异性,可以更直观地对比不同模型的性能。将数据集平均分为 5 组进行交叉验证,

图 6(a) 为不同方法的敏感性指标比较, M-MuTransformer 的敏感性对比其他方法具有明显优势, 最高达到了 85.75%, 相较其他模型中表现最好的 VGG16-BLSTM 提升了 2.98%。本文模型在所有测试数据中敏感性均高于其他方法, 可以说明本文方法在真阳性判断上的性能优势。图 6(b) 为在交叉验证中不同方法的特异性对比。M-MuTransformer 方法的特异性在数据子集 Fold 中略有降低, 但相较其他方法的整体特异性仍较有优势。多数方法的特异性在数据子集 Fold 达到最高, 在此数据子集下 M-MuTransformer 方法的特异性相比 VGG16-BLSTM 方法仅低了 1.22%。通过分析敏感性和特异性, 本文方法相比其他方法有更好的鲁棒性和稳定性。

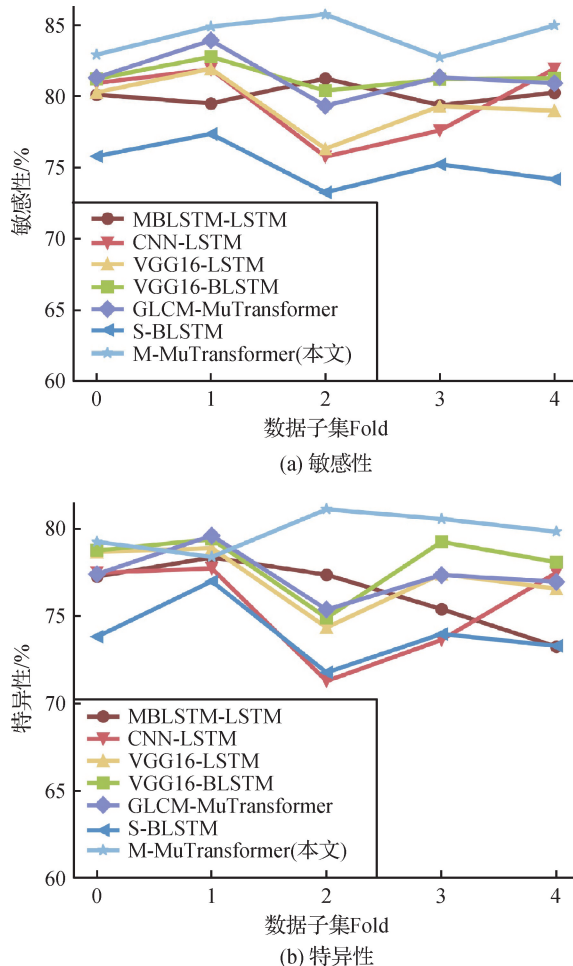


图 6 不同方法敏感性和特异性对比

Fig. 6 Comparison of sensitivity and specificity of different methods

### 3 结 论

目前关于急性缺血性卒中疾病的早期诊断的研

究还不够深入, 根据中医望诊理论基础在检测早期急性缺血性卒中方面有显著效果, 结合中医望诊中的颜色和纹理特征, 对山根和大鱼际感兴趣区域图像提出了基于双 Transformer 模块的端到端联合分类检测方法, 利用 YCbCr 颜色空间和 GLCM 获取高质量的数据, 加速模型的收敛过程; 采用患者脸部和手部图像提取特征信息; 通过自注意力机制学习特征间的关联并分配权重, 增强模型学习能力, 提升模型性能。本文模型诊断效果良好, 通过自动辅助诊断降低了主观因素影响, 在缺血性脑卒中辅助诊断研究中具有一定的价值, 为临床医生缺血性脑卒中疾病诊断提供了决策参考, 也为患者进行有效的自我筛查提供了一种新方法。本文方法在辅助检测急性缺血性卒中疾病中取得了一定效果, 但是该方法只结合中医望诊方法中的局部信息, 如何让中医望诊的整个体系融入辅助诊断中仍需要进一步研究。基于这些问题, 后续将会展开以下工作:

