



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

中国图象图形学报

2021
07
VOL.26

ISSN1006-8961
CN11-3758/TB



中国多媒体大会(ChinaMM2020)专刊



第26卷第7期（总第303期）

2021年7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

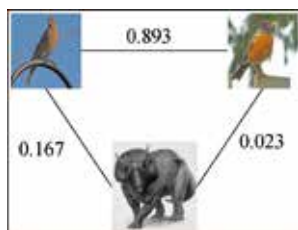
ISSN 1006-8961

CODEN ZTTXFZ

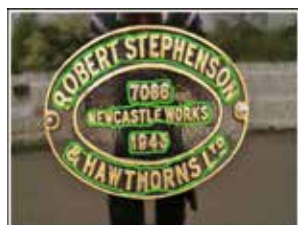
国外发行代号 M1406

国内邮发代号 82-831

国内定价 60.00元



类别语义相似性监督的小样本图像识别(第1594页)



像素聚合和特征增强的任意形状场景文本检测(第1614页)



时序特征融合的视频实例分割(第1692页)

综述

面向上行流媒体的压缩感知视频流技术前沿

刘浩, 黄荣, 袁浩东 1545

图像处理和编码

耦合保持投影哈希跨模态检索

闵康凌, 张国宾, 王磊, 李丹萍 1558

面向视觉搜索的空间局部敏感哈希方法

黄小燕, 孙彬, 杨展源, 朱映映, 田奇 1568

图像分析和识别

融合视觉关系检测的电力场景自动危险预警

高明, 左红群, 柏帆, 田清阳, 葛志峰, 董兴宁, 甘甜 1583

类别语义相似性监督的小样本图像识别

徐鹏帮, 桑基韬, 路冬媛 1594

局部双目视差回归的目标距离估计

张羽丰, 李昱希, 赵明璧, 喻晓源, 占云龙, 林巍峭 1604

像素聚合和特征增强的任意形状场景文本检测

师广琛, 巫义锐 1614

相位一致性指导的全参考全景图像质量评价

夏雨蒙, 王永芳, 王闯 1625

图像理解和计算机视觉

特征金字塔结构的时序行为识别网络

何嘉宇, 雷军, 李国辉 1637

多模态多层次事件网络的谣言检测

李莎, 张怀文, 钱胜胜, 方全, 徐常胜 1648

多模态零样本人体动作识别

吕露露, 黄毅, 高君宇, 杨小汕, 徐常胜 1658

足球视频球员感知跟踪算法

冯思佳, 宋子恺, 于俊清, 何云峰, 管涛 1668

融合时空图卷积的多人交互行为识别

成科扬, 吴金霞, 王文杉, 荣兰, 詹永照 1681

时序特征融合的视频实例分割

黄泽涛, 刘洋, 于成龙, 张加佳, 王轩, 漆舒汉 1692

医学图像处理

多尺度深度特征提取的肝脏肿瘤CT图像分类

毛静怡, 宋余庆, 刘哲 1704

3D多尺度深度卷积神经网络肺结节检测

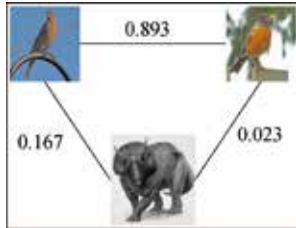
孙华聪, 彭延军, 郭燕飞, 张晓庆 1716

结合多通道注意力的糖尿病性视网膜病变分级

顾婷菲, 郝鹏翼, 白琮, 柳宁 1726

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Few shot image recognition based on class semantic similarity supervision(P1594)



Arbitrary shape scene-text detection based on pixel aggregation and feature enhancement (P1614)



Video instance segmentation based on temporal feature fusion (P1692)

Review

Survey on compressive sensing video stream for uplink streaming media

Liu Hao, Huang Rong, Yuan Haodong 1545

Image Processing and Coding

Structure-preserving hashing with coupled projections for cross-modal retrieval

Min Kangling, Zhang Guobin, Wang Lei, Li Danping 1558

Locality-sensitive hashing approach based on semantic space for visual retrieval

Huang Xiaoyan, Sun Bin, Yang Zhanyuan, Zhu Yingying, Tian Qi 1568

Image Analysis and Recognition

Visual relationship detection-based emergency early-warning description generation in electric power industry

Gao Ming, Zuo Hongqun, Bai Fan, Tian Qingyang, Ge Zhifeng, Dong Xingning, Gan Tian 1583

Few shot image recognition based on class semantic similarity supervision

Xu Pengbang, Sang Jitao, Lu Dongyuan 1594

Object distance estimation based on stereo regional disparity regression

Zhang Yufeng, Li Yuxi, Zhao Mingbi, Yu Xiaoyuan, Zhan Yunlong, Lin Weiyao 1604

Arbitrary shape scene-text detection based on pixel aggregation and feature enhancement

Shi Guangchen, Wu Yirui 1614

Phase consistency guided full-reference panoramic image quality assessment algorithm

Xia Yumeng, Wang Yongfang, Wang Chuang 1625

Image Understanding and Computer Vision

Temporal action detection based on feature pyramid hierarchies

He Jiayu, Lei Jun, Li Guohui 1637

Multi-modal multi-level event network for rumor detection

Li Sha, Zhang Huaiwen, Qian Shengsheng, Fang Quan, Xu Changsheng 1648

Multimodal-based zero-shot human action recognition

Lyu Lulu, Huang Yi, Gao Junyu, Yang Xiaoshan, Xu Changsheng 1658

Players-aware tracking algorithm in soccer video

Feng Sijia, Song Zikai, Yu Junqing, He Yunfeng, Guan Tao 1668

Multi-person interaction action recognition based on spatio-temporal graph convolution

Cheng Keyang, Wu Jinxia, Wang Wenshan, Rong Lan, Zhan Yongzhao 1681

Video instance segmentation based on temporal feature fusion

Huang Zetao, Liu Yang, Yu Chenglong, Zhang Jiajia, Wang Xuan, Qi Shuhan 1692

Medical Image Processing

CT image classification of liver tumors based on multi-scale and deep feature extraction

Mao Jingyi, Song Yuqing, Liu Zhe 1704

3D multi-scale deep convolutional neural networks in pulmonary nodule detection

Sun Huacong, Peng Yanjun, Guo Yanfei, Zhang Xiaoqing 1716

Diabetic retinopathy grading based on multi-channel attention

Gu Tingfei, Hao Pengyi, Bai Cong, Liu Ning 1726

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2021)07-1704-12

论文引用格式: Mao J Y, Song Y Q and Liu Z. 2021. CT image classification of liver tumors based on multi-scale and deep feature extraction. Journal of Image and Graphics, 26(07): 1704-1715 (毛静怡, 宋余庆, 刘哲. 2021. 多尺度深度特征提取的肝脏肿瘤 CT 图像分类. 中国图象图形学报, 26(07): 1704-1715) [DOI:10.11834/jig.200500]

