



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

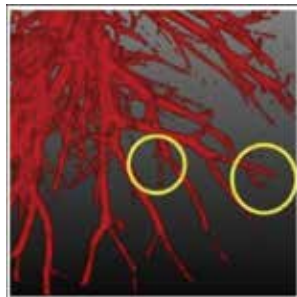
Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

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

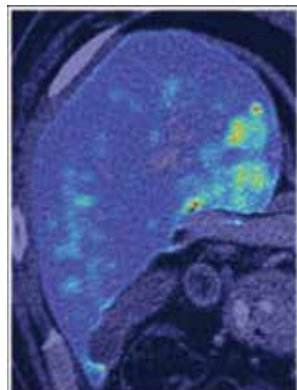
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

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

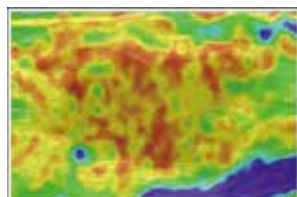
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



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

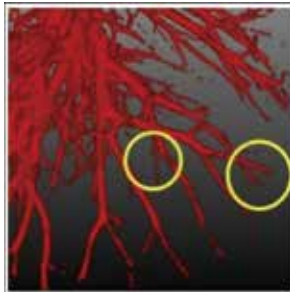
肖照林, 杨志林, 刘欢, 金海燕 0988

基于图像的介入器械前端变形建模与受力估计

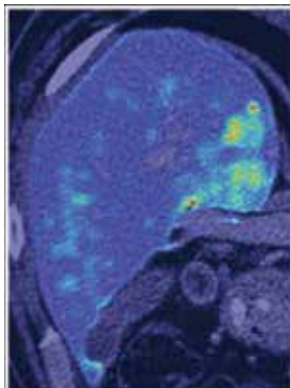
任龙飞, 孟偲, 刘博, 白相志, 周付根 1001

CONTENTS

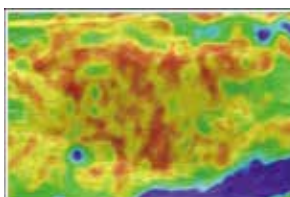
JOURNAL OF IMAGE AND GRAPHICS



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: TP391.4 文献标识码: A 文章编号: 1006-8961(2022)03-0898-13

论文引用格式: Gong R L, Shi J, Zhou W J and Wang C. 2022. Two-stage deep transfer learning for human breast ultrasound computer-aided diagnosis. Journal of Image and Graphics, 27(03): 0898-0910 (贡荣麟, 施俊, 周玮珺, 汪程. 2022. 面向乳腺超声计算机辅助诊断的两阶段深度迁移学习. 中国图象图形学报, 27(03): 0898-0910 [DOI:10.11834/jig.210674])

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¹, 施俊^{1*}, 周玮珺², 汪程¹

1. 上海大学通信与信息工程学院, 上海 200444; 2. 安徽医科大学第一附属医院超声科, 合肥 230032

摘要: **目的** 为了提升基于单模态 B 型超声(B 超)的乳腺癌计算机辅助诊断(computer-aided diagnosis, CAD)模型性能, 提出一种基于两阶段深度迁移学习(two-stage deep transfer learning, TSDTL)的乳腺超声 CAD 算法, 将超声弹性图像中的有效信息迁移至基于 B 超的乳腺癌 CAD 模型之中, 进一步提升该 CAD 模型的性能。 **方法** 在第 1 阶段的深度迁移学习中, 提出将双模态超声图像重建任务作为一种自监督学习任务, 训练一个关联多模态深度卷积神经网络模型, 实现 B 超图像和超声弹性图像之间的信息交互迁移; 在第 2 阶段的深度迁移学习中, 基于隐式的特权信息学习(learning using privileged information, LUPI)范式, 进行基于双模态超声图像的乳腺肿瘤分类任务, 通过标签信息引导下的分类进一步加强两个模态之间的特征融合与信息交互; 采用单模态 B 超数据对所对应通道的分类网络进行微调, 实现最终的乳腺癌 B 超图像分类模型。 **结果** 实验在一个乳腺肿瘤双模超声数据集上进行算法性能验证。实验结果表明, 通过迁移超声弹性图像的信息, TSDTL 在基于 B 超的乳腺癌诊断任务中取得的平均分类准确率为 $87.84 \pm 2.08\%$ 、平均敏感度为 $88.89 \pm 3.70\%$ 、平均特异度为 $86.71 \pm 2.21\%$ 、平均约登指数为 $75.60 \pm 4.07\%$, 优于直接基于单模态 B 超训练的分类模型以及多种典型迁移学习算法。 **结论** 提出的 TSDTL 算法通过两阶段的深度迁移学习, 将超声弹性图像的信息有效迁移至基于 B 超的乳腺癌 CAD 模型, 提升了模型的诊断性能, 具备潜在的应用可行性。

关键词: B 型超声成像; 超声弹性成像; 乳腺癌计算机辅助诊断; 特权信息学习(LUPI); 深度迁移学习; 自监督学习(SSL)

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

Gong Ronglin¹, Shi Jun^{1*}, Zhou Weijun², Wang Cheng¹

1. School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China;

2. Department of Ultrasound, the First Affiliated Hospital of Anhui Medical University, Hefei 230032, China

Abstract: **Objective** B-mode ultrasound (BUS) provides information about the structure and morphology information of human breast lesions, while elastography ultrasound (EUS) can provide additional bio-mechanical information. Dual-modal ultrasound imaging can effectively improve the accuracy of the human breast cancer diagnosis. The single-modal ultrasound-based computer-aided diagnosis (CAD) model has its potential applications. Deep transfer learning is a significant branch of transfer learning analysis. This technique can be utilized to guide the information transfer between EUS images and BUS images. However, clinical image samples are limited based on training deep learning models due to the high cost of data

收稿日期: 2021-08-13; 修回日期: 2021-11-18; 预印本日期: 2021-11-25

* 通信作者: 施俊 junshi@shu.edu.cn

基金项目: 国家自然科学基金项目(81830058, 81627804)

Supported by: National Natural Science Foundation of China (81830058, 81627804)

