



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所

中国图象图形学学会

北京应用物理与计算数学研究所

中国图象 图形学报

2020

08

VOL.25

ISSN1006-8961

CN11-3758/TB



三维建模和步态识别 P1539



第25卷第8期（总第292期）

2020年8月16日

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

版权声明

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

Copyright

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

主管单位 中国科学院

主办单位 中国科学院遥感与数字地球研究所

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 顾行发

编辑出版《中国图象图形学报》编辑出版委员会

通信地址 北京市海淀区北四环西路19号

邮 编 100190

电子信箱 jig@radi.ac.cn

电 话 010-58887035

网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司

（邮政信箱：北京399信箱 邮编：100048）

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing and Digital Earth, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Gu Xingfa

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

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

Zip code 100190

E-mail jig@radi.ac.cn

Telephone 010-58887035

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

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

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

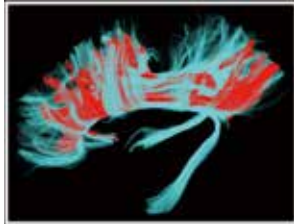
ISSN 1006-8961

CODEN ZTTXFZ

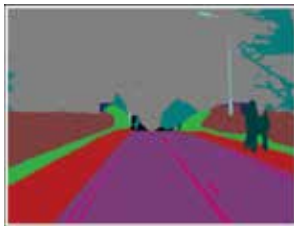
国外发行代号 M1406

国内邮发代号 82-831

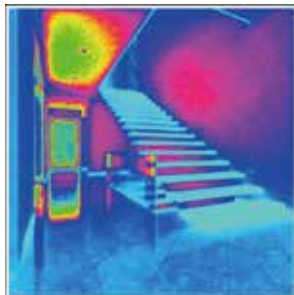
国内定价 60.00元



神经纤维跟踪算法研究进展
(第1513页)



结合注意力机制的双路径语义分割(第1627页)



突变策略下多通路Metropolis光照传播(第1658页)

综述

神经纤维跟踪算法研究进展

李茂, 何建忠, 冯远静 1513

多模态生物特征提取及相关性评价综述

杨雪鹤, 刘欢喜, 肖建力 1529

图像分析和识别

异常步态3维人体建模和可变视角识别

罗坚, 黎梦霞, 罗诗光 1539

扩展点态卷积网络的点云分类分割模型

张新良, 付陈琳, 赵运基 1551

特征一致性约束的视频目标分割

征煜, 陈亚当, 郝川艳 1558

增量角度域损失和多特征融合的地标识别

毛雪宇, 彭艳兵 1567

部件检测和语义网络的细粒度鞋类图像检索

陈前, 刘骊, 付晓东, 刘利军, 黄青松 1578

图像理解和计算机视觉

图注意力网络的场景图到图像生成模型

兰红, 刘秦邑 1591

跨层多模型特征融合与因果卷积解码的图像描述

罗会兰, 岳亮亮 1604

Haar-like特征双阈值Adaboost人脸检测

刘禹欣, 朱勇, 孙结冰, 王一博 1618

结合注意力机制的双路径语义分割

翟鹏博, 杨浩, 宋婷婷, 余亢, 马龙祥, 黄向生 1627

自学习规则下的多聚焦图像融合

刘子闻, 罗晓清, 张战成 1637

车路视觉协同的高速公路防碰撞预警算法

蔡创新, 高尚兵, 周君, 黄子赫 1649

计算机图形学

突变策略下多通路Metropolis光照传播

贺怀清, 赵煜桢, 刘浩翰, 王旭 1658

融合DNN与AABB-圆形包围盒自碰撞检测

靳雁霞, 程琦甫, 张晋瑞, 齐欣, 马博, 贾瑶 1674

医学图像处理

压缩感知下溃疡性结肠炎辅助诊断

杨海清, 孙道洋 1684

特征重用和注意力机制下肝肿瘤自动分类

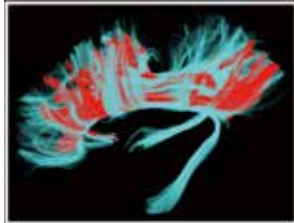
冯诺, 宋余庆, 刘哲 1695

融合注意力机制和高效网络的糖尿病视网膜病变识别与分类

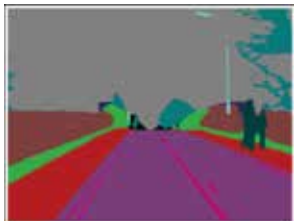
张子振, 刘明, 朱德江 1708

CONTENTS

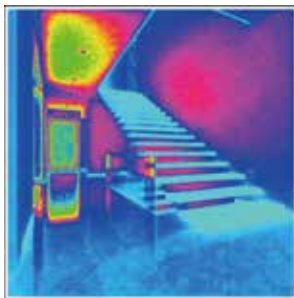
JOURNAL OF IMAGE AND GRAPHICS



Research progress of neural fiber tracking(P1513)



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



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

Review

Research progress of neural fiber tracking

Li Mao, He Jianzhong, Feng Yuanjing 1513

Extraction and relevance evaluation for multimodal biometric features

Yang Xuehe, Liu Huanxi, Xiao Jianli 1529

Image Analysis and Recognition

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

Luo Jian, Li Mengxia, Luo Shiguang 1539

Extended pointwise convolution network model for point cloud classification and segmentation

Zhang Xinliang, Fu Chenlin, Zhao Yunji 1551

Video object segmentation algorithm based on consistent features

Zheng Yu, Chen Yadang, Hao Chuanyan 1558

Landmark recognition based on ArcFace loss and multiple feature fusion

Mao Xueyu, Peng Yanbing 1567

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

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

Image Understanding and Computer Vision

Image generation from scene graph with graph attention network

Lan Hong, Liu Qinyi 1591

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

Luo Huilan, Yue Liangliang 1604

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

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

Two-path semantic segmentation algorithm combining attention mechanism

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

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

Liu Ziwen, Luo Xiaoqing, Zhang Zhancheng 1637

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

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

Computer Graphics

Improved multiplexed Metropolis light transport based on fusion mutation strategy

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

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

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

Medical Image Processing

Adjunctive diagnosis of ulcerative colitis under compressed sensing

Yang Haiqing, Sun Daoyang 1684

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

Feng Nuo, Song Yuqing, Liu Zhe 1695

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

Zhang Zizhen, Liu Ming, Zhu Dejiang 1708

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2020)08-1695-13

论文引用格式: Feng N, Song Y Q and Liu Z. 2020. Automatic classification of liver tumors by combining feature reuse and attention mechanism. Journal of Image and Graphics, 25(08): 1695-1707 (冯诺, 宋余庆, 刘哲. 2020. 特征重用和注意力机制下肝肿瘤自动分类. 中国图象图形学报, 25(08): 1695-1707) [DOI: 10.11834/jig.190634]

特征重用和注意力机制下肝肿瘤自动分类

冯诺, 宋余庆, 刘哲

江苏大学计算机科学与通信工程学院, 镇江 212013

摘要: **目的** 肝肿瘤分类计算机辅助诊断技术在临床医学中具有重要意义, 但样本缺乏、标注成本高及肝脏图像的敏感性等原因, 限制了深度学习的分类潜能, 使得肝肿瘤分类依然是医学图像处理领域中具有挑战性的任务。针对上述问题, 本文提出了一种结合特征重用和注意力机制的肝肿瘤自动分类方法。**方法** 利用特征重用模块对计算机断层扫描(computed tomography, CT)图像进行伪自然图像的预处理, 复制经 Hounsfield 处理后的原通道信息, 并通过数据增强扩充现有数据; 引入基于注意力机制的特征提取模块, 从全局和局部两个方面分别对原始数据进行加权处理, 充分挖掘现有样本的高维语义特征; 通过迁移学习的训练策略训练提出的网络模型, 并使用 Softmax 分类器实现肝肿瘤的精准分类。**结果** 在 120 个病人的 514 幅 CT 扫描切片上进行了综合实验。与基准方法相比, 本文方法平均分类准确率为 87.78%, 提高了 9.73%; 与肝肿瘤分类算法相比, 本文算法针对转移性肝腺瘤、血管瘤、肝细胞癌及正常肝组织的分类召回率分别达到 79.47%、79.67%、85.73% 和 98.31%; 与主流分类模型相比, 本文模型在多种评价指标中均表现优异, 平均准确率、召回率、精确率、F1-score 及 AUC(area under ROC curve)分别为 87.78%、84.43%、84.59%、84.44% 和 97.50%。消融实验表明了本文设计的有效性。**结论** 本文方法能提高肝脏肿瘤的分类结果, 可为临床诊断提供依据。