多尺度深度特征提取的肝脏肿瘤 CT 图像分类

毛静怡, 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 肝脏肿瘤是人体最具侵袭性的恶性肿瘤之一, 传统的肿瘤诊断依靠观察患者的 CT (computed tomography) 图像, 工作量大时易造成疲劳, 难免会产生误诊, 为此使用计算机辅助的方法进行诊断, 但现有的深度学习方法中存在肿瘤分类准确率低、网络的特征表达能力和特征提取能力较弱等问题。对此, 本文设计了一种多尺度深度特征提取的分类网络模型。方法 首先在原始 CT 图像中选取感兴趣区域, 然后根据 CT 图像的头文件进行像素值转换, 并进行数据增强来扩充构建数据集, 最后将处理后的数据输入到本文提出的分类网络模型中输出分类结果。该网络通过多尺度特征提取模块来提取图像的多尺度特征并增加网络的感受野, 使用深度特征提取模块降低背景噪声信息, 并着重关注病灶区域有效特征, 通过集成并行的空洞卷积使得尺度多元化, 并将普通卷积用八度卷积替换来减少参数量, 提升分类性能, 最终实现了对肝脏肿瘤的精确分类。结果 本文模型达到了 87.74% 的最高准确率, 比原始模型提升了 9.92%; 与现有主流分类网络进行比较, 多项评价指标占优, 达到了 86.04% 的召回率, 87% 的精准率, 86.42% 的 F1 分数; 此外, 通过消融实验进一步验证了所提方法的有效性。结论 本文方法可以较为准确地对肝脏肿瘤进行分类, 将此方法结合到专业的医疗软件当中去, 能够为医生早期的诊断和治疗提供可靠依据。

关键词: 深度学习; 肝脏肿瘤分类; 多尺度特征; 特征提取; 空洞卷积

CT image classification of liver tumors based on multi-scale and deep feature extraction

Mao Jingyi, Song Yuqing*, Liu Zhe

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

Abstract: **Objective** Liver tumors are the most aggressive malignancies in the human body. The definition of lesion type and lesion period based on computed tomography (CT) images determines the diagnosis and strategy of the treatment, which requires professional knowledge and rich experience of experts to classify them. Fatigue is easily experienced when the workload is heavy, and even experienced senior experts have difficulty avoiding misdiagnosis. Deep learning can avoid the drawbacks of traditional machine learning that takes a certain amount of time to manually extract the features of the image and perform dimensionality reduction, and is capable of extracting high-dimensional features of an image. Using deep learn-

收稿日期: 2020-08-22; 修回日期: 2021-01-13; 预印本日期: 2021-01-20

* 通信作者: 宋余庆 yqsong@ujs.edu.cn

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)

ing to assist doctors in diagnosis is important. In the existing medical image classification task, the challenge of the low accuracy of tumor classification, the weak capability of the feature extraction, and the rough dataset still remain. To address these tasks, this study presents a method with a multi-scale and deep feature extraction classification network. **Method** First, we extract the region of interest (ROI) according to the contours of the liver tumors that were labeled by experienced radiologists, along with the ROI of healthy livers. The ROI is extracted to capture the features of the lesion area and surrounding tissue, which is relative to the size of the lesion. Due to the different sizes of the lesion area, the size of the extracted ROI is also different. Then, the pixel value is converted and data augmentation is performed. The dataset is Hounsfield windows, the range of CT values is $(-1\ 024, 3\ 071)$, and the range of digital imaging and communications in medicine (DICOM) image is $(0, 4\ 096)$. The pixel values of DICOM images have to be converted to CT values. First, we read `rescale_intercept` and `rescale_slope` from the DICOM header file, and then we use the formula to convert. Thereafter, we limit the CT values of liver datasets to $[-100, 400]$ Hounsfield HU to avoid the influence of the background noise of the unrelated organs or tissues. We perform several data augmentation methods such as flipping, rotation, and transforming to expand the diversity of the datasets. Then, these images are sent into the MD_SENet for classification. The MD_SENet network is a SE_ResNet-like convolution neural network that can achieve end-to-end classification. The SE_ResNet learns the important features automatically from each channel to strengthen the useful features and suppress useless ones. MD_SENet network is much deeper than SE_ResNet. Our contributions are the following: 1) Hierarchical residual-like connections are used to improve multi-scale expression and increase the receptive field of each network layer. In the study, the image features after 1×1 convolution layers are divided into four groups. Each group of features passes through the 3×3 residual-like convolution groups, which improves the multi-scale feature extraction of networks and enhances the acquisition of focus areas features. 2) Channel attention and spatial attention are used to further focus on effective information on medical images. We let the feature images first go through the channel attention module, then we multiply its input and output to go through the spatial attention module. Then, we multiply the output of the spatial attention module and its input, which can pay more attention to the features of the lesion area and reduce the influence of background noise. 3) Atrous convolutions connected in parallel which refer to the spatial pyramid pooling, then we use 1×1 convolution layers to strengthen the feature. Finally, we concatenate the output and use softmax in classification. In this way, we can expand the receptive field and increase the image resolution, which can improve the feature expression ability and prevent the loss of information effectively. 4) The ordinary convolution is replaced by octave convolution to reduce the number of parameters and improve the classification performance. In this study, we compared the results of DenseNet, ResNet, MnasNet, MobileNet, ShuffleNet, SK_ResNet, and SE_ResNet with those of our MD_SENet, all of which were trained on the liver dataset. During the experiment, due to the limitation of graphics processing unit (GPU) memory, we set a batch size of 16 with Adam optimization and learning rate of 0.002 for 150 epochs. We used the dataset in Pytorch framework, Ubuntu 16.04. All experiments used the NVIDIA GeForce GTX 1060 Ti GPU to verify the effectiveness of our proposed method. **Result** Our training set consists of 4 096 images and the test set consists of 1 021 images for the liver dataset. The classification accuracy of our proposed method is 87.74% and is 9.92% higher than the baseline (SE_ResNet101). Our module achieves the best result compared with the state-of-the-art network and achieved 86.04% recall, 87% precision, 86.42% F1-score under various evaluation indicators. Ablation experiments are conducted to verify the effectiveness of the method. **Conclusion** In this study, we proposed a method to classify the liver tumors accurately. We combined the method into professional medical software so that we can provide a foundation that physicians can use in early diagnosis and treatment.

Key words: deep learning; liver lesion classification; multi-scale features; feature extraction; dilated convolution

0 引言

肝癌即肝脏恶性肿瘤,分为原发性和继发性两大类,肝细胞癌是我国高发的原发性肝癌中的一种,

根据世界卫生组织的报告(Zuckerman 和吉千秋, 1983),每年全球近 70 万人死于肝癌。计算机断层扫描(computed tomography, CT)是检测和诊断恶性肿瘤患者的一种常见的医学成像方式,广泛应用于临床实验。诊断方法通常是通过观察患者的 CT 图

像来进行判断。但由于器官的复杂性及其周围区域对比度等原因,工作量大时可能造成误诊。因此,如何利用计算机辅助实现对肝脏肿瘤的精准分类依然是一项非常具有挑战性的工作。

医学图像的分类方法按图像特征的提取方式分为基于传统机器学习的分类方法和基于深度学习的分类方法。基于传统机器学习的分类方法主要是运用机器学习进行分类,通常是利用 CT 图像的灰度值属性人为手动提取并研究病灶区域的统计量特征,然后构建分类模型,最后实现对图像的分类。Al-Kadi(2008)分别使用 3 种基于统计以及两种基于模型的方法从图像中提取纹理特征,并单独和成对使用每种方法后运用朴素贝叶斯对提取到的纹理特征进行分类。聂亚娜(2012)对通过传统方法计算得到的 4 类图像特征进行组合来获取综合特征,最后利用支持向量机(support vector machine, SVM)进行分类,确定最优特征组合。Zou(2015)先使用灰度共生矩阵提取图像的纹理特征并进行降维,再通过遗传算法构建决策树用来分离数据的模糊区域和提升识别率,并对决策树的每个节点都使用 SVM 训练分类器,最后由分类器分类。传统机器学习方法普遍存在以下问题:1)对图像的纹理、形状等方面进行特征提取,花费时间长且对有效特征的选取需慎重考虑;2)不同降维方法花费的计算时间与得到的效果不一致,降维方法不当可能造成数据冗余;3)缺乏对病灶区域深度特征的研究,同时由于病灶区域各有差异,过于微小的病灶区域很难通过传统特征进行分类。