collection and annotation. Self-supervised learning (SSL) is an effective solution to demonstrate its potential in a variety of medical image analysis tasks. In respect of the SSL pipeline, the backbone network is trained based on a pretext task, where the supervision information is generated from the train samples without manual annotation. Based on the weight parameters of the trained backbone network, the obtained results are then transferred to the downstream network for further fine-tuning with small size annotated samples. A step-based correlation multi-modal deep convolution neural network (CorrMCNN) has been facilitated to conduct a self-supervised image reconstruction task currently. In the training process, the model transfers the effective information between two modalities to optimize the correlation loss through SSL-based deep transfer learning. Since each BUS and EUS scan the same lesion area for the targeted patient simultaneously, the analyzed results are demonstrated in pairs and share labels. Learning using privileged information (LUPI) is a supervised transfer learning paradigm for paired source domain (privileged information) and target domain data based on shared labels. It can exploit the intrinsic knowledge correlation between the paired data in the source domain and target domain with shared labels, which guides knowledge transfer to improve model capability. Since the label information is used to guide transfer learning in classifiers, the current LUPI algorithm focus on the classifier. Feature representation is also the key step for a qualified CAD system. A two-stage deep transfer learning (TSDTL) algorithm is demonstrated for human breast ultrasound CAD, which transfers the clear information from EUS images to the CAD model of BUS-based human breast cancer and further improves the performance of the CAD model. **Method** In the first stage of deep transfer learning, an SSL task is first designed based on dual-modal ultrasonic image reconstruction, which trains the CorrMCNN model to conduct the interactive transfer of information between the two modal images of BUS and EUS images. The bi-channel encoder networks are adopted to learn the feature representation derived from the dual-modal images, respectively. The high-level learned features are used following for concatenation to obtain the joint representation. The original BUS and EUS images are reconstructed from the joint feature representation through the bi-channel decoder networks. In the training process, the network implicitly implements deep transfer learning via the correlation loss optimization amongst high-level features derived from two channels. In the second stage of deep transfer learning, the pre-training backbone network is reused followed by a sub-network for classification. The BUS and EUS images are input into this new network for targeted breast cancer classification based on dual-modal ultrasound images. In this training process, the data of the source domain and target domain can be applied to supervised transfer learning with the shared labels, and this strategy belongs to the general LUPI paradigm. Consequently, it can be considered that this deep transfer learning stage implicitly conducts knowledge transfer under the LUPI paradigm, which is based on the dual-modal ultrasound breast cancer classification task. In the final stage, the corresponding channel sub-network is fine-tuned with single-modal ultrasound data, which obtains an accurate breast cancer B-mode image classification model. The obtained single-channel network is the final network model of BUS-based breast cancer CAD, which can be directly applied to the diagnostic tasks of the emerging BUS images. **Result** The performance of the algorithm is demonstrated on a breast tumor dual-modal ultrasound dataset. Our illustrated TSDTL achieves the mean classification accuracy of $87.84 \pm 2.08\%$, sensitivity of $88.89 \pm 3.70\%$, specificity of $86.71 \pm 2.21\%$, and Youden index of $75.60 \pm 4.07\%$ respectively, which develops the classification model trained on single-modal BUS images and a variety of typical deep transfer learning algorithms. **Conclusion** Our TSDTL algorithm analysis can effectively transfer the information of EUS to the BUS-based human breast cancer CAD model through our illustrated two-stage deep transfer learning.

Key words: B-mode ultrasound imaging; elastography ultrasound imaging; computer-aided diagnosis of breast cancer; learning using privileged information (LUPI); deep transfer learning; self-supervised learning (SSL)

0 引言

乳腺癌是女性最常见的癌症之一,严重危害女性身体健康。其准确诊断对于后续的治疗以及预后具有重要意义(Siegel等,2020)。医学影像是临床

进行乳腺癌筛查和鉴别诊断的重要工具。其中,B型超声(B超)是一种无创、无辐射的实时成像诊断技术,主要对浅表和内脏器官等软组织进行实时成像,在临床诊断中为医生提供病灶的结构形态及病理学改变所引起的结构变化的信息,是目前临床诊断乳腺癌的常规工具(Wang等,2019)。随着人

工智能技术的飞速发展,基于B超的乳腺癌计算机辅助诊断(computer-aided diagnosis, CAD)系统的研究也已经广泛开展,其具有可重复性和一致性高的优点,可以辅助医生提高诊断的准确性(Sutton, 1989)。

值得注意的是,单一模态的B超所提供的信息类型主要局限于病灶的结构形态和声像,缺乏与病灶相关的其他方面的诊断信息,这在一定程度上制约了CAD模型的性能。而实际上,乳腺癌在其发展过程中,细胞外基质逐渐变硬,胶原纤维的形状和排列发生改变,导致乳腺癌病灶本身及周围基质发生力学上的改变,硬度或弹性随之而增加,但这种硬度的改变在B超中并不能有效显示和表征(Shi等, 2018)。近年来,基于超声弹性成像的乳腺癌诊断在临床逐步开展。该技术以加压等方式探测不同组织的硬度或弹性的改变,从而以伪彩编码的形式对不同组织和病灶的硬度信息进行实时成像显示,提供病灶的生物力学相关的信息(Sigrist等, 2017)。因此,弹性成像技术可以对B超成像提供补充的诊断信息,联合应用二者能有效提高乳腺癌诊断的准确性(Carlsen等, 2013)。于是,基于双模超声的CAD模型也是目前的研究热点之一,其性能也优于单模超声CAD模型(Mahmood等, 2020)。然而,由于多种因素的制约,目前超声弹性成像在临床的应用尚未完全普及,特别是在基层医院通常缺乏超声弹性成像设备以及能熟练操控该设备的有经验的超声医生。因此,研发基于单模态B超的乳腺癌CAD模型仍然具有更为广泛和灵活的应用前景,但其性能也还有待进一步提高。

在各种机器学习方法中,迁移学习可以通过将在源域中学习到的知识或训练的模型迁移至目标域,实现从源域到目标域的知识迁移,从而提升目标域模型的性能(Weiss等, 2016; 杜章锦等, 2020)。因此,可以将超声弹性图像作为源域,通过迁移学习技术将该图像所提供的生物力学相关信息迁移至基于B超的乳腺癌CAD模型,具有提升其性能的可行性。

深度迁移学习是目前迁移学习领域的前沿方向,已经应用于医学影像CAD(Yu等, 2017)。从临床数据应用角度,医学影像CAD中的深度迁移学习可以分为不同疾病之间的迁移和不同模态之间的迁移。前者指目标域(诊断模态)和源域数据属于同

模态的影像,对应疾病类型不同;后者指源域和目标域数据对应的疾病相同,但两者影像属于不同模态。超声弹性图像至B超图像之间的信息迁移属于不同模态之间的迁移,为深度迁移学习方法的使用提供了可行性。

然而,临床影像数据通常面临数据采集周期长、成本高和标记困难等问题(Sahiner等, 2008),导致用于训练深度学习模型的样本一般较少。为了缓解小样本学习问题,多种机器学习方法已经相继提出,自监督学习(self-supervised learning, SSL)就是其中一种有效的方法(Jing和Tian, 2021)。该方法通常利用数据自身蕴含的属性和信息,从无标签数据中自主生成伪标签,以此来构建代理任务预训练骨干网络。然后,将骨干网络的权重参数迁移至下游目标任务,使用少量标记的训练样本进行微调训练,最终提升模型在目标任务上的性能(Liu等, 2020)。为了应对临床影像数据样本数有限的问题,研究人员将自监督学习引入医学图像领域,并将其应用于分类、分割和定位等任务(Chen等, 2019),研究结果表明其具有有效性。

Bhatt等人(2019)提出了一种关联多模态深度卷积神经网络(step-based correlation multi-modal deep convolution neural network, CorrMCNN),分别通过两个编解码结构的子网络重建两个模态的信息,并使用关联损失函数对不同模态的高维表征进行约束,以此来最大化两者之间的相关性。目前,CorrMCNN主要应用于多模态数据的表征学习,而基于CorrMCNN的图像重建操作本质上属于一种自监督学习范式,并且是一种新的多模态图像的自监督学习任务。网络能在训练重建图像信息的过程中,自主学习到多模数据内在的高维特征表达。而将CorrMCNN应用于双模态超声图像重建,在进行自监督学习的同时,也实现了两个超声模态数据之间的信息迁移学习,即基于自监督学习的深度迁移学习。

另外,由于双模态的B超和超声弹性成像是针对同一患者的同一病灶区域进行扫描成像,两者成对出现,共享标签。利用标签信息的有监督迁移学习通常较无监督的迁移学习具有更好的迁移性能。而特权信息学习(learning using privileged information, LUPI)则是一种专门针对共享标签的成对源域(即特权信息)和目标域数据的有监督迁移学习范式,能通过共享标签来挖掘源域和目标域成对数据之间

内在的知识关联性,并引导知识迁移,提升迁移性能(Vapnik 和 Vashist,2009)。因此,LUPI 范式可以应用于共享标签的双模态超声图像之间的知识迁移。