关键词: 深度学习; 肝肿瘤分类; 注意力机制; 特征重用; 特征提取

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

Feng Nuo, Song Yuqing, Liu Zhe

School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang 212013, China

Abstract: **Objective** Liver, which is the largest organ in the abdomen, plays a vital role in the metabolism of the human body. Early detection, accurate diagnosis, and further treatment of liver disease are important and helpful to increase the chances for survival. Computed tomography (CT) is an effective tool to detect focal liver lesions due to its robust and accurate imaging techniques. Multiphase CT scans are generally divided into four phases, namely, noncontrast, arterial, portal, and delay. Radiological practice mainly relies on clinicians to analyze the liver. Physicians need to look back or forward in different phases. This task is time and energy consuming. Liver lesion detection also depends on experienced professional physicians. The identification and diagnosis of different liver lesions are challenging tasks for inexperienced doctors due to the similarity among CT images. Therefore, an effective computer-aided diagnosis (CAD) method for doctors should be

收稿日期: 2019-12-05; 修回日期: 2020-01-19; 预印本日期: 2020-01-26

基金项目: 国家自然科学基金项目(61976106, 61772242, 61572239); 中国博士后科学基金项目(2017M611737); 江苏省“六大人才高峰”高层次人才项目(DZXX-122); 镇江市卫生计生科技重点项目(SHW2017019)

Supported by: National Natural Science Foundation of China (61976106, 61772242, 61572239)

designed and developed. Most existing research methods based on CAD are mainly deep learning. Convolutional neural network in deep learning is a data-driven approach, which means it requires much training data to make a model learn the good features for a specific classification. However, a large-scale and well-annotated dataset is extremely difficult to construct due to the lack of data and the cost of labeling data. In accordance with research and analysis, the current methods for liver lesion classification can be divided into two categories, namely, methods based on data and features. The former focuses on expanding data to increase data diversity. The latter mainly studies the way to modify a network to improve the classification ability. These methods have two major drawbacks. First, appearance invariance cannot be controlled when new samples of a specific class are generated. Second, the feature extraction of lesion region cannot be enhanced adaptively. **Method** To solve the above-mentioned problems and improve classification performance, this study proposes a novel method for liver tumor classification by combining feature reuse and attention mechanism. Our contributions are threefold. First, we design a feature reuse module to preprocess medical images. We limit the image intensity values of all CT scans to the range of $[-100, 400]$ Hounsfield unit to eliminate the influences of unrelated tissues or organs on the classification of liver lesions in CT images. A new spatial dimension is added to the 2D pixel matrix, and we concatenate three times for the pixel matrix along the new dimension to generate an efficient feature map with a pseudo-RGB channel. We conduct data augmentation of natural images (such as cut, flip, and fill) to expand medical images and increase their diversity. The feature reuse module not only can enhance the overall representation of original image features but also can effectively avoid the overfitting problem caused by a small sample. Second, we introduce the feature extraction module from two aspects, namely, local and global feature extraction. The local feature extraction block is a pixel-to-pixel modeling, which can enhance the extraction of lesion features by generating a weight factor for each pixel. This block mainly includes two branches, namely, trunk and weight. We feed any given feature map into two group convolution layers in the trunk branch to extract deep and high-dimensional features. It is also fed into the weight branch of encoder-decoder to generate a coefficient factor for each pixel. Lesion features are acquired adaptively by weighting the two branches. The global feature extraction block focuses on the relationship among channels. It can generate a weighting factor by pooling spatially to selectively recalibrate the importance of each feature channel. The ways of local and global feature extraction blocks are processed in parallel to fully mine semantic information with data. Third, we use the training strategy of transfer learning to train the proposed classification model. During training, we transfer the same network layer parameters of SENet (squeeze-and-excitation networks) as those of the proposed network model. **Result** We perform comprehensive experiments on 514 CT slices from 120 patients to evaluate the proposed method thoroughly. The average classification accuracy of the proposed method is 87.78%, which shows an improvement of 9.73% over the baseline model (SENet34). The classification recall rates of metastasis, hemangioma, hepatocellular carcinoma, and healthy liver tissues are 79.47%, 79.67%, 85.73%, and 98.31%, respectively, by using the algorithm in this paper. Compared with the current mainstream classification models, DensNet, SENet, SE_Resnext, CBAM, and SKNet, the proposed model is excellent in many evaluation indicators. The average accuracy, recall rate, precision, F1-score, and area under ROC curve (AUC) are 87.78%, 84.43%, 84.59%, 84.44%, and 97.50%, respectively, by utilizing the proposed architecture. The ablation experiment proves the effectiveness of the proposed design. **Conclusion** The feature reuse module can preprocess medical images to alleviate the overfitting problems caused by lack of data. The feature extraction module based on attention mechanism can mine data features effectively. Our proposed method significantly improves the results of the classification of liver tumors in CT images. Thus, it provides a reliable basis for clinical diagnosis and makes computer-assisted diagnosis possible.

Key words: deep learning; liver lesion classification; attention mechanism; feature reuse; feature extraction

0 引言

世界卫生组织调查显示,肝癌发病率与死亡率位居世界第4和第2位(Bray等,2018)。医学影像

分析对于肝病的临床诊断、病理分析和治疗决策等方面具有重要意义。计算机断层扫描(computed tomography, CT)、核磁共振成像(magnetic resonance imaging, MRI)和肝脏穿刺活检技术(liver biopsy, LB)是临床医学分析、确诊肝肿瘤的主要方法。其

中, LB 不仅实施难度较大, 而且会对肝脏造成一定的损伤。与 MRI 相比, CT 因其鲁棒性和高分辨率的成像技术广泛应用于临床肝病分析, 多期的 CT 扫描一般分为实质期、动脉期、静脉期和延迟期 (Liang 等, 2018a)。传统人工分析往往需要反复对比各个时期、各幅 CT 切片, 费时费力, 而且严重依赖医生的主观判断。此外, 对于早期的、外观相似的以及小的病理特征 (如图 1 所示, 黄色为肝脏位置, 红色为病变位置), 医生很难做出正确判断, 从而错过最佳治疗期。因此, 对计算机辅助无创诊断技术的研究具有很大价值。

基于深度卷积神经网络 (deep convolutional neural network, DCNN) 提取的高维特征明显优于基于传统手工提取到的低维或中维特征 (He 等, 2016; Litjens 等, 2017; Krizhevsky 等, 2017)。现有的深度学习肝肿瘤分类方法分为基于数据扩充的肝肿瘤分类方法和基于特征提取的肝肿瘤分类方法。

基于数据扩充的肝肿瘤分类方法是指运用数据增强 (data augmentation)、生成对抗网络 (generative adversarial network, GAN) (Frid-Adar 等, 2018a, 2018b)、patch (Liang 等, 2018a) 等方法增加数据的多样性。Frid-Adar 等人 (2018a) 利用 GAN 合成新的病变区域, 在 182 个 3 种类型的病灶上取得了平均 85.7% 的灵敏度和 92.4% 的特异度。Ben-Cohen 等人 (2018b) 将不同类别的图像混合训练, 生成具有不同对比度和纹理特征的新样本, 在 239 幅四分

类的 CT 图像上获得了平均 75% 的准确率, 比原实验提高了 7.4%。Liang 等人 (2018a) 基于 patch 的方法, 将整个病变区域和局部病变区域作为模型输入, 并在 480 幅四分类的 CT 切片上取得了 87.58% 的准确率。虽然上述方法在一定程度上可以缓解小样本问题, 但是仍然存在以下问题: 1) 训练波动性较大, 新合成的未知病变需要医生再次进行标注。2) 网络很难学习到不同病灶的分类特性。