基于深度学习的方法解决了传统机器学习方法需要花费时间手动提取图像特征并选择合适方法进行降维的问题,同时能够获得高级特征。近年来,使用深度学习解决医学图像分类问题已成为主流手段。Zhang 等人(2019)提出一种协同的深度学习模型(synergic deep learning, SDL)并使用多个深度卷积神经网络(deep convolutional neural networks, DCNNs)(Litjens 等,2017)解决图像在类内的差异性和类间的相似性问题。首先将每一对 DCNNs 提取的图像特征并联并作为 SDL 的输入,通过全连接预测输入图像是否为同一类别。当一对 DCNNs 中有一个 DCNNs 分类正确,则另一个分类错误会导致协同错误,并作为额外力量来更新模型。使用二分类数据集,并将图像统一为 224×224 像素大小进行实

验,分类准确率比基准方法提升了 2.1%。为了区别存在较大差异的囊肿和转移,Romero 等人(2019)提出一种用于肝脏病变分类的端到端的判别性网络框架,结合使用 InceptionV3(Szegedy 等,2016)从不同大小的卷积中提取特征,并迁移使用了在 ImageNet 上预训练的权重,最后利用并行处理的池化操作和辅助分类器(auxiliary classifiers)来提高模型的收敛性。Ghoneim 等人(2020)将图像输入到卷积神经网络中提取深层次的图像特征,然后使用极限学习(extreme learning machine, ELM)的分类器对输入图像进行分类,最后对网络进行微调。以上方法主要通过提取医学图像病灶区域的特征进行分类,存在以下不足:1)网络训练过程中未有效利用病灶区域的上下文信息和学习到的不同病灶区域的深度特征;2)只提取了图像的整体特征,没有着重关注病灶区域的局部特征;3)受不相关信息干扰,没有关注图像中的辅助判断信息;4)训练过程中造成一定位置、细节等底层特征信息丢失,降低了分类精度。

对医学图像分类任务,如何根据病灶区域周边的上下文信息对病灶特征进行补充加强,聚焦病灶区域本身,全面深度地提取病灶区域的整体以及局部特征,同时降低背景噪声带来的影响,避免细节以及病变区域边缘特征信息的丢失,是值得深入研究的方向。为此,本文提出一种多尺度深度特征提取的肝脏肿瘤 CT 图像分类方法,主要有以下贡献:1)引入多尺度特征提取模块提取图像的多尺度特征,深度挖掘病灶区域与周边组织的上下文信息增强网络的特征提取能力;2)使用深度特征提取模块从通道和空间两方面着重关注病灶区域的局部特征,削弱不相关信息影响,加强对病灶区域辨识能力;3)运用空洞卷积并联构造增强特征提取模块,在不损失图像特征信息前提下提升分类精度;4)运用卷积替换策略在减少参数量的同时加强分类性能。

1 本文方法

本文使用 SE_ResNet101(Hu 等,2018)作为基础网络架构,提出了一种多尺度深度特征提取的腹部肿瘤 CT 图像分类方法(multi-scale and deep feature extraction based on SENet, MD_SENet),具体步骤如下:

1) 通过提升网络的多尺度表达能力以及增加网络的感受野来加强病灶区域上下文之间的联系;

2) 通过加入注意力机制模块来着重提取病灶区域的局部特征并缓解背景噪声影响;

3) 使用并联的空洞卷积来获得较大的图像感受野, 在提取图像多尺度特征的同时保留更多原始细节特征, 提高分类准确率;

4) 将网络中普通卷积用八度卷积进行替换, 达到减少参数量、提升分类性能的目的。

1.1 多尺度特征提取模块

医学图像中的病灶区域在单幅图像中以不同尺寸出现, 依靠病灶区域的上下文信息可以更好地判断提取出的 ROI (region of interest) 归属于哪一类。对此, 采用 Gao 等人(2019)提出的 Res2Net 来感知不同尺度的信息, 提升网络多尺度表达能力, 增加每个网络层的感受野。原始的 bottleneck 主要通过 $1 \times 1, 3 \times 3, 1 \times 1$ 的卷积来进行特征映射。本文方法将原本的单个 3×3 卷积替换为多个 3×3 卷积组。具体是将之前卷积层提取到的粗粒度特征经过一个 1×1 卷积后分为 s 个部分, 每个部分代表一个特征子集 a_i , 且每个特征子集 a_i 具有相同的空间大小。然后将原始的 3×3 卷积用更小的、以残差方式连接的 3×3 卷积组进行了替换, 用 $m_i()$ 进行表示。为了在减少参数量的同时增加对特征的重复利用, 省略了 a_1 后面的 3×3 卷积。令 b_i 为 $m_i()$ 的输出, $m_{i-1}()$ 的输出与特征子集 a_i 相加后送入 $m_i()$ 。因此, b_i 表示为

$$b_i = \begin{cases} a_i & i = 1 \\ m_i(b_{i-1} + a_i) & 1 < i \leq s \end{cases} \quad (1)$$

式中, i 为整数且 $i \in \{1, 2, \dots, s\}$ 。最后将输出 b_i 进行 concat 之后通过一个 1×1 卷积。本文取 $s=4$, 即将特征图平均分为 4 部分, 此时获得最佳性能, 能够在不同尺度下获得更多的特征信息, 并对其进行高效处理。

1.2 深度特征提取模块

为了增加特征的表现力, 在关注重要特征的同时抑制不重要特征, Woo 等人(2018)引入注意力模块, 从空间和通道两个维度生成注意力图。通道注意力模块与空间注意力模块以串联方式执行。通过空间注意力的特征与之前通过通道注意力模块的特征进行相乘来自适应特征细化, 具体为

$$F' = A_s(A_c(F) \otimes F) \otimes A_c(F) \otimes F \quad (2)$$

$$F = f^{1 \times 1}[b_1; b_2; b_3; b_4] \quad (3)$$

式中, $A_c \in \mathbf{R}^{C \times 1 \times 1}$ 表示通道注意力模块, $A_s \in \mathbf{R}^{1 \times H \times W}$ 表示空间注意力模块, $\otimes$ 表示元素之间 element-wise 点乘操作。 F' 表示通过通道和空间注意力模块后的输出特征。

在通道注意力模块中, 首先将特征图 $F \in \mathbf{R}^{C \times H \times W}$ 分别经过基于宽和高的全局平均池化和全局最大池化, 并经过一个多层感知器 (multi-layer perception, MLP), 然后将输出得到的特征进行基于点乘的加和操作, 最终通过激活函数生成通道注意力特征图, 并融合通道注意力模块的输入 $F \in \mathbf{R}^{C \times H \times W}$ 一起送入空间注意力模块。具体为

$$A_c(F) = \sigma(M(P_{GA}(F)) + M(P_{GM}(F))) \quad (4)$$

式中, P_{GA} 表示全局平均池化, P_{GM} 表示全局最大池化, M 表示多层感知器, σ 表示 sigmoid 激活函数。在空间注意力模块中, 本文将特征图经过池化操作后分别经过两个 1×1 的卷积, 再通过 sigmoid 激活函数进行激活, 这样能够在提升网络表达力的同时, 解决由于使用 MLP 而造成的输入向量时丢失空间信息的问题, 即式(4)可表示为

$$A_c(F) = \sigma(f^{1 \times 1}(f^{1 \times 1}(P_{GA}(F))) + f^{1 \times 1}(f^{1 \times 1}(P_{GM}(F)))) \quad (5)$$