由于在分类器中易于实现利用标签信息引导迁移学习,目前的 LUPI 研究主要提出多种新的分类器算法(Pechyony 和 Vapnik,2012;Li 等,2018)。其中,支持向量机 + (support vector machine plus,SVM +) 是一种典型的 LUPI 分类器,并且已有多种改进的 SVM + 算法相继提出,应用于各种迁移学习分类任务之中。目前,基于 LUPI 的分类算法研究已经成功应用于医学影像 CAD 领域。例如,Duan 等人(2014)将单核苷酸作为特权信息,以常规检查中更易获得的眼底图像作为诊断模态,共同训练 SVM + 分类器模型,通过迁移单核苷酸信息,提升了基于眼底图像的青光眼诊断准确性;Zhang 等人(2021)以 3 个典型时相的超声造影图像作为特权信息,以 B 超图像作为诊断模态,提出了一种多核学习 SVM + 算法,以有效地从多视图超声造影图像中传递多源域知识,从而提升基于 B 超的肝癌 CAD 模型性能。上述研究表明了 LUPI 算法在 CAD 领域的可行性和应用潜力。但是,目前的 LUPI 算法集中于分类器的研究。实际上,特征表达是决定一个 CAD 系统性能的关键环节,而基于深度学习的 LUPI 算法的研究尚处于起步阶段,有待进一步的研究。

综上,为了提升基于单一模态 B 超的乳腺 CAD 的诊断性能,提出一种基于两阶段深度迁移学习(two-stage deep transfer learning,TSDTL)算法,实现超声弹性图像向基于 B 超的 CAD 模型的知识迁移。在第 1 阶段的深度迁移学习中,提出将双模态超声图像重建任务作为一种自监督学习任务,训练一个关联多模态深度卷积神经网络模型,实现 B 超图像和超声弹性图像之间的信息交互迁移;然后,在第 2 阶段的深度迁移学习中,进行基于双模态超声图像的乳腺肿瘤分类任务,通过标签信息引导下的分类进一步加强两个模态之间的特征融合与信息交互,以隐式的 LUPI 范式进行特征级迁移;最后,采用单模态 B 超数据对所对应通道的分类网络进行微调,实现最终的乳腺癌 B 超图像分类模型。

本文工作的主要创新点包括:

1)提出一种新的两阶段深度迁移学习算法应用于基于 B 超的乳腺癌 CAD 模型,有机结合自监督迁移学习和特征级 LUPI 这两种迁移模式,实现从

弹性超声图像到乳腺 B 超 CAD 模型的知识迁移,从而有效缓解双模态超声数据的小样本学习问题,提升 CAD 模型的诊断准确性,并且具有简单易行的优点。

2)在两级迁移学习过程中,均通过隐式的特征级深度迁移学习方式有效实现从弹性超声图像到基于 B 超的乳腺 CAD 模型的知识迁移,即通过构建双模态超声图像的自监督重建任务实现隐式的特征级迁移学习;然后利用双模态超声图像共享标签的特性,通过多模态分类任务以隐式 LUPI 的范式实现两个骨干网络之间的信息迁移,增强 B 超通道网络的特征表达,使其内在蕴含对应超声弹性图像的信息。

1 方 法

1.1 算法流程

提出的 TSDTL 算法流程如图 1 所示。该算法主要包含两个迁移学习阶段,以及一个单通道网络微调阶段。其具体核心步骤如下:

1)第 1 阶段自监督迁移学习。在本阶段的深度迁移学习中,设计了一个基于双模超声图像重建的自监督学习任务,以此来训练得到关联多模态深度卷积神经网络模型,从而实现 B 超和超声弹性成像两个模态图像之间的信息交互迁移。该 CorrMC-NN 模型首先使用双通道的编码器网络分别学习 B 超图像和超声弹性图像特征表达;然后将模型在两个通道中学习得到的高维特征进行拼接操作,得到联合特征表达;再采用双通道解码器网络从联合特征表达中分别重建出原有的 B 超图像和超声弹性图像。其中,网络在训练过程中通过优化相关性损失隐式地实现深度迁移学习。需要注意的是双通道网络中的编码器部分和联合特征表达层将作为第 2 阶段迁移学习的预训练骨干网络。

2)第 2 阶段多模态分类迁移学习。第 1 阶段获得的预训练骨干网络将用于第 2 阶段多模态分类任务。在其之后再添加新的全连接层以降低特征维度,并用 Softmax 函数计算分类结果。将 B 超和超声弹性图像输入该新网络进行基于双模超声的乳腺癌分类任务。在该阶段的网络训练中,由于双模超声图像样本共享标签,因此两个通道的网络之间可以在共同优化参数的过程中进行进一步的信息交互

迁移。在这一过程中,源域和目标域数据共享标签,并且有监督地进行迁移学习,符合一般的 LUPI 范式。因此,可以认为,该阶段是通过进行基于双模超声的乳腺癌分类任务,隐式地实现了 LUPI 范式下的特征级迁移学习。在完成此阶段迁移学习以后,再单独将 B 超对应的网络通道(不包含联合特征表达层)作为最终的基于 B 超的乳腺癌 CAD 模型的预训练骨干网络。

3) 单通道网络微调阶段。在得到新的预训练骨干网络后,添加全连接层和 Softmax 函数运算,构建最终的基于 B 超的乳腺癌 CAD 模型网络。在该

微调阶段中,只采用单模态的 B 超图像对该网络进行训练,以进一步提升针对 B 超数据的分类性能。

4) 测试阶段。将步骤 3) 中微调后的单通道网络作为最终的基于 B 超的乳腺癌 CAD 的网络模型,该模型可以直接应用于新的 B 超图像的诊断任务。

1.2 第 1 阶段自监督迁移学习

第 1 阶段借助图像恢复的自监督学习任务,对交叉网络进行预训练,以此来学习乳腺 B 超和超声弹性模态的公共表征信息。该阶段的 CorrMCNN 包含 4 个子模块:特征提取、非线性映射、联合特征表达和重构。