基于特征提取的肝肿瘤分类方法是指运用深度神经网络提取样本的深层抽象特征。Liang 等人 (2018b) 在原始方法中 (Liang 等, 2018a) 添加了一个双向长短记忆模型 (bi-directional long short term memory, BDLSTM), 从序列数据中提取有用信息, 刷新了原始分类精度。Romero 等人 (2019) 提出将 InceptionV3 (Szegedy 等, 2016) 与 InceptionResNet-V2 (Szegedy 等, 2017) 相结合, 进行有效的特征提取, 并迁移了 ImageNet 预训练的权重, 最终在 230 幅二分类的肝脏病变图像上取得了平均 96% 的准确率和 92% 的 F1-score。Yasaka 等人 (2017) 利用 CNN 在 460 个病人的 1 068 幅五分类的图像上提取病灶特征, 最终获得了 84% 的精度。Ben-Cohen 等人 (2018a) 使用全卷积 (fully convolutional networks, FCN) 上下文网络和基于超像素局部 patch 的方法对肝损害进行分类与检测。以上方法只能进行整体特征的提取, 不能自适应地关注病变区域, 缓解背景及噪声的影响。

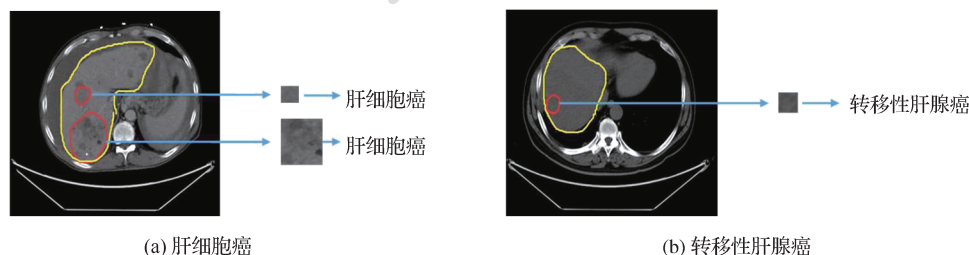


图1 不同病变的 CT 切片

Fig. 1 CT slices of different lesions ((a) hepatocellular carcinoma; (b) metastasis carcinoma)

本文针对肝脏病灶分类任务中存在的训练样本缺乏和已有方法不能充分挖掘有效特征等问题, 提出一种结合特征重用和注意力机制的肝肿瘤自动分类方法, 通过特征重用和特征提取两个模块提高分类性能。首先, 利用特征重用模块复制原通道信息, 首次尝试在增强原始图像信息表达的基础上对数据

进行扩充, 避免因样本较少而引起的过拟合问题。其次, 利用改进的注意力机制从全局和局部两个方面自适应地提取病灶特征。最后, 为了更好地缓解小样本带来的影响, 迁移了 ImageNet 预训练的部分模型权重。在肝细胞癌、血管瘤、转移性肝腺癌及正常肝组织的 514 幅 CT 切片上进行实验。实验结果

表明,该方法可以充分增强病变特征,并在仅有的小样本分类任务上取得了良好的表现性能。

1 本文方法

本文提出的肝脏肿瘤分类方法主要包括以下步骤:1)通过提出的特征重用模块进行数据预处理,增强病灶特征并扩充原始数据;2)使用基于注意力机制的特征提取模块,过滤冗余信息、加强病灶特征提取;3)根据提取到的特征,通过线性分类器对局灶性病变(focal liver lesions, FLLs)进行分类。算法整体流程如图2所示。

1.1 基于特征重用模块的数据预处理

在训练数据量较小的前提下,数据的预处理是模型取得良好分类结果的关键。CT图像与自然图像不同,是由从黑到白不同灰度组成的2D像素矩阵,而自然图像是3通道的3D像素矩阵。为了充分利用原始图像特征并扩充现有数据样本,提出了一种特征重用模块对医学图像进行一种伪自然图像的预处理。

可视化过程如图3所示,图中网格里的数值表示经过每步操作后 patch(局灶性病变中方框标注位置)的像素值。进入分类网络之前,所有数据均进行以下处理:

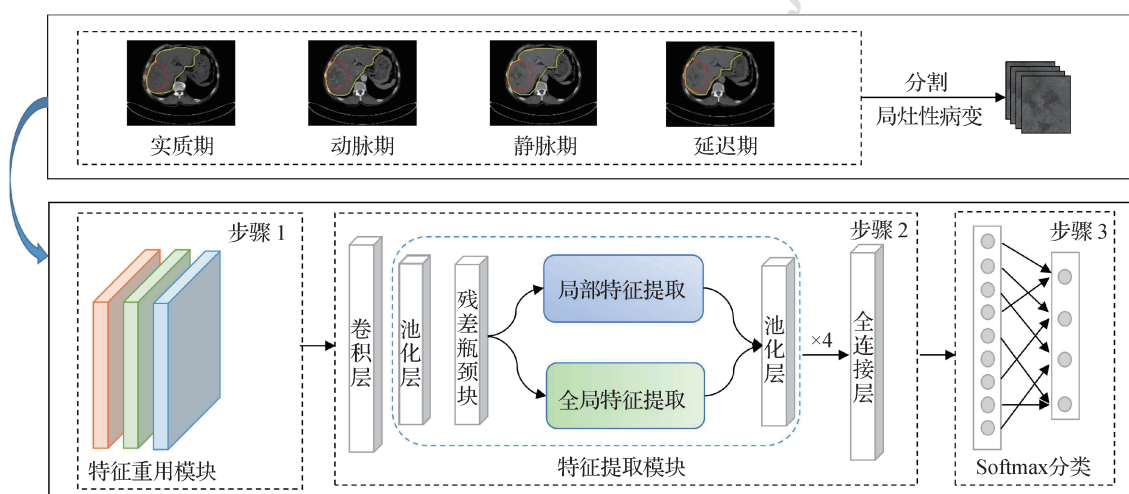


图2 本文算法示意图

Fig. 2 Illustration of the proposed method

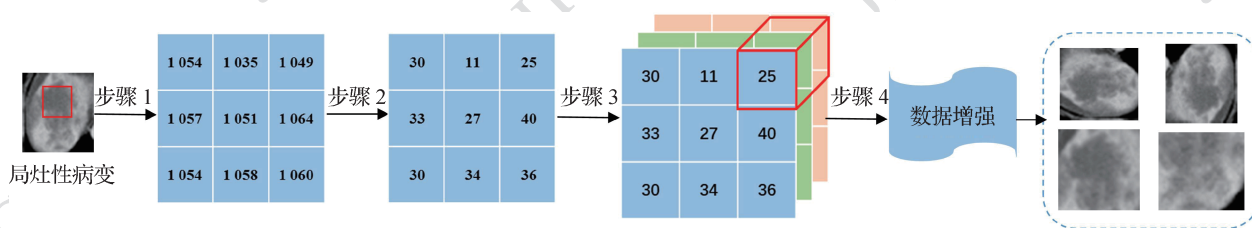


图3 可视化数据预处理的过程

Fig. 3 Visualizing the processing of data preprocessing

1) 获取 FLLs 的 2D 像素矩阵。为了更好地捕获与周围组织相对大小的病变区域,需根据医生标注,从各个时期、不同类别的 CT 切片中分割出 FLLs。读入 DICOM(digital imaging and communications in medicine)格式的 FLLs 并转换为 2D 像素矩阵。如图3中步骤1所示。

2) 获取 CT 值并限制其强度范围。CT 值是测定人体某一局部组织或器官密度大小的一种计量单

位,通常称为亨氏单位(Hounsfield unit, HU)。像素值与 CT 值(用 HU 表示 CT 值)之间的转换关系为

$$HU = PV \times RS + RI \quad (1)$$

式中, RS 表示 RescaleSlope, RI 表示 RescaleIntercept,都可从 CT 图像的头文件中获取; PV 表示像素值。实验中,针对肝脏 CT 数据,取 $RS = 1$, $RI = -1024$ 。

在医学图像中,不同器官或组织具有不同范围

的CT值,但相同部位的器官或组织共享同一范围的CT值,如软组织为30~50 HU。为了避免腹部CT切片中其他非相关器官或组织的影响,在获取HU之后,需要对肝脏进行HU的限制。具体为

$$L_{HU} = \begin{cases} -100 & HU \leq -100 \\ HU & -100 < HU < 400 \\ 400 & HU \geq 400 \end{cases} \quad (2)$$