将通道注意力的输出与其输入点乘后作为空间注意力模块的输入, 首先经过了基于通道的全局平均池化和全局最大池化, 再将通过这两个池化操作后得到的特征图进行 concat, 并且接着通过一个 7×7 的卷积进行聚合, 最后通过激活函数生成空间注意力特征图。具体为

$$G = A_c(F) \otimes F \quad (6)$$

$$A_s(G) = \sigma(f^{7 \times 7}[P_{GA}(G); P_{GM}(G)]) \quad (7)$$

式中, $f^{7 \times 7}$ 表示核为 7 的卷积, G 表示空间注意力模块的输入, 即通道注意力模块的输出与其输入进行 element-wise 点乘后得到的特征图。通过使用该模块, 能够更进一步地挖掘医学图像的有效信息而不受到背景噪声的干扰, 即能够聚焦所需要的详细病灶的深层次特征来进行正确分类, 然后送入增强特征提取模块。

1.3 增强特征提取模块

现有的分类任务网络结构往往都是通过卷积池化全连接层来输出分类预测结果。网络结构在训练

中若能获得不同大小的感受野,则能够捕捉不同尺度的信息,从而提升分类精度。

为了扩大感受野,获得多尺度特征,缓解空间信息的丢失,Chen 等人(2018)提出了空洞空间金字塔池化对特征进行增强。为了提升特征的分辨率,将空洞卷积并行连接。在空洞空间金字塔池化中,包括 1×1 卷积和 3 个 3×3 空洞卷积。先池化再卷积后,通过上采样恢复原分辨率,最后 concat 后送入 1×1 卷积。具体过程如图 1 所示。

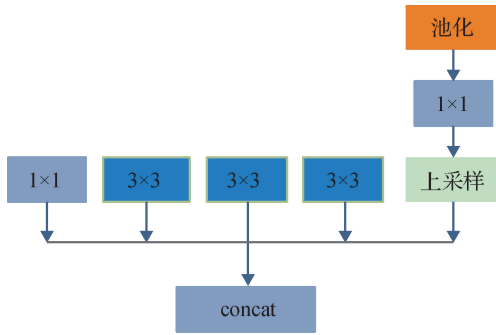


图 1 空洞空间金字塔

Fig. 1 Atrous spatial pyramid pooling

本文对 Chen 等人(2018)提出的空洞空间金字

塔进行改进,MD_SENet 的主干瓶颈层如图 2 所示。首先,将特征图并行地经过 1 个 1×1 卷积和 3 个 3×3 的空洞卷积,通过 1×1 卷积获取原始特征图,3 个 3×3 空洞卷积的膨胀率分别设置为 6、12、18 (Chen 等,2018),使用不同的膨胀率有利于考虑不同尺寸图像的多个不同比例,能够在得到图像不同尺度的同时提升网络的感受野。然后,将通过并行空洞卷积得到的特征分别经过 1×1 卷积来加强提取到的多尺度特征,并与平均池化后得到的特征图进行融合。将融合得到的输出经过 1×1 卷积,得到固定大小的输出 F'' 。此时,不但能够获得原始空洞空间金字塔具有的图像级特征,还能更加充分地挖掘病灶区域的语义信息。具体表示为

$$F'' = \sigma(f^{1 \times 1} [f^{1 \times 1}(P_A(F'))]; f^{1 \times 1}(f^{3 \times 3, d=6}(F')); f^{1 \times 1}(f^{3 \times 3, d=12}(F')); f^{1 \times 1}(f^{3 \times 3, d=18}(F')); P_A(F')]) \quad (8)$$

式中, P_A 表示平均池化, d 表示膨胀率 (dilation rate)。为了能够有效利用之前病灶区域的全局以及局部的特征信息,本文将融合后经过 1×1 卷积得到的输出与之前输入的特征进行一个相加的操作来防止梯度消失。

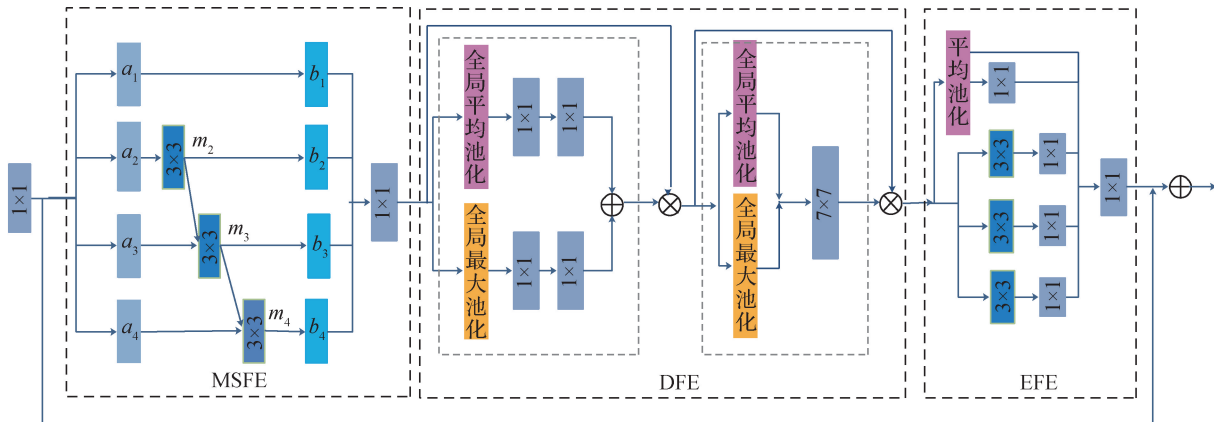


图 2 详细瓶颈结构图

Fig. 2 Detail of the bottleneck

1.4 整体模块设计以及网络架构

本文算法流程如图 3 所示,网络整体由 3 个 3×3 卷积、4 个 block、增强特征提取模块 (enhanced feature extraction module, EFE) 以及 softmax 构成。原始的 SE_ResNet101 的输入层是一个 7×7 卷积,为了与 7×7 卷积获得相同感受野并且捕获更多特征,将其替换为 3 个 3×3 卷积 (He 等,2019),这样能够减少参数量,学习到更具有区分性的病灶边缘

区域特征。为了更好地深层次提取病灶区域的多尺度及其深度特征,扩大卷积层感受野,避免损失图像特征和缓解背景噪声的影响,本文结合多尺度特征提取模块 (multi-scale feature extraction module, MSFE)、深度特征提取模块 (deep feature extraction module, DFE) 以及增强特征提取模块对 SE_ResNet (Hu 等,2018) 的网络瓶颈部分进行了优化,具体为

$$B = \sigma(f_c(\delta(f_c(P_{GA}(F''))))) \otimes F'' + X_i \quad (9)$$

式中, B 表示通过 bottleneck 后的输出, f_c 表示全连接层, δ 表示 ReLU 激活函数, $X_i \in \mathbf{R}^{C \times H \times W}$ 表示送入 bottleneck 之前的原始特征图。