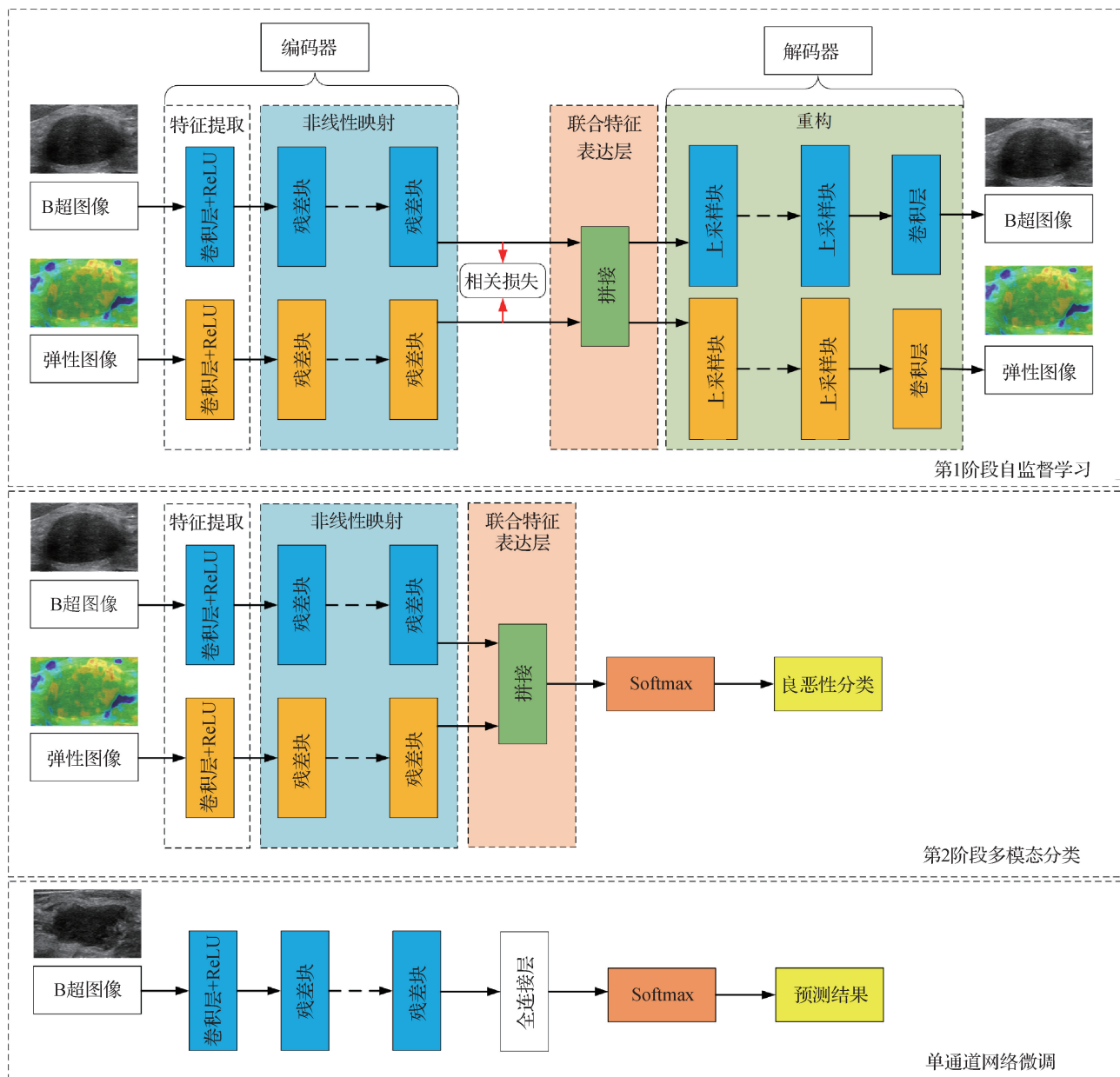


图 1 两阶段深度迁移学习算法流程图

Fig. 1 The flowchart of TSDTL algorithm

1.2.1 特征提取

利用特征提取模块学习 B 超和超声弹性模态的图像高维特征。特征提取模块由一个卷积层和一个 ReLU 激活层组成。其中,卷积层从图像中提取特征,而 ReLU 激活层负责实现特征的非线性映射。将输入的 B 超图像记为 I_B ,输入的超声弹性图像记为 I_E ,从 B 超图像提取的特征记为 F_B ,从 B 超图像提取的特征记为 F_E ,计算为

$$F_B = G_B(I_B, W_B) \quad (1)$$

$$F_E = G_E(I_E, W_E) \quad (2)$$

式中, G_B 为 B 超模态通道的特征提取映射函数, G_E 为超声弹性模态通道的特征提取映射函数, W_B 为特征提取子模块中 B 超模态通道的参数, W_E 为特征提取子模块中超声弹性模态通道的参数。

1.2.2 非线性映射

将特征提取模块输出的高维特征送入非线性映射模块,继而使用 8 个连续的残差块学习得到高级特征表达。第 1 阶段预训练网络中,每一通道的非线性映射都由 8 个残差块组成。

其中,非线性映射子模块中每个残差块的组成结构如图 2 所示。残差块使用残差学习策略,模块只需要学习输入特征与目标特征之间的差异,而两者间相同的信息可以通过跳连接直接传到残差块的输出端。假设输入为 x ,映射函数为 m ,最终输出 y 为

$$y = x + m(x) \quad (3)$$

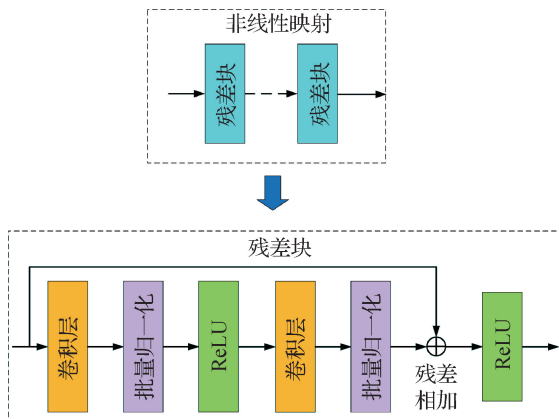


图2 非线性映射结构

Fig.2 Nonlinear mapping structure

将特征提取模块输出的高维特征 F_B 和 F_E 送入非线性映射的首个残差块,这些特征在第 1 个残差块进行计算,得到相对输入特征来说更高维度的特

征表达。然后,将首个残差块中得到的特征送入第 2 个残差块,以此类推,两个模态的特征分别经过 8 个残差块的计算,然后将得到的高级特征表达作为联合表征模块的输入。其中,将 B 超模态通道获得的高级特征表达记为 T_B ,超声弹性模态通道获得的高级特征表达记为 T_E ,计算为

$$T_B = M_B(F_B, \theta_B) \quad (4)$$

$$T_E = M_E(F_E, \theta_E) \quad (5)$$

式中, M_B 为 B 超模态通道的非线性映射函数, M_E 为超声弹性模态通道的非线性映射函数, θ_B 为非线性映射模块中 B 超模态通道的参数, θ_E 为非线性映射模块中超声弹性模态通道的参数。

1.2.3 联合特征表达

输入的 B 超模态图像和超声弹性模态图像分别经过了特征提取和非线性映射模块,获得对应的高级特征表达。然后,在联合特征表达模块中,利用拼接操作对两个模态通道输出的高级特征进行特征融合,即投影到公共子空间获得联合特征表达,计算式为

$$T = [T_B, T_E] \quad (6)$$

式中, T 为联合特征表达,即拼接操作的结果。

1.2.4 重构

在重构模块中,将联合表征模块输出的特征表达 T_B 和 T_E 输入双通道网络,分别重建出原先作为输入的 B 超与超声弹性图像。重建模块包含两个网络通道,分别由 5 个上采样层和 1 个卷积层组成,每个上采样层结构如图 3 所示。

利用多层上采样结构,两个网络通道分别将联

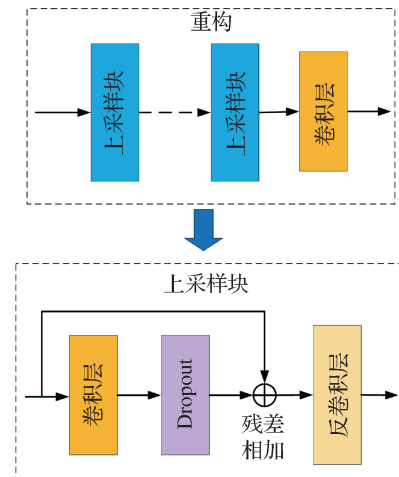


图3 上采样结构

Fig.3 Upsampling structure

合特征表达上采样至原输入图像尺寸,分别用做重建 B 超与超声弹性图像。恢复的 B 超图像记为 B_r , 恢复的超声弹性图像记为 E_r , 计算为

$$B_r = H_B(T, \beta_B) \quad (7)$$

$$E_r = H_E(T, \beta_E) \quad (8)$$

式中, H_B 为 B 超模态通道的重构映射函数, H_E 为超声弹性模态通道的重构映射函数, β_B 为重构模块中 B 超模态通道参数, β_E 为重构模块中超声弹性模态通道的参数。