通过上述操作,将HU的值限制在 $[-100, 400]$ (Christ等,2016;刘哲等,2018)。如图3中步骤2所示。

3)重用单像素通道为3通道。在原2D的像素矩阵上添加一个新的空间维度,将其转换为3D的像素矩阵。之后沿着新的空间轴,复制3次原2D像素矩阵,并串联为一个整体,作为伪RGB的3通道的特征图(3D像素矩阵)。如图3中步骤3所示。

4)处理数据以扩充样本。以自然图像数据增强的方式预处理CT图像。对3D像素矩阵执行各种转换(transforms)操作,如:旋转、裁剪和填充等,以生成新的样本数据。如图3中步骤4所示。

使用上述方法预处理数据主要有两大优点:1)增强原始特征。通过重用通道信息不仅能以3倍强度突出原始特征,而且可以作为一种信息补偿机制,缓解某一通道经CNN提取特征不足的问题。2)扩充数据,避免过拟合。基于原始数据新生成的病变区域无需医生重新标注,可直接用于训练。此外,不同于卷积操作会引入大量参数,它是通过数值计算(numerical python, numpy)来复制重用厚特征信息,几乎没有引入任何参数,可为医学图像的预处理提供参考。

1.2 基于注意力机制的特征提取模块

为充分挖掘FLLs的语义信息,提出一种基于注意力机制的特征提取模块,从局部和全局两个方面加强特征提取。通过局部特征提取为每个像素生成一个权重因子加强特征提取,通过全局特征提取为每个通道生成一个2D的权重矩阵调整每个通道的重要性。

1.2.1 局部特征提取

针对CNN在反向传播中每一层参数根据梯度更新不同以及每个像素对梯度响应的不同(Fu等,2019;纪超等,2019),提出一种像素—像素的局部特征提取方法,结构如图4(a)所示,包括主干和权重两个分支。对任意输入的特征图 $x_t \in \mathbf{R}^{c \times h \times w}$,主干分支采用2个 3×3 的分组卷积(group convolution)以充分挖掘图像的深层语义信息 $T(x_t)$ 。权重分支由一组 2×2 的分组卷积与反卷积构成,具体操作如下:1)通过分组卷积对图像特征进行快速编码,得到尺度减半的特征图。基于该特征图,通过反卷积以恢复至与主干维度一致的特征图,从而得到具有强语义信息的深度特征图 $M(x_t)$ 。2)根据残差设置,聚合原始信息 X_t (X_t 与 x_t 表示经某一特征映射后获得的同一特征图)。3)将融合的信息进行Sigmoid激活以得到每个空间位置上像素的权重因子 $F(x_t)$ 。该分支计算为

$$F(x_t) = \sigma(M(x_t) + X_t) \quad (3)$$

式中, σ 是门控机制函数Sigmoid, $X_t \in \mathbf{R}^{c \times h \times w}$,其中, c 为通道数, h 为图像高度, w 为图像宽度, $t \in \{1, 2, \dots, c\}$ 为通道索引。

通过权重因子与主干分支上每个不同位置的像素

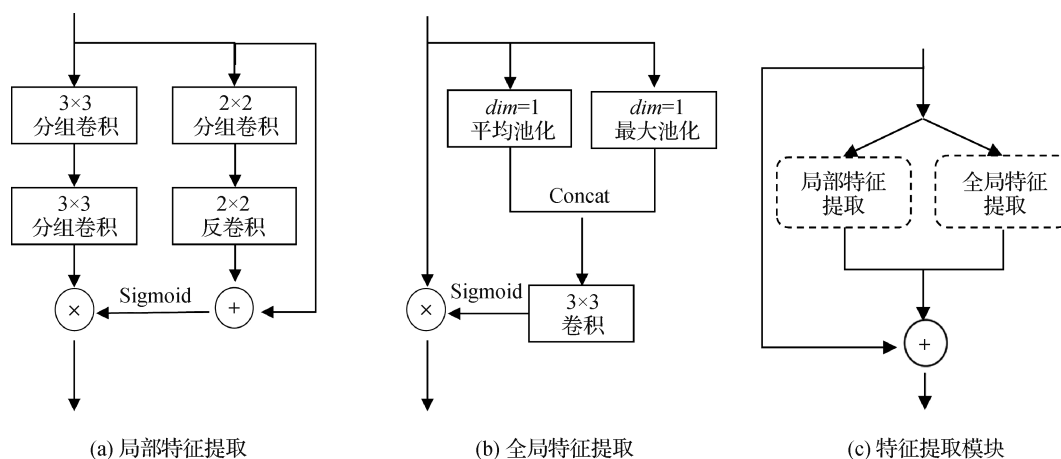


图4 特征提取模块的结构

Fig.4 The structure of feature extraction module((a) local feature extraction; (b) global feature extraction; (c) feature extraction module)

点乘进行加权,得到局部的注意力特征图 $A(x_i)$,即

$$A(x_i) = F(x_i) * T(x_i) \quad (4)$$

式中, $F(x_i)$ 与 $T(x_i)$ 具有相同维度 $\mathbf{R}^{c \times h \times w}$, $*$ 表示 3 维矩阵中同一位置对应的像素进行点乘。

通过对主干特征进行加权,权重分支不仅可以作为前向传播过程中的特征选择器,而且可以作为反向传播过程中的梯度更新滤波器,即

$$\frac{\partial M(x_i, \theta) T(x_i, \varphi)}{\partial \varphi} = M(x_i, \theta) \frac{\partial T(x_i, \varphi)}{\partial \varphi} \quad (5)$$

式中, θ 和 φ 分别是权重分支和主干分支的参数。通过权重分支生成权重因子,可以有效避免背景区域对主干梯度更新的影响。

1.2.2 全局特征提取

由于整个特征空间的特征不受噪声影响(Li 等,2019a),为避免诸如噪声等非相关分类因素的影响,提出一种全局特征提取的方法,结构如图 4(b)所示,采用基于注意力机制的空间池化自适应地关注特征空间中每个病变的通道信息。空间池化与 CNN 的池化操作不同,CNN 的池化操作分别对 3 维特征图中每个通道执行池化,而空间池化是基于空间沿通道轴($dim = 1$)上的一种池化。全局特征提取的详细操作步骤如下:

1) 计算平均空间池化与最大空间池化,得到两个 2D 的像素矩阵 G_m 和 G_n ,具体计算为

$$G_m = \frac{1}{t} \sum_{i=1}^t X_i \quad (6)$$

$$G_n = M_d(X_i) \quad (7)$$

式中, i 为空间索引,取遍空间 t 中所有索引位置, $M_d(\cdot)$ 表示利用空间最大池化操作对其中的像素矩阵进行压缩。

2) 两个 2D 矩阵通过 Concat 得到一个 2 通道的特征矩阵,之后被送入一个 3×3 的标准卷积层中

进行聚合。此后,经 Sigmoid 激活以得到空间中每个通道的 2D 权重矩阵,即

$$Y = \sigma(f^{3 \times 3}[G_m; G_n]) \quad (8)$$

式中, $Y \in \mathbf{R}^{1 \times h \times w}$, $f^{3 \times 3}$ 表示卷积核为 3×3 的卷积操作, σ 表示 Sigmoid 激活函数。

3) 将通道的权重矩阵加权到原始特征图 X_i 中的每个通道上,得到全局的注意力特征图 $G(x_i)$,即

$$G(x_i) = X_i * Y \quad (9)$$

1.2.3 整体模块设计及网络架构

整体特征提取模块设计如图 4(c)所示,为了深入挖掘仅有样本的高维特征,以并行的方式提取局部和全局两种特征。此外,遵循残差网络的设计思想(He 等,2016),聚合原始信息 X_i ,即

$$U = A(x_i) + YG(x_i) + X_i \quad (10)$$

式中, Y 是一个自学习参数,初始化为 0,随着网络的训练,会根据梯度下降算法更新相应的系数。

网络的整体结构及参数设置如表 1 所示,其中 C(convolution layer)表示卷积层、AP(average pooling layer)表示平均池化层、FE(feature extraction module)表示特征提取模块、MP(max pooling)表示最大池化层以及 FC(fully connected layer)表示全连接层,表 1 中 C、S、P 后的数值依次为核大小、步长和填充值。本文将特征提取模块 FE 作为嵌入模块,嵌入至 SENet34(Hu 等,2018)不同维度(64,128,256,512)的瓶颈块(He 等,2016)中。对网络结构进一步优化(He 等,2019),用 3 个 3×3 的卷积层替代前网络层中原 7×7 的卷积层,以精确提取病灶的边缘特征。然后,将瓶颈残差块中第 1 个 1×1 ,stride 为 2 的卷积层替换为在第 2 个 3×3 的卷积处进行下采样,避免特征信息的丢失。此外,训练的过程中,迁移了 ImageNet 预训练 ResNet34 的部分参数,仅迁移与本文网络架构相同层的权重参数。