本文优化的 bottleneck 由多尺度特征提取模块、深度特征提取模块以及增强特征提取模块构成。训练时使用八度卷积 (Chen, 2019) 代替了普通卷积,

用于降低资源消耗, 提升分类精度。算法的网络结构及其参数设置如表 1 所示, 其中 C (convolution layer)、AP (average pooling)、MP (max pooling)、FC (fully connected layer) 分别表示卷积层、平均池化、最大池化和全连接层, 后面的参数依次为核大小、步长、填充值。

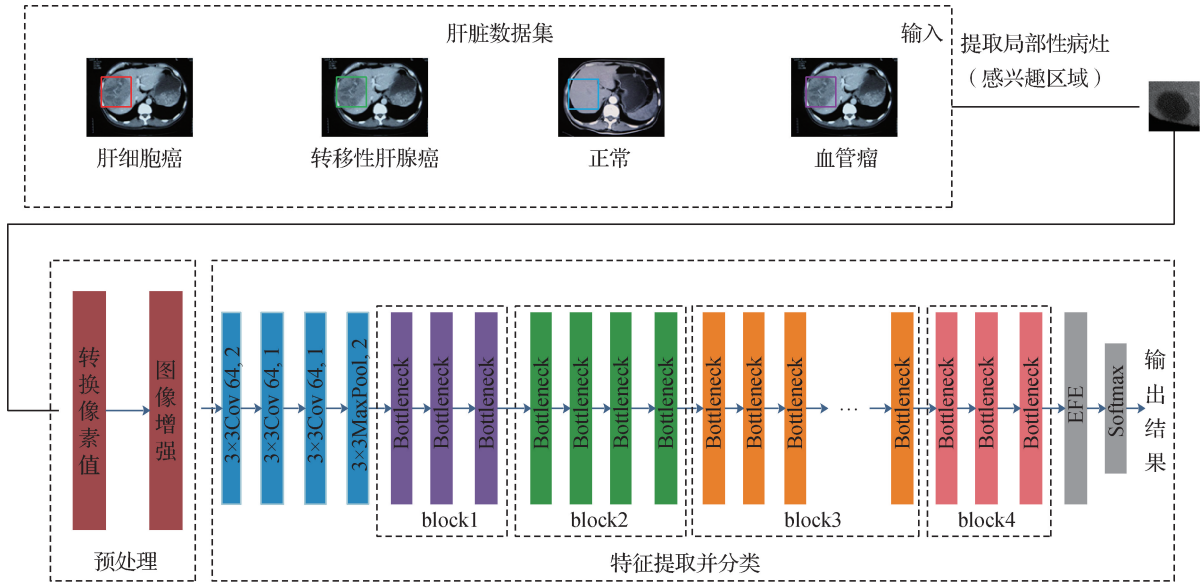


图 3 本文算法演示图

Fig. 3 Overview of the proposed method

表 1 网络结构及其参数设置

Table 1 Network structure and parameter setting

MD_SENet 网络结构及参数				
名称				
前网络层	$\begin{bmatrix} C: 3, S: 2, P: 1 \\ \text{BatchNorm} \end{bmatrix}$	$\begin{bmatrix} C: 3, S: 1, P: 1 \\ \text{BatchNorm} \end{bmatrix}$	$\begin{bmatrix} C: 3, S: 1, P: 1 \\ \text{BatchNorm} \\ \text{ReLU} \end{bmatrix}$	$\begin{bmatrix} \text{MP}: 3, S: 2, P: 1 \end{bmatrix}$
瓶颈模块	$\begin{bmatrix} \text{MSFE} \\ \text{DFE} \\ \text{EFE}: 256 \end{bmatrix}$	$\begin{bmatrix} \text{MSFE} \\ \text{DFE} \\ \text{EFE}: 512 \end{bmatrix}$	$\begin{bmatrix} \text{MSFE} \\ \text{DFE} \\ \text{EFE}: 1\,024 \end{bmatrix}$	$\begin{bmatrix} \text{MSFE} \\ \text{DFE} \\ \text{EFE}: 2\,048 \end{bmatrix}$
后网络层	$\begin{bmatrix} \text{EFE}: 4 \end{bmatrix}$	$\begin{bmatrix} \text{Softmax} \end{bmatrix}$		

2 实验结果与分析

2.1 数据集

实验使用的数据集为 2015—2018 年在镇江市第一人民医院采集的 5 127 幅腹部 CT 扫描图像。CT 扫描采用 5~7 mm 切片, 512 × 512 像素矩阵, 平面分辨率为 0.57~0.89 mm。数据集包括 190 例患者门静脉 CT 扫描的 4 种病灶类型, 其中 35 例转

移性肝腺癌 (metastasis, MET)、40 例血管瘤 (hemangioma, HEM)、62 例肝细胞癌 (hepatocellular carcinoma, HCC) 和 53 例正常组织 (healthy)。放射科专家对每个病灶的边缘进行了标记, 并通过活检或临床随访确定相应诊断, 数据样本如图 4 所示。

2.2 训练设置及实验设备

2.2.1 预处理

将所有数据样本进行预处理, 具体步骤如下:

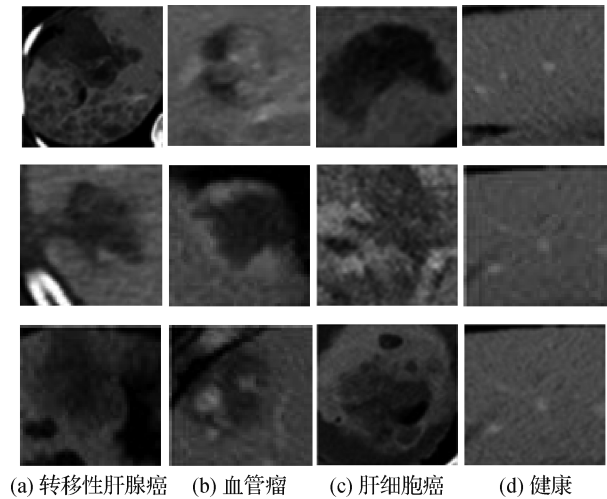


图4 肝脏数据样本

Fig. 4 Typical samples of liver dataset

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

1)感兴趣区域(ROI)提取。根据经验丰富的影像科医生标记的肝脏肿瘤轮廓,提取出所需要的ROI,为了增加样本的多样性,同时从健康的肝脏组织中由医生勾勒出健康肝脏的感兴趣区域。

2)转换像素值。CT值的单位是亨斯菲尔德(Hounsfield, HU),范围是(−1 024, 3 071),反映了组织对X射线的吸收程度。医学数字图像与通讯(digital imaging and communications in medicine, DICOM)格式的图像范围通常是(0, 4 096)。转换时需要从DICOM头文件中读取RI(rescale intercept)和RS(rescale slope)这两个标签信息。而CT值与像素值之间的转换关系可表达为

$$HU = PV \times RS + RI \quad (10)$$

式中, PV表示像素值,在本次实验中,针对本文所使用的数据,取RS = 1, RI = −1 024。并将肝脏的范围限制为[−100, 400](Christ等, 2016; 刘哲等, 2018)。