1.2.5 多损失函数联合训练

在第 1 阶段预训练中,本文算法计算了 3 个损失函数,包括两个模态图像的自重建损失函数和双模态之间的关联损失函数。均方误差(mean square error, MSE)是一种像素级优化损失函数,它是图像重建任务中最常用的损失函数。因此,本文算法采用 MSE 作为自监督重建任务的损失函数。将 B 超模态自重建损失函数记为 $Loss_B$, 超声弹性模态自重建损失函数记为 $Loss_E$, 计算为

$$Loss_B = \frac{1}{2N} \sum_{i=1}^N \|B_{ri} - I_{Bi}\|_2 \quad (9)$$

$$Loss_E = \frac{1}{2N} \sum_{i=1}^N \|E_{ri} - I_{Ei}\|_2 \quad (10)$$

式中, N 为双模态训练样本对总数, i 表示样本对序号。

联合表征模块中 B 超和超声弹性模态之间的关联损失记为 $Loss_C$, 其计算为

$$Loss_C = \frac{\sum_{i=1}^N (h_B(I_B) - \overline{h_B(T_B)})(h_E(I_E) - \overline{h_E(T_E)})}{\sqrt{\sum_{i=1}^N (h_B(I_B) - \overline{h_B(T_B)})^2 \sum_{i=1}^N (h_E(I_E) - \overline{h_E(T_E)})^2}} \quad (11)$$

式中, h_B 为 B 超模态通道的映射函数, h_E 为超声弹性模态通道的映射函数, $\overline{h_B(T_B)}$ 为 B 超模态高级特征表达的平均向量, $\overline{h_E(T_E)}$ 表示超声弹性模态高级特征表达的平均向量, N 为双模态训练样本对总数, i 表示样本对序号。

最终,第 1 阶段的损失函数 $Loss_1$ 由上述 3 个损失函数加权计算得到

$$Loss_1 = Loss_B + Loss_E + \mu Loss_C \quad (12)$$

式中, $\mu \in \mathbf{R}^+$ 为 $Loss_C$ 的权重。

在第 1 阶段预训练中,用总损失函数 $Loss_1$ 来优

化网络参数。其中,第 1 个分量 $Loss_B$ 计算 B 超模态图像自重建的损失,约束 B 超通道网络的参数优化;第 2 个分量 $Loss_E$ 计算超声弹性模态通道的自重建损失,约束超声弹性通道网络的参数优化;第 3 个分量 $\mu Loss_C$ 计算的是联合表征模块中公共子空间的关联损失,约束两个模态的表征相关性。由式(12)可以看出, $Loss_B$ 和 $Loss_E$ 越小,说明两个模态的自重建效果越好, $Loss_C$ 越小,说明两个模态投影到公共子空间的特征表达越相关。

1.3 第 2 阶段多模态分类迁移学习

如图 1 所示,第 2 阶段多模态分类是在第 1 阶段得到的预训练骨干网络基础上,添加新的全连接层以降低特征维度,并利用 Softmax 函数运算得到分类的预测值,然后输入 B 超和超声弹性图像,进行基于双模超声的乳腺癌分类任务。该阶段的网络包含特征提取和非线性映射两个子模块,这两部分的网络结构和网络参数等设置和第 1 阶段相同,故本节不再赘述。与第 1 阶段预训练网络不同的是,第 2 阶段将联合特征表达层输出的高级特征直接进行 Softmax 函数运算,实现共享标签的双模态分类任务。

在第 2 阶段多模态分类任务中,使用常用的交叉熵损失函数对网络进行训练,将其记为 $Loss_2$, 计算为

$$Loss_2 = \sum_{i=1}^N [y_i \log(a_i) + (1 - y_i) \log(1 - a_i)] \quad (13)$$

式中, y_i 表示第 i 对双模态图像样本所共享的分类标签,恶性标签为 1,良性标签为 0, a_i 表示第 i 对样本预测为恶性的概率值, N 为双模态训练样本对总数。

1.4 单通道网络微调

该阶段使用第 2 阶段迁移学习获得的预训练骨干网络(即 B 超对应的单通道网络),并在其后添加几层全连接层以及 Softmax 函数运算,构建基于 B 超单模态的乳腺癌 CAD 模型。该阶段是在第 2 阶段预训练得到的骨干网络模型基础上,使用 B 超单模态数据进行迁移微调,其中预训练模型已经蕴含了经过两阶段迁移得到的超声弹性图像信息。该阶段同样使用交叉熵损失函数,记为 $Loss_3$, 计算为

$$Loss_3 = \sum_{j=1}^N [y_j \log(a_j) + (1 - y_j) \log(1 - a_j)] \quad (14)$$

式中, y_j 表示 B 超模态图像的分类标签, 恶性标签为 1, 良性标签为 0, a_j 表示预测为恶性的概率值。

2 实验和结果

2.1 实验数据与预处理

使用的乳腺肿瘤双模态超声数据集是由南京鼓楼医院采集, 两种模态的数据同时采集自 264 位患者, 其中包括 129 位良性乳腺肿瘤患者和 135 位恶性乳腺肿瘤患者, 已获得医院伦理委员会的批准, 所有患者均签署了知情同意书。乳腺图像是由 Mindray Resona7 超声扫描仪通过 L11-3 探头采集的, 包括 B 超和超声弹性两种模态的医学影像信息, 所有恶性肿瘤均已经过专业医生的病理诊断证实。

医生手动截取了原始数据 B 超和超声弹性图像中的感兴趣区域(region of interest, ROI), 因此本文算法直接在 ROI 区域数据上进行实验验证。图 4 所示分别为良恶性双模态超声图像 ROI 区域的数据示例样本。

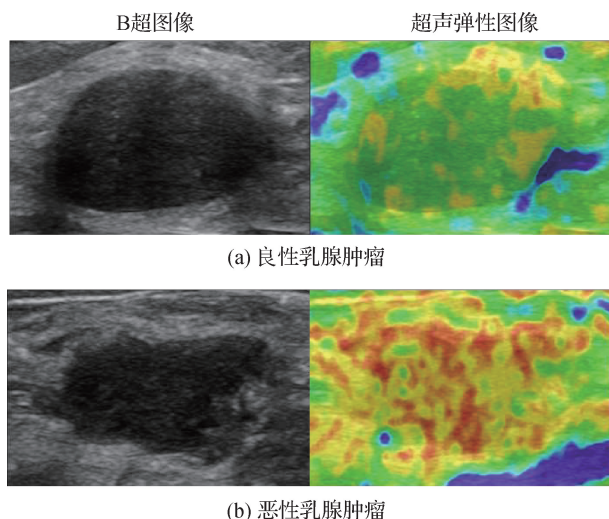


图 4 双模态超声图像 ROI 区域样本

Fig. 4 The samples of the bimodal ultrasound image
((a) benign breast tumor; (b) malignant breast tumor)

2.2 实验设计

为了评估 TSDTL 算法的性能, 实验与多种目前常用的深度迁移学习算法进行了比较。由于 B 超是临床中使用最广泛的肿瘤诊断模态, 且 B 超图像相比超声弹性图像更容易获得, 因此, 实验以 B 超图像作为诊断模态, 超声弹性图像作为源域

模态。

与 TSDTL 对比的算法有:

1) ResNet-18 (He 等, 2016)。对比实验采用一个单通道的 ResNet-18 应用于单模态的 B 超进行乳腺癌良恶性分类, 其结果作为基准算法。

2) 深度域混淆网络 (deep domain confusion, DDC) (Tzeng 等, 2014)。这是一种引入了自适应层和域混淆损失函数的深度迁移网络。本文使用 ResNet-18 作为该算法的骨干网络。