1) 结合中医望诊的理论知识, 将更多的特征有机结合, 进一步提高模型性能。

2) 模型整体使用的数据量偏少, 加大对数据的收集来促进模型性能的提升。

3) 目前, 模型体量较大且不方便大规模的应用, 将对模型进一步优化来满足实际需求。

### 参考文献 (References)

- Boling B and Keinath K. 2018. Acute ischemic stroke. AACN Advanced Critical Care, 29(2): 152-162 [DOI: 10.4037/aacnacc2018483]
- Chen Y W. 2019. Face Detection and Facial Landmark Localization Based on Improved MTCNN Model. Shanghai: Donghua University (陈雨薇. 2019. 基于改进 MTCNN 模型的人脸检测与面部关键点定位. 上海: 东华大学)
- Cordonnier J B, Loukas A and Jaggi M. 2020. Multi-head attention: collaborate instead of concatenate [EB/OL]. [2020-01-29]. <https://arxiv.org/pdf/2006.16362v1.pdf>
- Gao J M, Lyu M, Xie W W, Liu X Y, Zhao B C and Zhu Y. 2019. Regularity of traditional Chinese medicine prescriptions for same treatment for cardiovascular and cerebrovascular diseases. China Journal of Chinese Materia Medica, 44(1): 193-198 (高佳明, 吕明, 解微微, 刘昕彦, 赵步长, 朱彦. 2019. 中医药心脑血管疾病同治的方剂用药规律分析. 中国中药杂志, 44(1): 193-198) [DOI: 10.19540/j.cnki.cjmm.20181101.007]
- Gao L, Wang P P and Li N. 2007. Observation for cause; clinical considerations in a case of stroke. Chinese Journal of Integrated Traditional and Western Medicine in Intensive and Critical Care, 14(4):