3)图像增强。为了加强病灶特征,提升网络的泛化能力,将提取的ROI通过随机翻转、填充、裁剪和仿射变换等方式进行数据增强。

2.2.2 数据划分与参数设置

预处理后,将数据集随机分为训练集和测试集两部分,其中训练集占样本总数的80%,测试集占样本总数的20%,并将样本裁剪为64 × 64像素大小,数据分配如表2所示。实验中,由于GPU(graphics processing unit)显存的限制,本文将batch_

size设置为16,采用Adam(Kingma和Ba, 2015)优化算法,在150个epoch中的初始学习率(learning_rate)设置为0.002,指数衰减率(gamma)设置为0.98,下降周期设置为1。将此数据集在pytorch框架下进行实验,所有实验都使用NVIDIA GeForce GTX 1060 Ti GPU以验证方法的有效性。

表2 肝脏数据集分布划分

Table 2 Distribution of the liver dataset

类型	/幅			
	转移性肝腺癌	血管瘤	肝细胞癌	健康
训练集	1 303	659	1 466	668
测试集	325	164	366	166
合计	1 628	823	1 832	834

2.3 评价指标

采用准确率(accuracy)、召回率(recall)、精准率(precision)、F1分数(F1-score)和ROC曲线(receiver operating characteristic curve)图这几个评价指标来对分类结果进行评价,具体定义为

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

$$R = \frac{TP}{TP + FN} \quad (12)$$

$$P = \frac{TP}{TP + FP} \quad (13)$$

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (14)$$

$$TPR = \frac{TP}{TP + FN} \quad (15)$$

$$FPR = \frac{FP}{TN + FP} \quad (16)$$

式中, A, R, P, F1分别表示准确率、精准率、召回率和F1分数。TP表示正确预测肿瘤的样本数, TN表示正确预测背景的样本数, FP表示将背景预测为肿瘤的样本数, FN表示将肿瘤预测为背景的样本数。指标在范围[0, 1]之间,数值越大,效果越佳。FPR为假阳率(false positive rate, FPR),指在所有非此类别下预测为正确的概率。TPR为真阳率(true positive rate, TPR),指在该类别下真正预测出来的概率。ROC是根据FPR和TPR绘制的曲线图,称为受试工作者曲线。AUC(area under the curve)是ROC曲线下的面积(Romero等, 2019), AUC越接近1,检验方法的真实性越高。

2.4 实验分析

2.4.1 与基准实验比较

采用 SE_ResNet101 (Hu 等, 2018) 作为基准实验, 在此基础上增加了多尺度特征提取模块、深度特征提取模块和增强特征提取模块, 并使用八度卷积 (Chen, 2019) 替换普通卷积。图 5 是本文方法与基准方法对肝细胞癌 (HCC)、转移性肝腺癌 (MET)、正常 (normal) 和血管瘤 (HEM) 等情况的混淆矩阵 (confusion matrix) 比较。其中, 纵轴代表真实标签 (true label), 横轴代表预测标签 (predict label), 每

行表示在真实标签情况下预测正确的概率, 即召回率 (recall), 每列表示在预测标签情况下正确预测的概率, 即精确率 (precision)。从图 5 可以看出, 本文方法对肝脏每类病灶的召回率和精确率均高于或等于基准实验, 且每一类样本均衡, 表明使用本文算法能够实现更好的整体分类, 增强网络的特征提取能力, 缓解背景噪声造成的影响, 加强对图像特征的利用。但是由于病变区域的相似性太过于接近, 导致在区分病变类型时仍然存在一些错误分类。

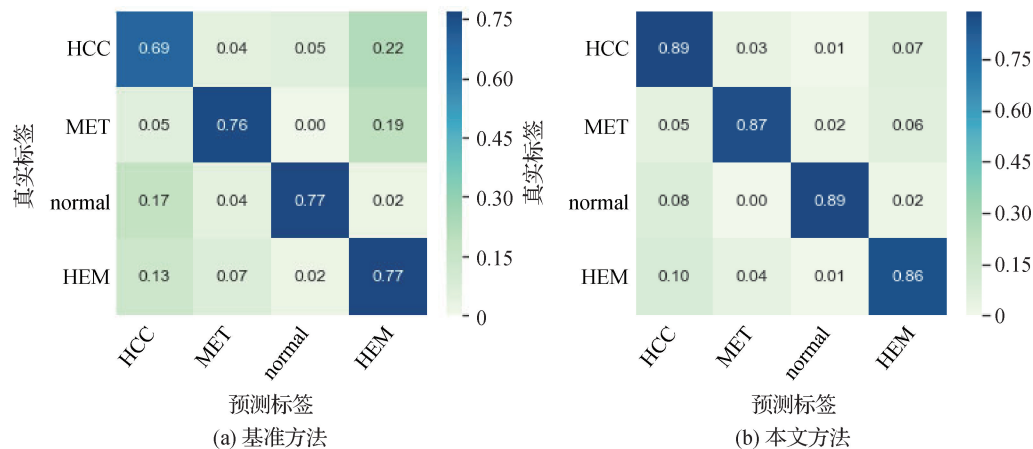


图5 本文方法与基准方法的混淆矩阵比较

Fig. 5 Comparison of confusion matrix between baseline method and ours ((a) baseline method; (b) ours)

本文方法与基准模型对不同类型肝脏病变及整体 ROC 曲线如图 6 和图 7 所示。可以看出, 本文方法获得的 ROC 曲线图比基准表现更加稳定和光滑, 并具有更大的 AUC, 表明本文方法具有更好的分类效果。

2.4.2 消融实验

本文方法主要包括多尺度特征提取模块 (MSFE)、深度特征提取模块 (DFE)、增强特征提取模块 (EFE) 以及卷积替换策略 (convolution substitution strategy, CSS) 等 4 部分, 在同一数据集上进行消融实验来验证每一部分的有效性。结果如表 3 所示。可以看出, 通过使用 MSFE, 准确率提升了 2.92%, 表明加强网络的多尺度提取能力和获取更多的上下文信息能够有效提升分类精度。MSFE 与 DFE 结合, 相比仅使用 MSFE, 准确率提升了 2.94%, 可知使用深度特征提取模块能够加强对病灶区域特征的有效利用, 降低背景噪声对分类任务造成的影响。MSFE、DFE 和 EFE 三者结合与 MSFE 和 DFE 二者

结合相比较, 准确率提升了 2.7%。EFE 模块使得网络在训练过程中扩大了感受野, 增加了多尺度特性, 因此降低了精度损失。将网络结合 MSFE、DFE、EFE 和 CSS, 精度提升了 1.36%, 并且参数量减少了 3.67 M, 证明结合 CSS 能够在减少冗余的同时有效提升分类精度。

为了验证本文改进的增强特征提取模块 (EFE) 的分类性能, 在 SE_ResNet101 (Hu 等, 2018) 和 SK_ResNet101 (Li 等, 2019) 中进行嵌入, 结果如表 4 所示。可以看出, 使用 EFE 模块后, 两种模型准确率分别提高了 4.28% 和 3.11%, 表明添加 EFE 模块能够增加多尺度信息, 扩大网络的感受野, 缓解特征信息的丢失, 提升分类准确率。

2.4.3 与经典网络比较

为进一步验证分类性能, 将本文方法与 DenseNet (Huang 等, 2017)、ResNet101 (He 等, 2016)、MnasNet (Tan 等, 2018)、MobileNet2 (Sandler 等, 2018)、ShuffleNetV2 (Zhang 等, 2017)、SK_Res-

Net101 (Li 等, 2019) 和 SE_ResNet101 (Hu 等, 2018) 等经典网络在同一数据集下进行比较, 结果如表 5 所示。可以看出, 通过加强病灶区域的多尺度以