3) 深度自适应相关对齐网络 (correlation alignment for deep domain adaptation, DeepCORAL) (Sun 和 Saenko, 2016)。这是一种将 CORAL 方法 (Sun 等, 2016) 和深度神经网络相结合的深度迁移网络。本文使用 ResNet-18 作为该算法的骨干网络。

4) 深度自适应网络 (deep adaptation network, DAN) (Long 等, 2015)。这是一种将多核最大均值差异度量应用于多层域自适应网络中的深度迁移网络。本文使用 ResNet-18 作为该算法的骨干网络。

5) 动态对抗自适应网络 (dynamic adversarial adaptation network, DAAN) (Yu 等, 2019)。这是一种使用对抗模型来动态地学习域不变表达的深度迁移学习网络。原始论文中使用 ResNet-50 作为基础网络, 本文将其替换为 ResNet-18 网络。

6) 深度子域自适应网络 (deep subdomain adaptation network, DSAN) (Zhu 等, 2021)。这是一种基于局部最大均值差异的深度迁移网络。原始论文中使用 ResNet-50 作为基础网络, 本文将其替换为 ResNet-18 网络。

在上述实验的基础上, 为了进一步验证本文提出的 TSDTL 算法的有效性, 设计了两个只包含一阶段迁移学习的消融实验:

1) SSL-TL。在这个消融实验中, 只保留第 1 阶段基于自监督学习的迁移学习 (self-supervised learning based transfer learning, SSL-TL), 而将基于双模超声的分类任务的第 2 阶段迁移学习过程删除, 这是为了验证第 2 阶段迁移学习的有效性。在实际训练过程中, 直接将如图 1 中第 1 阶段网络训练以后得到的 B 超对应的单通道编码器网络部分直接应用到单通道网络微调阶段, 后续步骤不变。该阶段算法记为 SSL-TL。

2) 基于双模分类的迁移学习 (dual-modal classi-

fication transfer learning, DMC-TL)。在这个消融实验中,只保留基于双模超声的乳腺癌分类任务第2阶段迁移学习,而将基于自监督学习的第1阶段迁移学习过程删除,这是为了验证第1阶段迁移学习的有效性。在实际训练过程中,跳过如图1中所示的第1阶段图像重建任务,直接进行第2阶段的基于双模超声的乳腺癌分类任务,后续步骤不变。该阶段算法记为DMC-TL。

为了对比不同算法针对B超图像的分类效果,采用5折交叉验证方法进行实验,结果均以5次结果的平均±标准差的形式呈现。最终结果使用了如下指标:分类准确率(accuracy, ACC)、敏感度(sensitivity, SEN)、特异度(specificity, SPE)和约登指数(Youden index, YI)。计算式为

$$\begin{cases} ACC = (TP + TN)/(TP + FN + TN + FP) \\ SEN = TP/(TP + FN) \\ SPE = TN/(TN + FP) \\ YI = SEN + SPE - 1 \end{cases}$$

(15)

式中,TP为真阳性,FP为假阳性,TN为真阴性,FN为假阴性。ACC能够直观反映算法的性能;SEN反映了对该疾病的敏感程度和漏诊程度;SPE反映了误诊程度;YI综合考虑SEN和SPE的结果,评判了算法的真实性,该值越大,说明分类效果越好,真实性越大。利用真正例率和假正例率描述两者之间的受试者工作特征曲线(receiver operating characteristic curve, ROC),并计算对应曲线下面积(area under

ROC curve, AUC)作为评价指标。AUC值越大,表明该模型的分类性能越强。

此外,TSDTL算法第1阶段的总损失由3个损失函数加权得到。其中,分量 $Loss_c$ 为联合表征模块中公共子空间的关联损失。为了确定其权重 μ ,对该超参数的不同取值进行了对比实验。

对公共子空间关联损失的应用位置进行了消融实验。分别在TSDTL算法的第1、第2阶段以及两阶段加入关联损失,从而验证本文的TSDTL算法仅在第1阶段使用该损失的有效性和合理性。

2.3 算法实现细节

实验中所有对比算法都是基于图像ROI进行训练和测试的,图像的大小统一为224×224像素,并做归一化操作。使用深度学习Pytorch框架构建模型并完成实验,所有模型的参数初始化采用Msra算法(He等,2015),使用Adam算法优化参数,学习率设为0.001,批量大小设置为16。另外,为了对比公平,对所有算法的网络层数和其余各项超参数的设置也尽可能相同。

2.4 实验结果

表1为不同分类算法实现乳腺肿瘤分类诊断的结果。

由表1可以发现,本文提出的TSDTL算法获得了整体最优的分类结果,其对应的平均分类准确率、敏感度、特异度和约登指数结果分别为87.84±2.08%、88.89±3.70%、86.71±2.21%和75.60±

表1 不同算法在乳腺肿瘤超声数据集的分类结果
Table 1 Classification results of different algorithms on breast tumor ultrasound dataset

算法	准确率	敏感度	特异度	约登指数
ResNet-18(He等,2016)	84.10±2.81	85.19±2.62	82.99±6.34	68.17±5.73
DDC(Tzeng等,2014)	85.61±2.12	86.89±2.41	84.22±3.31	71.11±4.20
DeepCORAL(Sun和Saenko,2016)	86.00±3.39	86.15±6.23	85.72±6.13	71.87±6.61
DAN(Long等,2015)	86.37±1.50	87.41±6.20	85.26±6.90	72.67±2.99
DAAN(Yu等,2019)	86.73±2.37	87.41±4.22	86.06±1.98	73.47±4.66
DSAN(Zhu等,2021)	86.74±3.03	87.40±8.92	85.97±7.79	73.38±6.06
TSDTL(本文)	87.84±2.08	88.89±3.70	86.71±2.21	75.60±4.07

注:加粗字体为每列最优值。

4.07%。与直接采用 ResNet-18 进行单模态 B 超分类的结果相比, TSDTL 算法在分类准确率、敏感度、特异度和约登指数上分别提高了 3.74%、3.70%、3.72% 和 7.43%, 表明了 TSDTL 算法的有效性; 与其他深度迁移学习算法相比, 本文提出的 TSDTL 算法在分类准确率、敏感度、特异度和约登指数上至少提高了 1.10%、1.48%、0.65% 和 2.13%。表明了 TSDTL 算法能更充分地将超声弹性图像中补充性信息迁移至基于 B 超的乳腺癌 CAD 模型中。

图 5 所示为在乳腺肿瘤超声数据集上不同算法的 ROC 曲线, 图中标注了各个算法的 AUC 值, 由图 5 可知, 本文算法对应 ROC 下面积最大, 即 AUC 值最高, 为 0.918, 表明了 TSDTL 算法的有效性。

表 2 为消融实验的结果。由表 2 可以发现, TSDTL 算法虽然在敏感度指标上的结果比 DMC-TL 算法降低了 2.22%, 但其在分类准确率、特异度和约登指数上的结果均分别比 DMC-TL 算法提升了 1.08%、4.51% 和 2.29%, TSDTL 算法整体性能优于 DMC-TL 算法, 表明了第 1 阶段自监督学习在信息迁移中的有效性。TSDTL 算法虽然在敏感度指标上的结果比 SSL-TL 算法降低了 0.74%, 但其在

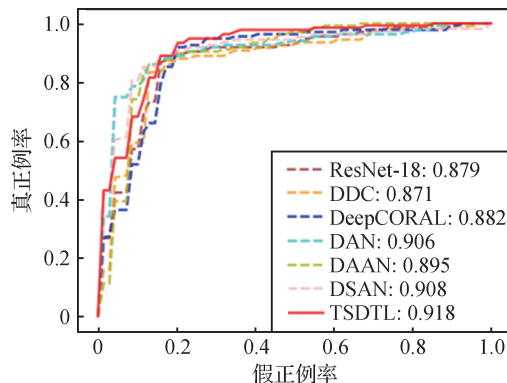


图 5 不同算法在对比实验中的 ROC 曲线及其 AUC 值

Fig. 5 ROC curves and values of different algorithms in the comparison experiment