- 252-253 (高利, 王平平, 李宁. 2007. 望诊寻因——1例脑卒中患者引发的临床思考. 中国中西医结合急救杂志, 14(4): 252-253) [DOI: 10.3321/j.issn:1008-9691.2007.04.026]
- Gao Y R, Han X J, Wang L Y, Liu D S and Ren C. 2020. Research progress of the ear inspection. *China Journal of Traditional Chinese Medicine and Pharmacy*, 35(8): 4052-4054 (邵亚茹, 韩学杰, 王丽颖, 刘大胜, 任聪. 2020. 大鱼际望诊法研究进展. 中华中医药杂志, 35(8): 4052-4054)
- Hao P P, Jiang F, Chen Y G, Yang J M, Zhang K, Zhang M X, Zhang C, Zhao Y X and Zhang Y. 2015. Traditional Chinese medication for cardiovascular disease. *Nature Reviews Cardiology*, 12(6): #318 [DOI: 10.1038/nrcardio.2015.60]
- Hsieh M J, Chen Y J, Tang S C, Chen J H, Lin L C, Seak C J, Lee J T, Chang K C, Lien L M, Chan L, Liu C H, Hsieh C Y, Chern C M, Chen J C, Chiu T F, Hung S C, Ng C J and Jeng J S. 2021. 2020 Guideline for Prehospital management, emergency evaluation and treatment of patients with acute ischemic stroke: a guideline for healthcare professionals from the Taiwan society of emergency medicine and Taiwan stroke society. *Journal of Acute Medicine*, 11(1): 12-17 [DOI: 10.6705/j.jacme.202103\_11(1).0002]
- Karthik R and Menaka R. 2018. Computer-aided detection and characterization of stroke lesion—a short review on the current state-of-the-art methods. *The Imaging Science Journal*, 66(1): 1-22 [DOI: 10.1080/13682199.2017.1370879]
- Lin Y. 2020. Complexion classification based on convolutional neural network. *Journal of Artificial Intelligence Practice*, 3(1): 22-30 [DOI: 10.23977/jaip.2020.030105]
- Liu Z W. 2015. Research on Stroke Recurrence Prediction Based on Machine Learning. Changsha: Hunan University (刘泽文. 2015. 基于机器学习的脑卒中复发预测模型研究. 长沙: 湖南大学)
- Luo X Z, Wen X P, He J Y, Huang J T and Tang C Z. 2017. Analysis of the influencing factors of post-stroke depression: based on machine learning. *Journal of Traditional Chinese Medicine*, 58(17): 1478-1481 (罗晓舟, 温小鹏, 何家扬, 黄健婷, 唐纯志. 2017. 基于机器学习的卒中后抑郁影响因素分析. 中医杂志, 58(17): 1478-1481) [DOI: 10.13288/j.11-2166/r.2017.17.011]
- Luo Y S, Shao Y Y and Chen D H. 2020. Diagnosis model of ischemic stroke based on LSTM with multi-feature combination. *Intelligent Computer And Applications*, 10(10): 74-79 (骆轶姝, 邵圆圆, 陈德华. 2020. 基于 LSTM 多特征联合的缺血性脑卒中诊断模型. 智能计算机与应用, 10(10): 74-79) [DOI: 10.3969/j.issn.2095-2163.2020.10.018]
- Seshadri S and Wolf P A. 2007. Lifetime risk of stroke and dementia: current concepts, and estimates from the Framingham study. *The Lancet Neurology*, 6(12): 1106-1114 [DOI: 10.1016/S1474-4422(07)70291-0]
- Shaik K B, Ganesan P, Kalist V, Sathish B S and Jenitha J M M. 2015. Comparative study of skin color detection and segmentation in HSV and YCbCr color space. *Procedia Computer Science*, 57: 41-48 [DOI: 10.1016/j.procs.2015.07.362]
- Strong K, Mathers C and Bonita R. 2007. Preventing stroke: saving lives around the world. *The Lancet Neurology*, 6(2): 182-187 [DOI: 10.1016/S1474-4422(07)70031-5]
- Virani S S, Alonso A, Aparicio H J, Benjamin E J, Bittencourt M S, Callaway C W, Carson A P, Chamberlain A M, Cheng S, Delling F N, Elkind M S V, Evenson K R, Ferguson J F, Gupta D K, Khan S S, Kissela B M, Knutson K L, Lee C D, Lewis T T, Liu J X, Loop M S, Lutsey P L, Ma J, Mackey J, Martin S S, Matchar D B, Mussolino M E, Navaneethan S D, Perak A M, Roth G A, Samad Z, Satou G M, Schroeder E B, Shah S H, Shay C M, Stokes A, VanWagner L B, Wang N Y and Tsao C W. 2021. Heart disease and stroke statistics—2021 update: a report from the American heart association. *Circulation*, 143(8): e254-e743 [DOI: 10.1161/CIR.0000000000000950]
- Wu M H, Song R R and Liu M. 2017. Video shadow elimination algorithm combining HSV with texture feature. *Journal of Image and Graphics*, 22(10): 1373-1380 (武明虎, 宋冉冉, 刘敏. 2017. 结合 HSV 与纹理特征的视频阴影消除算法. 中国图象图形学报, 22(10): 1373-1380) [DOI: 10.11834/jig.170151]
- Xi L. 2016. Classification of Stroke Based on Microwave Experiment Platform. Shanghai: Donghua University (席恋. 2016. 基于微波实验平台的脑中风分类判断研究. 上海: 东华大学)
- Yu C Y, Wen L F, Yang G and Wang Y T. 2019. Video person reidentification based on BiLSTM and attention mechanism. *Journal of Image and Graphics*, 24(10): 1703-1710 (余晨阳, 温林凤, 杨钢, 王玉涛. 2019. 结合 BiLSTM 和注意力机制的视频行人再识别. 中国图象图形学报, 24(10): 1703-1710) [DOI: 10.11834/jig.190637]
- Zhang C, Li W W, Xiao J and Sun C L. 2020. Application of TCM inspection in clinical practice. *Henan Traditional Chinese Medicine*, 40(6): 839-843 (张超, 李唯薇, 肖静, 孙成力. 2020. 中医望诊在临床中的应用. 河南中医, 40(6): 839-843) [DOI: 10.16367/j.issn.1003-5028.2020.06.0212]
- Zhang J, Wang C, Zhuo L and Yang Y C. 2014. Uniform color space based facial complexion recognition for traditional Chinese medicine//Proceedings of the 13th International Conference on Control Automation Robotics and Vision. Singapore, Singapore: IEEE: 631-636 [DOI: 10.1109/ICARCV.2014.7064377]
- Zhang W T, Gao K, Liu J W, Zhao H H, Wang J A, Li Y B, Murtaza G, Chen J X and Wang W. 2013. A review of the pharmacological mechanism of traditional Chinese medicine in the intervention of coronary heart disease and stroke. *African Journal of Traditional, Complementary and Alternative Medicines*, 10(6): 532-537 [DOI: 10.4314/ajtcam.v10i6.24]

## 作者简介



赵琛琦, 1997年生, 男, 硕士研究生, 主要研究方向为医学图像处理与深度学习、自然语言处理。

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



赵涓涓,通信作者,女,教授,主要研究方向  
为智能信息处理、图像识别与机器学习。

E-mail:zhaojuanjuan@ 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

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

王麒达,男,硕士研究生,主要研究方向为医学图像处理与深  
度学习。E-mail:wangqida0375@ link. tyut. edu. cn

李慧芝,女,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,博士研究生,主要研究方向为中医大数据分析  
与深度学习。E-mail:zhaozjuan0064@ link. tyut. edu. cn