及深度特征, 能够削弱背景噪声的影响, 增加有用信息的利用, 在不损失精度的前提下提高分类性能。

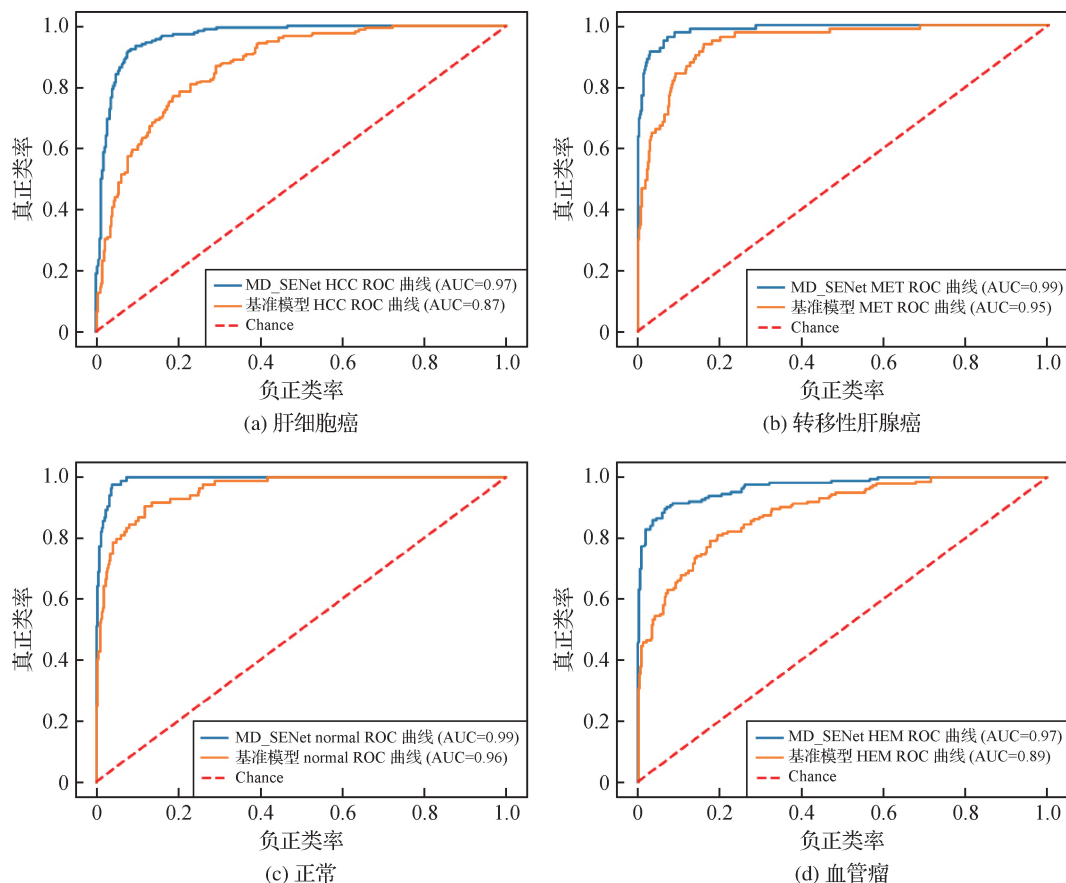


图 6 不同类型肝脏病变中本文方法与基准模型之间的 ROC 曲线对比图

Fig. 6 Comparison of ROC curves of each type between backbone model and proposed model

((a) HCC; (b) MET; (c) normal; (d) HEM))

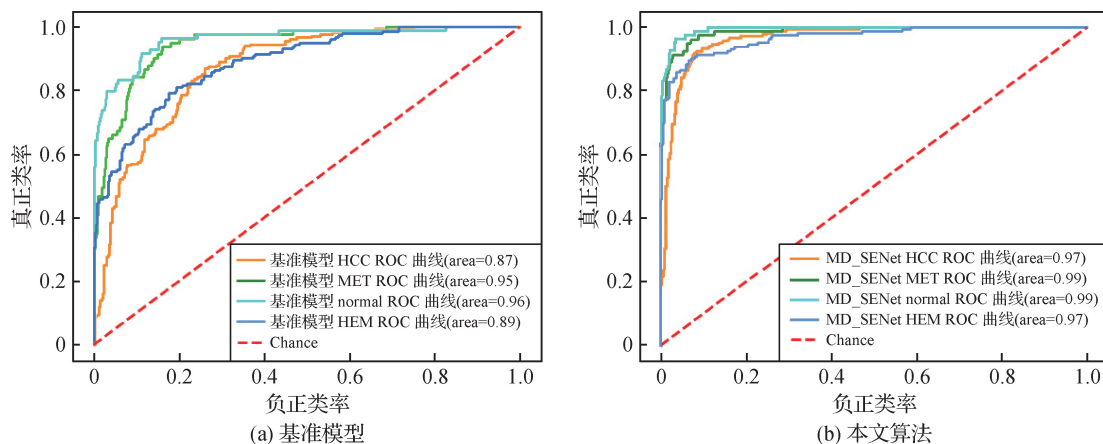


图 7 本文方法与基准模型对不同类型肝脏病变的整体 ROC 曲线对比图

Fig. 7 Comparison of ROC curves between backbone model and proposed model ((a) baseline method; (b) ours)

表3 肝脏数据消融实验

Table 3 Liver dataset with ablation experiments

方法	准确率/%	召回率/%	精准率/%	F1 分数/%	AUC/%	参数量/M
SE_ResNet101	77.82	77.78	78.75	78.75	91.00	49.12
MSFE	80.74	81.76	79.35	80.34	94.00	34.53
MSFE + DFE	83.68	83.00	84.36	83.85	95.00	56.14
MSFE + DFE + EFE	86.38	86.75	86.07	86.38	96.00	58.83
MSFE + DFE + EFE + CSS	87.74	86.04	87.00	86.42	96.00	55.16

注:加粗字体为各列最优结果。

表4 肝脏数据集上基于 EFE 模块的消融实验

Table 4 Ablation experiment based on EFE module on the liver dataset

方法	准确率	召回率	精准率	F1 分数	ACC 增幅
SE_ResNet101	77.82	77.78	78.75	78.75	4.28
SE_ResNet101 + EFE	82.10	81.24	82.05	81.59	
SK_ResNet101	81.32	82.22	81.03	81.50	3.11
SK_ResNet101 + EFE	84.43	84.70	84.88	84.76	

注:加粗字体表示添加 EFE 模块后的波动值。

表5 肝脏数据集分类效果

Table 5 Classification performance of the liver dataset

方法	准确率	召回率	精准率	F1 分数	AUC
DenseNet	79.58	79.17	78.33	79.50	93.00
ResNet101	77.63	76.40	77.47	76.28	93.00
MnasNet	73.54	73.57	73.62	73.62	91.00
MobileNet2	81.13	81.05	80.31	80.31	95.00
ShuffleNetV2	76.07	76.56	76.89	76.89	91.00
SK_ResNet101	81.32	82.22	81.03	81.50	91.00
SE_ResNet101	77.82	77.78	78.75	78.75	91.00
本文	87.74	86.04	87.00	86.42	96.00

注:加粗字体为各列最优结果。

图8为本文方法及现有经典分类算法的整体ROC曲线图。可以看出,在所有经典分类模型中,本文方法的ROC曲线波动幅度较小且更平滑,获得了比其他经典分类网络模型更大的AUC,进一步证明了本文算法的优越性。