表 1 网络结构与参数设置

Table 1 Network structure and parameter setting

名称	结构			
前网络层	$\begin{bmatrix} C: 3, S: 2, P: 1 \\ \text{BatchNorm} \end{bmatrix}$	$\begin{bmatrix} C: 3, S: 1, P: 1 \\ \text{BatchNorm} \end{bmatrix}$	$\begin{bmatrix} C: 3, S: 1, P: 1 \\ \text{BatchNorm} \\ \text{ReLU} \end{bmatrix}$	$[\text{MP}: 3, S: 2, P: 1]$
嵌入模块	$\begin{bmatrix} C: 1, S: 1 \\ C: 3, S: 2 \\ C: 1, S: 1 \\ \text{FE}: 64 \end{bmatrix}$	$\begin{bmatrix} C: 1, S: 1 \\ C: 3, S: 2 \\ C: 1, S: 1 \\ \text{FE}: 128 \end{bmatrix}$	$\begin{bmatrix} C: 1, S: 1 \\ C: 3, S: 2 \\ C: 1, S: 1 \\ \text{FE}: 256 \end{bmatrix}$	$\begin{bmatrix} C: 1, S: 1 \\ C: 3, S: 2 \\ C: 1, S: 1 \\ \text{FE}: 512 \end{bmatrix}$
后网络层	$[\text{AP}: 2, S: 1]$	$[\text{FC}: 2048, 4]$	$[\text{Softmax}]$	

2 实验结果与分析

实验在 NVIDIA GeForce GTX 1080 Ti 显卡, 8 GB 显存, 32 GB 内存, python 3.7.1 环境下进行, 使用版本为 1.0.1 的 pytorch 深度学习开源框架, 实验任务是基于腹部肝脏 CT 图像对局灶性病变进行分类。

2.1 数据集

实验数据来自江苏大学附属医院, 由美国单光子扫描仪 (single-photon emission computed tomography, SPECT) 和德国西门子 52 环扫描仪 (positron emission tomography, PET-CT) 获取, SPECT 和 PET-CT 分别依靠扫描仪发射单光子及正电子得到动态多期的 CT 图像。实验数据包括 120 个病人的 430 个病灶, 其中肝细胞癌 (hepatocellular carcinoma, Hcc) 图像 184 幅、血管瘤 (hemangioma, Hem) 图像 83 幅和转移性肝腺癌 (metastasis, Met) 图像 163 幅。为了增加数据的多样性, 从非标注肝脏位置中截取 84 幅正常肝组织 (healthy, heal) 图像, 构成 514 幅的四分类数据集, 数据样本如图 5 所示, 图中病灶样本是根据医生标注截取的病灶区域。输入到分类网络的图像是根据放射科医生标注截取的感兴趣区域 (region of interests, ROIs)。由于病人的数据一般是私有的, 目前没有公开的肝肿瘤图像数据集, 因此本文所有实验均基于该数据集。

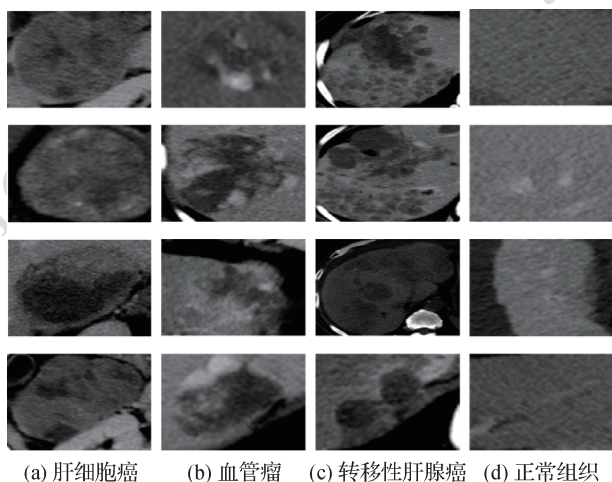


图5 不同类型的数据样本

Fig.5 Different types of data samples

((a) hepatocellular carcinoma; (b) hemangioma;
(c) metastasis; (d) healthy tissue)

2.2 模型参数设置及实验过程

实验过程中, 所有样本均采用 1.1 节中提出的方法对数据进行预处理, 并将不规则尺寸 FLLs 的大小裁剪为 64×64 像素, 采用随机中心裁剪与随机水平翻转的方式进行数据增强, 采用交叉熵作为损失函数, 每次从训练集中取 16 (batchsize) 个样本训练, 共训练 150 个 epoch (数据集所有图像训练 1 次为 1 个 epoch)。将自适应学习率优化算法 Adam (Kingma 和 Ba, 2015) 作为优化器, 初始学习率为 0.001, 学习率下降周期 stepsize 为 1, 衰减因子 gamma 取 0.98, 学习率衰减策略为每隔 stepsize 个 epoch 学习率变为原来的 0.98。

为了充分利用现有数据, 避免小样本影响, 采用 5 折交叉验证 (5-fold cross validation), 每折 (fold) 中随机选取 20% 的样本用来测试, 为避免实验结果的偶然性, 所有实验结果均取 5-fold 均值作为最终分类结果。实验分为训练和测试两个阶段: 1) 训练阶段。将 CT 图像中的 ROIs 输入基于特征重用和注意力机制的肝脏肿瘤分类网络, 通过特征重用模块对数据进行预处理, 采用特征提取模块提取病灶特征, 且分类网络部分层参数不随机初始化, 采用 ImageNet 预训练参数。通过训练, 得到在训练集上表现良好的模型。2) 测试阶段。将测试集中的 ROIs 输入训练好的模型中, 通过 Softmax 层自动对其进行分类, 将预测的病变类别与其真实标签对比, 计算分类准确率。

2.3 评价指标

为充分证明提出方法的有效性, 设计实施了综合实验并采用准确率、精确率、召回率 (灵敏度)、F1-score 及 AUC 等多项指标进行评价。

准确率 (accuracy) (Chen 等, 2017) 是分类模型中所有预测正确的分类样本占总样本的比重。精确率 (precision) (Frid-Adar 等, 2018a) 表示在模型预测为某一类别的结果中, 模型预测正确的比重。召回率 (recall) (Frid-Adar 等, 2018a) 指在真实标签为某一类别的结果中, 模型预测正确的比重。根据精确率和召回率, 可计算 F1-score (Chen 等, 2017)。各项指标具体计算为

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

$$P = \frac{TP}{TP + FP} \quad (12)$$

$$R = \frac{TP}{TP + FN} \quad (13)$$

$$F1 = \frac{2PR}{P + R} \quad (14)$$

式中, A 、 P 、 R 、 $F1$ 分别表示准确率、精确率、召回率和 $F1$ -score, TP 与 TN 分别表示当前类别及非当前类别下预测正确的样本数, FP 与 FN 分别表示当前类别与非当前类别下预测错误的样本数。上述 4 项评价指标的取值区间为 $[0, 1]$, 其值越大表示分类模型的性能表现越佳。

AUC (Romero 等, 2019) 是接受者操作特性 (receiver operating characteristic, ROC) 曲线下的面积, 表示当前分类算法根据计算得到的得分值将正样本排在负样本前面的概率。AUC 的值介于 0.1 ~ 1 之间, 越向 1 靠拢越好, 通过其值大小可以直观判断分类方法的优劣。

2.4 实验结果及分析

2.4.1 与基准实验比较

本文以 SENet34 (Hu 等, 2018) 为基准实验, 同时采用特征重用模块、特征提取模块及迁移参数训练策略。本文方法与基准方法的 5 折交叉验证测试准确率如表 2 所示。

从表 2 可以看出, 使用本文方法后, 每折的分类准确率均优于原始方法, 平均准确率提高了 9.7%, 数据方差降低了 0.46%, 表明采用特征重用的数据预处理模块及注意力机制的特征提取模块可以更好地扩充数据及增强特征, 从而缓解数据不足的问题, 提升了分类算法的鲁棒性。