分类准确率、特异度和约登指数上的结果均分别比 SSL-TL 算法提升了 1.84%、4.49% 和 3.75%, TSDTL 算法整体性能优于 SSL-TL 算法, 表明了第 2 阶段多模态分类对最终超声 CAD 模型分类性能的有效性。此外, SSL-TL 算法和 DMC-TL 算法分类准确率、敏感度、特异度和约登指数上的结果都优于 ResNet-18 算法, 再次验证了第 1 阶段自监督学习和第 2 阶段多模态分类 LUPI 对模型性能的有效性。

表 2 不同算法在消融实验中的分类结果

Table 2 Classification results of different algorithms in ablation experiment

算法	准确率	敏感度	特异度	约登指数
ResNet-18(He 等, 2016)	84.10 ± 2.81	85.19 ± 2.62	82.99 ± 6.34	68.17 ± 5.73
SSL-TL	86.00 ± 3.37	89.63 ± 5.49	82.22 ± 4.23	71.85 ± 6.67
DMC-TL	86.76 ± 3.03	91.11 ± 3.31	82.20 ± 5.72	73.31 ± 5.90
TSDTL(本文)	87.84 ± 2.08	88.89 ± 3.70	86.71 ± 2.21	75.60 ± 4.07

注: 加粗字体为每列最优值。

图 6 所示为在乳腺肿瘤超声数据集上不同算法的 ROC 曲线以及对应的 AUC 值。由图 6 可知, TSDTL 算法的 AUC 值高于 DMC-TL 算法, 验证了第 1 阶段自监督深度迁移学习的有效性。此外, TSDTL 算法的 AUC 值相较于 SSL-TL 算法同样提升明显, 表明了第 2 阶段多模态分类 LUPI 的有效性。

图 7 展示了在第 1 阶段的训练中损失函数使用不同 μ 取值时, TSDTL 算法分类结果的变化。由图 7 可知, $\mu = 10^{-4}$ 时, TSDTL 算法能达到最优的综合分类性能。此外, 值得注意的是, 为了达到较好的训练效果, μ 的数量级较小, 这是因为在第 1 阶段的

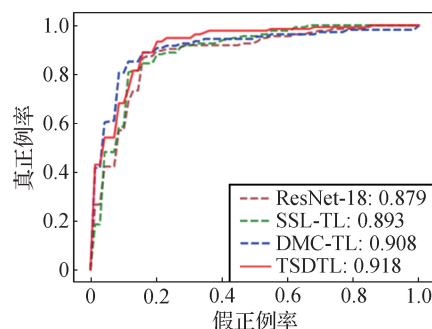


图 6 不同算法在消融实验中的 ROC 曲线及其 AUC 值

Fig. 6 ROC curves and values of different algorithms in the ablation experiment

3 个损失分量中,关联损失相比另外两个图像重建损失较大。因此,为了平衡这两类损失,超参数 μ 在第 1 阶段的训练中取值较小。

表 3 所示为 TSDTL 算法分别在不同的训练阶段使用关联损失的分类结果。为了表述方便,使用 TSDTL-Stage1 来表示仅在第 1 阶段使用关联损失;TSDTL-Stage2 表示仅在第 2 阶段使用关联损失;TSDTL-Stage1&2 表示在两阶段都使用关联损失。由该消融实验结果可知,TSDTL 算法仅在第 1 阶段使用该损失的方式在各项指标上取得了最优性能,验证了该训练策略的有效性。

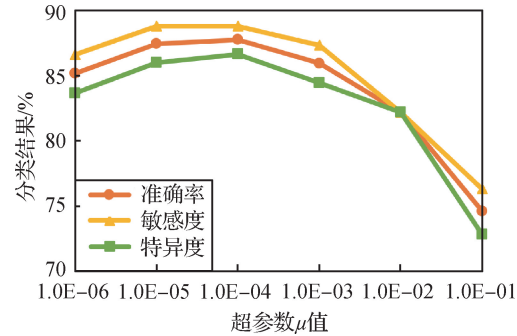


图 7 TSDTL 算法在使用不同 μ 值时的分类结果

Fig. 7 Classification results of TSDTL when different hyperparameter μ were used

表 3 在 TSDTL 不同阶段中使用关联损失的分类结果

Table 3 Classification results of TSDEL when the correlation loss was used in different stages

算法	准确率	敏感度	特异度	约登指数
TSDTL-Stage1	87.84 ± 2.08	88.89 ± 3.70	86.71 ± 2.21	75.60 ± 4.07
TSDTL-Stage2	85.25 ± 3.56	86.67 ± 2.03	83.78 ± 6.22	70.45 ± 7.24
TSDTL-Stage1&2	86.38 ± 3.58	88.15 ± 3.10	84.55 ± 5.99	72.70 ± 7.24

注:加粗字体为每列最优值。

在 TSDTL 的第 1 阶段预训练中,算法采用的是一种双模态图像重建的自监督学习任务。加入关联损失的目的是为了在双模态数据在分别自主编码过程中形成一种约束机制,提升两者特征表达的相关性,从而避免模型在训练过程中只通过单一模态(如超声弹性模态)完成重建双模图像的任务。而在训练第 2 阶段的分类任务时,已有预训练模型和类别标签,不存在第 1 阶段的问题,使用冗余的损失反而可能损伤其性能。因此,只在第 1 阶段的训练中使用该关联损失是 TSDTL 的最优训练策略。

3 结 论

提出了一种基于两阶段深度迁移学习的乳腺超声 CAD 算法,有效实现了将超声弹性图像所内涵的生物力学相关的信息迁移至基于 B 超的乳腺癌 CAD 模型之中,从而提升 CAD 模型的分类诊断性能。该算法创新性地结合了基于双模态超声图像的自监督学习任务 and 分类任务,分别在两阶段实现了隐式的特征级深度迁移学习,在有效缓解样本较少问题的同时,对乳腺超声 CAD 模型的信息完整性提

供了支持。其中,算法在第 1 阶段构建了一个双模态超声图像的自监督重建任务,并通过相关性损失函数实现了隐式的特征级迁移学习;第 2 阶段则是利用共享标签的特性,以隐式 LUPI 的范式加强两个通道骨干网络学习到特征之间的信息交互。TSDTL 分类算法在乳腺肿瘤双模态超声数据集上进行了实验。相较于其他深度迁移学习算法,本文算法在各项分类评价指标上取得了最佳性能,表明了提出的两阶段深度迁移学习能够使 B 超模态和超声弹性模态之间的信息进行充分的交互迁移。此外,消融实验的结果也分别验证了第 1 阶段自监督学习和第 2 阶段多模态分类对提升网络分类性能的有效性。

本文算法仍有待进一步改进。在 ROC 曲线图中,本文算法虽然均取得了最佳的 AUC 值,但相较于其他算法,对应的 ROC 曲线没有明显的优势。而消融实验的结果表明,第 1 阶段的自监督学习对整体性能有效,但仍然有提升空间。在未来的研究工作中,将探索自监督学习领域的另一个前沿分支,即对比学习在第 1 阶段迁移学习中的应用效果。基于共享标签的双模态超声数据,通过优化对比损失使

网络学习两者之间的内在关联性,从而进一步增强CAD模型的分类性能。

参考文献 (References)