3 结 论

针对现有分类方法中存在的对上下文信息的不

充分挖掘、受到背景噪声的影响以及损失图像特征信息等问题,本文提出了一种多尺度深度特征提取的肝脏肿瘤CT图像分类方法,从以下4个方面进行了改进:1)使用多尺度特征提取模块来加强病灶区域上下文信息之间的联系,在增加感受野的同时能够充分挖掘语义信息;2)使用深度特征提取模块来加强病灶区域特征的同时降低背景噪声的影响,深度关注有用的病灶信息;3)将空洞卷积并行连接,从而能够对不同尺度的特征图进行采样,扩大网

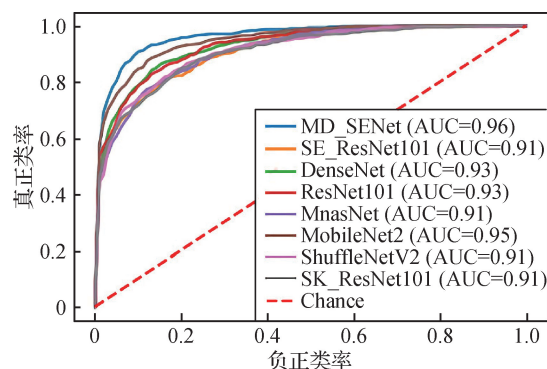


图8 不同算法下的ROC曲线比较

Fig. 8 Comparison of ROC curves among different methods

络的感受野,提升对原始图像的特征利用;4)将普通卷积替换为八度卷积,降低参数量的同时提升分类效果。通过以上几方面的改进,实现对肝脏肿瘤精准分类。本文方法的优越性在肝脏数据集上得到了验证,并在多项评价指标下获得了最佳性能。与主流经典网络相比,本文方法分类效果优于经典网络,未来将延伸到更多医疗领域,作为一种早期检测病变的方法,辅助医生进行诊断和治疗,具有深远意义。

本文方法仍存在不足,由于器官在不同时期如静脉期、动脉期、延迟期与平扫采集到的图像CT值有所不同,会导致分类结果下降,因此如何利用不同时期CT值准确对肿瘤精准分类还有待研究。

参考文献 (References)

- Al-Kadi O S. 2008. Combined statistical and model based texture features for improved image classification//Proceedings of the 4th IET International Conference on Advances in Medical, Signal and Information Processing. MEDSIP, Italy; IEEE; 175-178 [DOI: arXiv: 1512.08814]
- Chen L C, Papandreou G, Kokkinos I, Murphy K and Yuille A. 2018. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4): #834 [DOI: arXiv: 1606.00915v2]
- Chen Y P, Fang H Q, Xu B, Yan Z C, Yannis, Marcus, Yan S C and Feng J S. 2019. Drop an octave: reducing spatial redundancy in convolutional neural networks with octave convolution//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South); IEEE; 3434-3443 [DOI: 10.1109/ICCV.2019.00353]
- Christ P F, Elshaer M E A, Ettlinger F, Tatavarty S, Bickel M, Bilic P, Rempfler M, Armbruster M, Hofmann F, 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. Granada, Spain; Springer Verlag; 415-423 [DOI: 10.1007/978-3-319-46723-8_48]
- Gao S H, Cheng M M, Zhao K, Zhang X Y, Yang M H and Philip H. S. T. 2019. Res2 Net: a new multi-scale backbone architecture. IEEE Transactions on Pattern Analysis and Machine Intelligence. #2938758 [DOI: 10.1109/tpami.2019.2938758]
- Ghoneim A, Muhammad G and Hossain M S. 2020. Cervical cancer classification using convolutional neural networks and extreme learning machines. Future Generation Computer Systems, 102: 643-649 [DOI: https://doi.org/10.1016/j.future.2019.09.015]
- 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 (CVPR). Las Vegas, 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 of the 32nd IEEE Conference on Computer Vision and Pattern Recognition. Los Angeles, USA; IEEE; 558-567
- Hu J, Shen L and Sun G. 2018. Squeeze-and-excitation networks//Proceedings of the 31st IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA; IEEE; 7132-7141 [DOI: 10.1109/CVPR.2018.00745]
- Huang G, Liu Z, Maaten L V D, Weinberger and Kilian Q. 2017. Densely connected convolutional networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA; IEEE; 2261-2269 [DOI: 10.1109/CVPR.2017.243]
- Kingma D and Ba J. 2015. Adam: a method for stochastic optimization//Proceedings of the 3rd International Conference on Learning Representations. San Diego, USA; ICLR; 1-15
- 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]
- Li X, Wang W H, Hu X L and Yang J. 2019. Selective kernel networks//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). California, USA; IEEE; 510-519 [DOI: 10.1109/CVPR.2019.00060]
- Litjens G, Kooi T, Bejnordi B E, Setio A A A, Ciampi F, Ghafoorian M, Jeroen A M, Ginneken B and Sanchez C. 2017. A survey on deep learning in medical image analysis. Medical Image Analysis, 42: 60-88 [DOI: 10.1016/j.media.2017.07.005]

- Nei Y N. 2012 Research on Application of Support Vector Machine in Liver B Ultrasound Images Classification. Xi'an; Xi'an University of Science and Technology (聂亚娜. 2012. 支持向量机在肝脏 B 超图像分类中的应用研究. 西安: 西安科技大学)
- 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]
- Sandler M, Howard A, Zhu M, Zhmoginov A and Chen L C. 2018. MobileNetV2: Inverted residuals and linear bottlenecks//Proceedings of 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA; IEEE: 4510-4520 [DOI: 10.1109/CVPR.2018.00474]
- Szegedy C, Ioffe S, Vanhoucke V and Alemi, A. 2016. Inception-v4, Inception-ResNet and the impact of residual connections on learning//Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, USA; AAAI: 4278-4284
- Tan M X, Chen B, Pang R M and Vasudevan V. 2018. MnasNet: platform-aware neural architecture search for mobile//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA; IEEE: 2815-2823 [DOI: arXiv: 1807.11626]
- 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 Verlag: 3-19 [DOI: 10.1007/978-3-030-01234-2_1]
- Zhang J P, Xie Y T, Wu Q and Xia Y. 2019. Medical image classification using synergic deep learning. Medical Image Analysis. 54: 10-19 [DOI: 10.1016/j.media.2019.02.010]
- Zhang X Y, Zhou X Y, Lin M X and Sun J. 2017. ShuffleNet: an extremely efficient convolutional neural network for mobile devices//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA; IEEE: 6848-6856 [DOI: 10.1109/CVPR.2018.00716]
- Zou L. 2015 Research on Medical Image Classification Based On Twin Support Vector Machine. Gansu; Northwest Normal University (邹丽, 2015. 基于对支持向量机的医学图像分类研究. 甘肃: 西北师范大学)
- Zuckerman A J and Ji Q Q. 1983. Prevention of primary liver cancer—report on a meeting of the world health organization's scientific team. Guangzhou post-medical information. 1983 (Z1): 131-132 (Zuckerman A J, 吉千秋. 1983. 原发性肝癌的预防—世界卫生组织科研组的一次会议报告. 广后医学资料, 1983 (Z1): 131-132)

作者简介



毛静怡, 1996 年生, 女, 硕士研究生, 主要研究方向为医学图像处理。

E-mail: 2221808009@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, 教授, 主要研究方向为医学图像处理与分析。

E-mail: lzhe@ujs. edu. cn