为了从不同角度进一步表明改进方法优于基线

表 2 与基准实验准确率的比较
Table 2 Comparison of accuracy with baseline experiments

每折	基准方法	本文方法
1	77.63	87.94
2	78.40	88.52
3	79.57	87.16
4	77.82	87.74
5	76.85	87.56
均值	78.05 ± 0.91	87.78 ± 0.45

注: 加粗字体表示最优结果。

方法, 通过实验统计显示混淆矩阵加以验证, 如图 6 所示。纵轴表示真实标签, 为便于观察, 用颜色深浅衡量真实标签下被预测为各个类别的概率, 即召回率, 相应的概率值越大其对应的颜色越深, 当预测标签与真实标签一致时 (图中深色对角线位置), 表示分类准确。如图 6(b) 中, 当真实标签为肝细胞癌 (Hcc) 时, 被预测为肝细胞癌 (Hcc)、血管瘤 (Hem)、正常肝组织 (Heal)、转移性肝腺癌 (Met) 的召回率分别为: 85%、4%、1%、9%。从图 6 可以看出, 本文方法针对每一类病灶分类的召回率均优于基线方法, 因为特征重用及特征提取模块可以有效增强及挖掘病变区域的高维语义特征。

ROC 曲线的横轴为负正类率 (false positive rate), 即特异度, 表示非当前类别下预测正确的概率; 纵轴为真正类率 (true positive rate), 即灵敏度, 表示当前

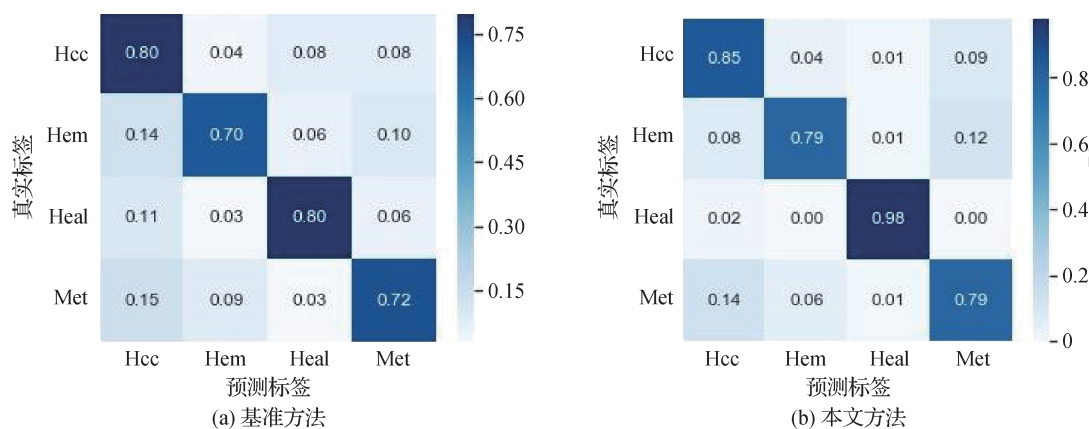


图 6 测试集上的混淆矩阵

Fig. 6 Confusion matrix on valid dataset ((a) baseline method; (b) ours)

类别下预测正确的概率。图7—图10分别是肝细胞癌、血管瘤、转移性肝腺癌和正常肝组织的 ROC 曲线。通过 ROC 曲线可以计算得到 AUC, 进而对分类算法做出评价。从图7—图10可以看出, 对于不同病变类型

的 ROIs, 本文方法与基线方法相比, 五折交叉验证的平均(mean) ROC 曲线更偏离 chance (45°角对角线), 横轴结果更靠近 1.0, 且 AUC 值更大, 曲线波动更小, 模型相对更加稳定, 进一步证明了本文方法的优越性。

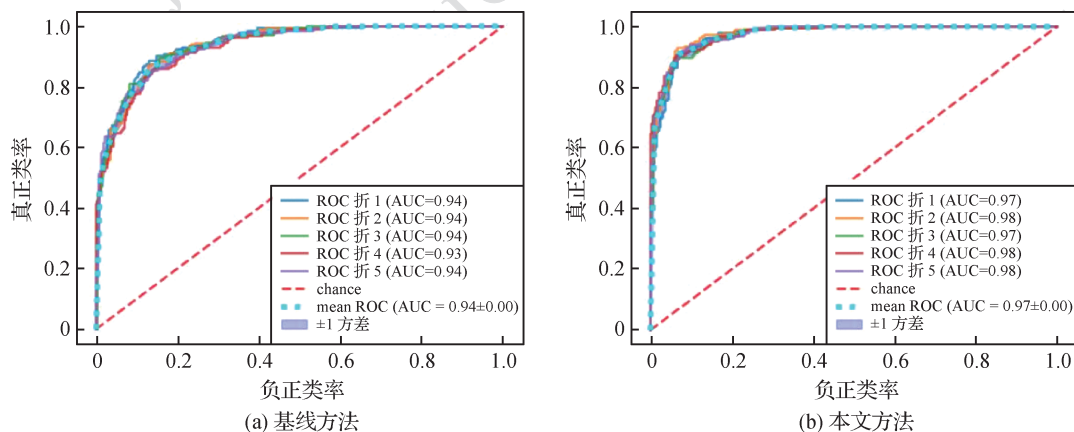


图7 肝细胞癌的 ROC 曲线

Fig. 7 ROC curves of hepatocellular carcinoma ((a) baseline method; (b) proposed)

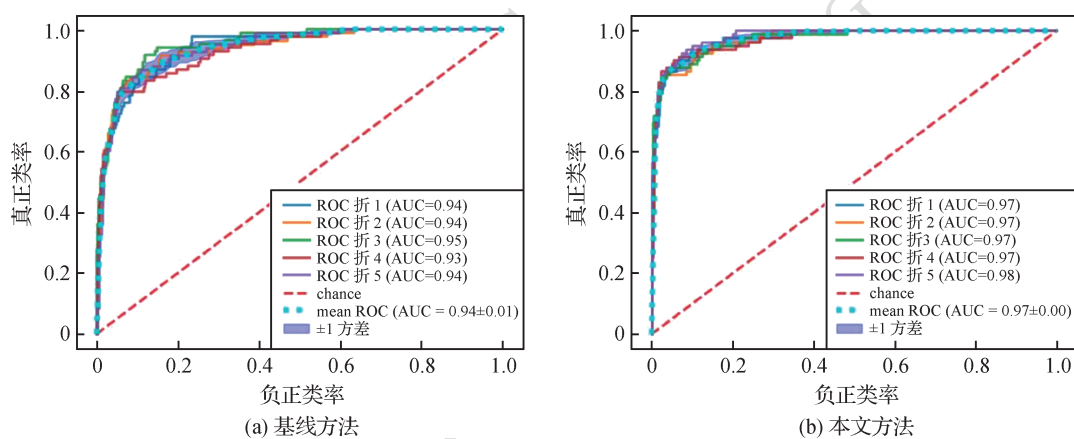


图8 血管瘤的 ROC 曲线

Fig. 8 ROC curves of hemangioma ((a) baseline method; (b) proposed)

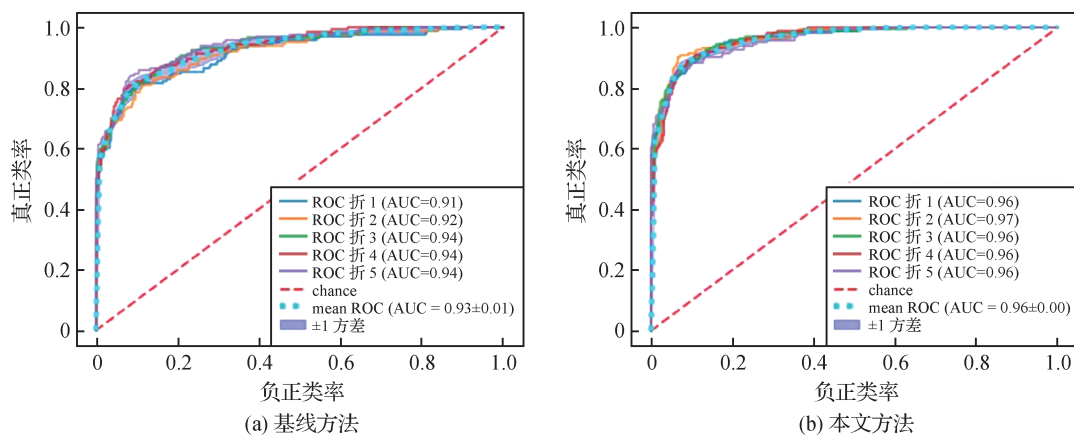


图9 转移性肝腺癌的 ROC 曲线

Fig. 9 ROC curves of metastasis ((a) baseline method; (b) proposed)