- Bhatt G, Jha P and Raman B. 2019. Representation learning using step-based deep multi-modal autoencoders. *Pattern Recognition*, 95: 12-23 [DOI: 10.1016/j.patcog.2019.05.032]
- Carlsen J F, Ewertsen C, Lönn L and Nielsen M B. 2013. Strain elastography ultrasound; an overview with emphasis on breast cancer diagnosis. *Diagnostics*, 3 (1): 117-125 [DOI: 10.3390/diagnostics3010117]
- Chen L, Bentley P, Mori K, Misawa K, Fujiwara M and Rueckert D. 2019. Self-supervised learning for medical image analysis using image context restoration. *Medical Image Analysis*, 58: #101539 [DOI: 10.1016/j.media.2019.101539]
- Du Z J, Gong X, Luo J, Zhang Z M and Yang F. 2020. Classification method for samples that are easy to be confused in breast ultrasound images. *Journal of Image and Graphics*, 25(7): 1490-1500 (杜章锦, 龚勋, 罗俊, 章哲敏, 杨菲. 2020. 乳腺超声图像中易混淆困难样本的分类方法. *中国图象图形学报*, 25(7): 1490-1500) [DOI: 10.11834/jig.190442]
- Duan L X, Xu Y W, Li W, Chen L, Wong D W K, Wong T Y and Liu J. 2014. Incorporating privileged genetic information for fundus image based glaucoma detection//*Proceedings of the 17th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Boston, USA; Springer: 204-211 [DOI: 10.1007/978-3-319-10470-6_26]
- He K M, Zhang X Y, Ren S Q and Sun J. 2015. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification//*Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago, Chile; IEEE: 1026-1034 [DOI: 10.1109/ICCV.2015.123]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA; IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Jing L L and Tian Y L. 2021. Self-supervised visual feature learning with deep neural networks; a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43 (11): 4037-4058 [DOI: 10.1109/TPAMI.2020.2992393]
- Li X, Du B, Xu C, Zhang Y P, Zhang L F and Tao D C. 2018. R-SVM+: robust learning with privileged information//*Proceedings of the 27th International Joint Conference on Artificial Intelligence*. Stockholm, Sweden; IJCAI: 2411-2417 [DOI: 10.24963/ijcai.2018/334]
- Liu X, Zhang F J, Hou Z Y, Mian L, Wang Z Y, Zhang J and Tang J. 2020. Self-supervised learning: generative or contrastive [EB/OL]. [2021-07-23]. <https://arxiv.org/pdf/2006.08218.pdf>
- Long M S, Cao Y, Wang J M and Jordan M I. 2015. Learning transferable features with deep adaptation networks//*Proceedings of the 32nd International Conference on Machine Learning*. Lille, France; JMLR.org: 97-105
- Mahmood T, Li J Q, Pei Y, Akhtar F, Imran A and Rehman K U. 2020. A brief survey on breast cancer diagnostic with deep learning schemes using multi-image modalities. *IEEE Access*, 8: 165779-165809 [DOI: 10.1109/ACCESS.2020.3021343]
- Pechyony D and Vapnik V. 2012. Fast optimization algorithms for solving SVM+//*Statistical Learning and Data Science*. London; Chapman and Hall/CRC: 27-42 [DOI: 10.1201/b11429-7]
- Sahiner B, Chan H P and Hadjiiski L. 2008. Classifier performance estimation under the constraint of a finite sample size: resampling schemes applied to neural network classifiers. *Neural Networks*, 21(2/3): 476-483 [DOI: 10.1016/j.neunet.2007.12.012]
- Shi X Q, Li J, Qian L, Xue X, Li J and Wan W. 2018. Correlation between elastic parameters and collagen fibre features in breast lesions. *Clinical Radiology*, 73 (6): 595.e1-595.e7 [DOI: 10.1016/j.crad.2018.01.019]
- Siegel R L, Miller K D and Jemal A. 2020. Cancer statistics, 2020. *CA: A Cancer Journal for Clinicians*, 70 (1): 7-30 [DOI: 10.3322/caac.21590]
- Sigrist R M S, Liao J, Kaffas A E, Chammas M C and Willmann J K. 2017. Ultrasound elastography: review of techniques and clinical applications. *Theranostics*, 7 (5): 1303-1329 [DOI: 10.7150/thno.18650]
- Sun B C, Feng J S and Saenko K. 2016. Return of frustratingly easy domain adaptation//*Proceedings of the 30th AAAI Conference on Artificial Intelligence*. Phoenix, Arizona, USA; AAAI Press: 2058-2065
- Sun B C and Saenko K. 2016. Deep CORAL: correlation alignment for deep domain adaptation//*Proceedings of European Conference on Computer Vision*. Amsterdam, the Netherlands; Springer: 443-450 [DOI: 10.1007/978-3-319-49409-8_35]
- Sutton G C. 1989. Computer-aided diagnosis: a review. *British Journal of Surgery*, 76(1): 82-85 [DOI: 10.1002/bjs.1800760126]
- Tzeng E, Hoffman J, Zhang N, Saenko K and Darrell T. 2014. Deep domain confusion: maximizing for domain invariance [EB/OL]. [2021-07-23]. <https://arxiv.org/pdf/1412.3474.pdf>
- Vapnik V and Vashist A. 2009. A new learning paradigm: learning using privileged information. *Neural Networks*, 22 (5/6): 544-557 [DOI: 10.1016/j.neunet.2009.06.042]
- Wang Y, Chen H D, Li N, Ren J S, Zhang K, Dai M and He J. 2019. Ultrasound for breast cancer screening in high-risk women: results from a population-based cancer screening program in China. *Frontiers in Oncology*, 9: #286 [DOI: 10.3389/fonc.2019.00286]
- Weiss K, Khoshgoftaar T M and Wang D D. 2016. A survey of transfer learning. *Journal of Big Data*, 3(1): #9 [DOI: 10.1186/s40537-

016-0043-6]

Yu C H, Wang J D, Chen Y Q and Huang M Y. 2019. Transfer learning with dynamic adversarial adaptation network//Proceedings of 2019 IEEE International Conference on Data Mining. Beijing, China; IEEE: 778-786 [DOI: 10.1109/ICDM.2019.00088]

Yu Y H, Lin H F, Meng J N, Wei X C, Guo H and Zhao Z H. 2017. Deep transfer learning for modality classification of medical images. Information, 8(3): #91 [DOI: 10.3390/info8030091]

Zhang H L, Guo L H, Wang D, Wang J, Bao L L, Ying S H, Xu H X and Shi J. 2021. Multi-source transfer learning via multi-kernel support vector machine plus for B-mode ultrasound-based computer-aided diagnosis of liver cancers. IEEE Journal of Biomedical and Health Informatics, 25(10): 3874-3885 [DOI: 10.1109/JBHI.2021.3073812]

Zhu Y C, Zhuang F Z, Wang J D, Ke G L, Chen J W, Bian J, Xiong H and He Q. 2021. Deep subdomain adaptation network for image classification. IEEE Transactions on Neural Networks and Learning Systems, 32(4): 1713-1722 [DOI: 10.1109/TNNLS.2020.2988928]

作者简介



贡荣麟, 1996年生, 男, 硕士研究生, 主要研究方向为深度学习、医学图像分析与处理。

E-mail: gongronglin@shu.edu.cn



施俊, 通信作者, 男, 教授, 主要研究方向为医学图像分析与处理、模式识别。

E-mail: junshi@shu.edu.cn

周玮珺, 女, 主治医师, 主要研究方向为超声影像分析。

E-mail: wyky1985@163.com

汪程, 男, 硕士研究生, 主要研究方向为深度学习、医学图像分析与处理。E-mail: wc1995@shu.edu.cn