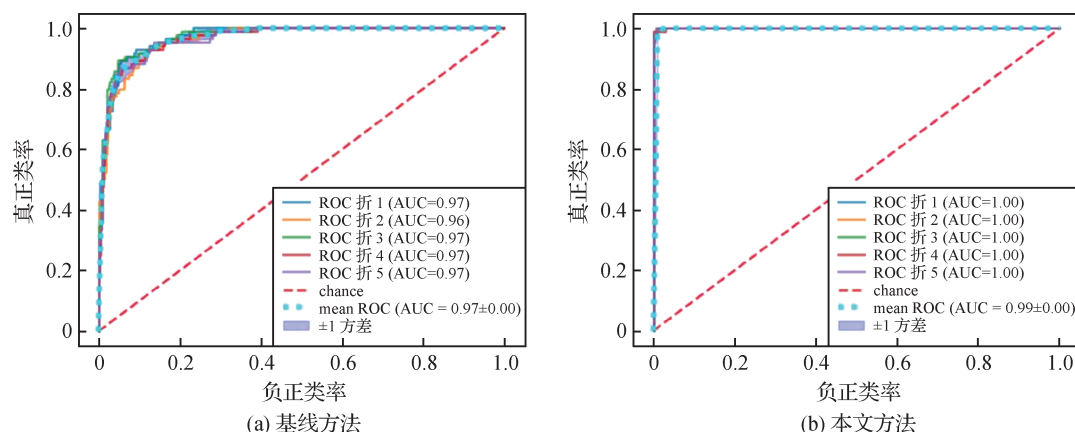


图10 正常肝组织的 ROC 曲线

Fig. 10 ROC curve of healthy liver tissue ((a) baseline method; (b) proposed)

2.4.2 不同训练策略下的消融实验

本文的训练策略主要包含特征重用模块 (feature reuse module, FRM)、特征提取模块 (feature extraction module, FEM) 和迁移参数训练策略 (transfer learning parameters, TLP)。为了表明各项策略的有效性,在相同数据集上,仅控制一项作为变量进行消融实验,结果如表 3 所示。可以看出,采用 FRM

后,准确率、召回率及精确率分别提高了 3.9%、4.15%、4.22%,表明 FRM 可以对数据进行有效扩充,缓解数据量不足问题。通过利用 FEM,评价指标平均提高了 4.47%,因为 FEM 可以自适应地加强病灶特征提取,缓解背景及噪声影响。TLP 的训练策略进一步缓解了当前医学数据缺乏的现状。

表3 不同训练策略下的消融实验

Table 3 Ablation experiments by different training strategies

方法	准确率	召回率	精确率
SENet34	78.05 ± 0.91	75.37 ± 0.64	74.67 ± 0.61
本文 FEM	82.98 ± 0.55	79.31 ± 0.60	79.21 ± 0.58
本文 FEM + FRM	86.88 ± 0.48	83.46 ± 0.23	83.43 ± 0.28
本文 FEM + FRM + TLP	87.78 ± 0.45	84.43 ± 0.26	84.59 ± 0.28

注:加粗字体表示各列最优结果。

本文对医学数据进行预处理构建的特征重用模块属于轻量级模块,几乎没有引入任何计算代价,可以嵌入到任何分类网络中对医学图像进行预处理。基于特征重用模块的消融实验结果如表 4 所示,可以看出,通过在原有模型中嵌入 FRM,原分类准确率均有很大程度的提升,说明 FRM 可以进行有效的数据预处理,显著增强并扩充原始图像,从而提高原分类准确率。

2.4.3 与其他分类方法比较

1) 与肝肿瘤分类算法比较。由于肝脏 CT 图像是私有的且缺乏针对肝肿瘤分类的开源代码,因此无法将本文方法与现有的肝脏 CT 肿瘤分类方法进

表4 基于特征重用模块的消融实验结果

Table 4 Results of ablation experiment based on FRM

方法	准确率	准确率 (+ FRM)	增幅
SENet34 (Hu 等, 2018)	78.05	82.21	4.16
CBAM34 (Woo, 2018)	84.31	88.23	3.92
SKNet34 (Li 等, 2019b)	82.23	86.04	3.81
本文	82.98	86.88	3.90

注:加粗字体表示各行最优结果。

行比较,表 5 列举的是当前分类方法对肝细胞癌、血管瘤、转移性肝腺癌、肝结节增生、肝囊肿及正

常肝组织的实验数据,可以看出,与目前分类算法相比,本文分类算法表现优异。

2)与主流分类模型比较。为验证本文模型的分类效果,将本文使用的肝脏数据分别输入至 DensNet121 (Huang 等,2017)、SENet (Hu 等,2018)、SE_Resnext (Hu 等,2018)、CBAM (Woo 等,2018)、

SKNet (Li 等,2019b)模型和本文方法,完成对模型的训练及测试,测试结果如表 6 所示。从表 6 可以看出,本文方法的分类评价指标均高于其他模型,原因在于本文方法使用了基于全局和局部的特征提取模块,可以充分挖掘肝 CT 的细节信息,缓解背景及噪声的干扰,从而提高了分类精度。

表 5 本文方法与其他肝肿瘤分类方法的比较

Table 5 Comparison of the proposed method with other methods for the classification of liver tumors

方法	数据类型数量/幅(召回率/%)						准确率/%
	肝细胞癌	血管瘤	转移性肝腺癌	肝结节增生	肝囊肿	正常肝组织	
Frid-Adar 等人 (2018a)	-	65 (72.32)	64 (68.71)	-	53 (98.14)	-	85.71
Ben-Cohen 等人 (2018b)	-	65 (52.31)	81 (53.08)	-	66 (90.93)	27 (92.59)	75.00
Liang 等人 (2018a)	132 (85.11)	124 (85.16)	-	114 (84.05)	110 (93.75)	-	85.46
本文	184 (85.73)	83 (79.67)	163 (79.47)	-	-	84 (98.31)	87.78

注:“-”表示未使用该类样本进行实验,加粗字体表示各列最优结果。

表 6 本文方法与主流分类模型的比较

Table 6 Proposed method comparison with mainstream classification models

方法	准确率	召回率	精确率	F1-score	AUC
DensNet121	83.43 ± 0.34	80.53 ± 0.85	79.93 ± 0.99	80.12 ± 0.91	95.56 ± 1.77
SENet34	78.05 ± 0.91	75.37 ± 0.64	74.67 ± 0.61	74.90 ± 0.61	94.40 ± 1.62
SENet50	80.16 ± 0.14	76.87 ± 0.32	76.62 ± 0.23	76.60 ± 0.26	94.58 ± 1.85
SE_Resnext	79.51 ± 0.44	76.52 ± 0.09	77.08 ± 0.08	76.69 ± 0.05	94.58 ± 2.09
CBAM34	84.31 ± 0.32	82.64 ± 0.29	83.08 ± 0.32	82.72 ± 0.30	96.13 ± 2.08
CBAM50	81.26 ± 0.44	78.58 ± 0.27	79.20 ± 0.21	78.74 ± 0.23	94.75 ± 2.13
SKNet34	82.23 ± 0.42	78.29 ± 0.23	78.66 ± 0.22	78.39 ± 0.22	95.01 ± 1.41
SKNet50	82.36 ± 0.29	79.49 ± 0.36	79.23 ± 0.37	79.30 ± 0.36	95.03 ± 1.73
本文	87.78 ± 0.45	84.43 ± 0.26	84.59 ± 0.28	84.44 ± 0.23	97.50 ± 1.07

注:加粗字体表示各列最优结果。

3 结 论

针对现有肝脏病灶分类方法中样本缺乏、病灶信息缺失等问题,本文提出了一种结合特征重用和注意力机制的肝肿瘤自动分类方法,从两个方面增强病灶特征。1)通过特征重用模块复制原始通道信息并扩充当前样本,不仅可以缓解小数据引起的过拟合问题,而且可以以 3 倍突出显示原始 CT 图像信息;2)通过基于注意力机制的特征提取模块,有效捕获病变区域的特征,该模块主要融合了局部

和全局两个方面的信息。此外,为了进一步缓解小样本问题,采用迁移学习的训练机制更新训练网络参数。在 120 个病人的数据集上进行了大量实验,结果表明,本文方法的准确率、F1-score 和 AUC 等多项评价指标结果均优于对比的分类算法及主流分类框架,能够有效加强图像语义信息的提取。但在以下方面存在优化空间:计算机辅助诊断中对各种病灶图像的处理速度尚不够快,各类样本数目及正异常病灶的差异导致分类结果不平衡,如图 6(b)中正常肝组织与其他病灶的分类召回率差异较大。因此,特征提取模型的优化及样本不平衡的问题在未

来仍需要研究解决。

参考文献 (References)

- Ben-Cohen A, Klang E, Kerpel A, Konen E, Amitai M M and Greenspan H. 2018a. Fully convolutional network and sparsity-based dictionary learning for liver lesion detection in CT examinations. *Neurocomputing*, 275: 1585-1594 [DOI: 10.1016/j.neucom.2017.10.001]
- Ben-Cohen A, Mechrez R, Yedidia N and Greenspan H. 2018b. Improving CNN training using disentanglement for liver lesion classification in CT [EB/OL]. [2019-11-01]. <https://arxiv.org/pdf/1811.00501.pdf>
- Bray F, Ferlay J, Soerjomataram I, Siegel R L, Torre L A and Jemal A. 2018. Global cancer statistics 2018; GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *A Cancer Journal for Clinicians*, 68(6): 394-424 [DOI: 10.3322/caac.21492]
- Chen Y F, Yue X D, Fujita H and Fu S Y. 2017. Three-way decision support for diagnosis on focal liver lesions. *Knowledge-Based Systems*, 127: 85-99 [DOI: 10.1016/j.knosys.2017.04.008]
- Christ P F, Elshaer M E A, Ettlinger F, Tatavarty S, Bickel M, Bilic P, Rempfler M, Armbruster M, Hofmann F, D'Anastasi M, Sommer W H, Ahmadi S A and Menze B H. 2016. Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields//Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention. Athens: Springer; 415-423 [DOI: 10.1007/978-3-319-46723-8_48]
- Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J and Greenspan H. 2018a. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*, 321: 321-331 [DOI: 10.1016/j.neucom.2018.09.013]
- Frid-Adar M, Klang E, Amitai M, Goldberger J and Greenspan H. 2018b. Synthetic data augmentation using GAN for improved liver lesion classification//Proceedings of the 15th IEEE International Symposium on Biomedical Imaging. Washington DC, USA: IEEE; 289-293 [DOI: 10.1109/ISBI.2018.8363576]
- Fu J, Liu J, Tian H J, Li Y, Bao Y J, Fang Z W and Lu H Q. 2019. Dual attention network for scene segmentation//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE; 3146-3149 [DOI: 10.1109/CVPR.2019.00326]
- 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, NV, USA: IEEE; 770-778 [DOI: 10.1109/CVPR.2016.90]
- He T, Zhang Z, Zhang H, Zhang Z Y, Xie J Y and Li M. 2019. Bag of tricks for image classification with convolutional neural networks//Proceedings the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE; 558-567 [DOI: 10.1109/CVPR.2019.00065]
- Hu J, Shen L and Sun G. 2018. Squeeze-and-excitation networks//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE; 7132-7141 [DOI: 10.1109/CVPR.2018.00745]
- Huang G, Liu Z, van der Maaten L and Weinberger K Q. 2017. Densely connected convolutional networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE; 2261-2269 [DOI: 10.1109/CVPR.2017.243]
- Ji C, Huang X B, Cao W, Zhu Y C and Zhang Y. 2019. Fusion of deep learning and global-local features of the image salient region calculation. *Journal of Computer-Aided Design and Computer Graphics*, 31(10): 1838-1846. (纪超, 黄新波, 曹雯, 朱永灿, 张烨. 2019. 结合深度学习和全局-局部特征的图像显著区域计算. *计算机辅助设计与图形学学报*, 31(10): 1838-1846) [DOI: 10.3724/SP.J.1089.2019.17544]
- Kingma D and Ba J. 2015. Adam: a method for stochastic optimization//Proceedings of the 3rd International Conference on Learning Representations. San Diego, CA, USA: ICLR; 1-15
- Krizhevsky A, Sutskever I and Hinton G E. 2017. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6): 84-90 [DOI: 10.1145/3065386]
- Li X, Hu X L and Yang J. 2019a. Spatial group-wise enhance: improving semantic feature learning in convolutional networks [EB/OL]. [2019-05-25]. <https://arxiv.org/pdf/1905.09646.pdf>
- Li X, Wang W H, Hu X L and Yan J. 2019b. Selective kernel networks//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE; 510-519 [DOI: 10.1109/CVPR.2019.00060]
- Liang D, Lin L F, Hu H J, Zhang Q W, Chen Q Q, Lwamoto Y, Han X H and Chen Y W. 2018a. Residual convolutional neural networks with global and local pathways for classification of focal liver lesions//Proceedings of the 15th Pacific Rim International Conference on Artificial Intelligence. Nanjing, China: Springer; 617-628 [DOI: 10.1007/978-3-319-97304-3_47]
- Liang D, Lin L F, Hu H J, Zhang Q W, Chen Q Q, Lwamoto Y, Han X H and Chen Y W. 2018b. Combining convolutional and recurrent neural networks for classification of focal liver lesions in multi-phase CT images//Proceedings of the 21st International Conference on Medical Image Computing and Computer Assisted Intervention. Granada, Spain: Springer; 666-675 [DOI: 10.1007/978-3-030-00934-2_74]
- Litjens G, Kooi T, Bejnordi B E, Setio A A A, Ciompi F, Ghafoorian M, van der Laak J A W M, van Ginneken B and Sánchez C I. 2017. A survey on deep learning in medical image analysis. *Medical*

- Image Analysis, 42: 60-88 [DOI: 10. 1016/j. media. 2017. 07. 005]
- Liu Z, Zhang X L, Song Y Q, Zhu Y and Yuan D Q. 2018. Liver segmentation with improved U-Net and Morphsnakes algorithm. Journal of Image and Graphics, 23(8): 1254-1262 (刘哲, 张晓林, 宋余庆, 朱彦, 袁德琪. 2018. 结合改进的 U-Net 和 Morphsnakes 的肝脏分割. 中国图象图形学报, 23(8): 1254-1262) [DOI: 10. 11834/jig. 170585]
- Romero F P, Diler A, Bisson-Gregoire G, Turcotte S, Lapointe R, Vandenbroucke-Menu F, Tang A and Kadoury S. 2019. End-to-end discriminative deep network for liver lesion classification//Proceedings of the 16th IEEE International Symposium on Biomedical Imaging. Venice, Italy; IEEE: 1243-1246 [DOI: 10. 1109/ISBI. 2019. 8759257]
- Szegedy C, Ioffe S, Vanhoucke V and Alemi A A. 2017. Inception-v4, inception-ResNet and the impact of residual connections on learning//Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, CA, USA; AAAI: 4278-4284
- Szegedy C, Vanhoucke V, Ioffe S, Shlens J and Wojna Z. 2016. Rethinking the inception architecture for computer vision//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA; IEEE: 2818-2826 [DOI: 10. 1109/CVPR. 2016. 308]
- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: Convolutional block attention module//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany; Springer: 3-19 [DOI: 10. 1007/978-3-030-01234-2_1]
- Yasaka K, Akai H, Abe O and Kiryu S. 2017. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study. Radiology, 286(3): 887-896 [DOI: 10. 1148/radiol. 2017170706]

作者简介



冯诺, 1992年生, 女, 硕士研究生, 主要研究方向为医学图像处理。

E-mail: 2221808007@stmail. ujs. edu. cn



宋余庆, 通信作者, 男, 教授, 主要研究方向为医学图像处理与分析、数据挖掘、模式识别、计算机医学应用。

E-mail: yqsong@ujs. edu. com

刘哲, 女, 副教授, 主要研究方向为医学图像处理与分析。

E-mail: lzhe@ujs. edu. cn