



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

09

VOL.26

ISSN1006-8961
CN11-3758/TB



“深度学习+智慧医疗”专刊



第26卷第9期（总第305期）
2021年9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

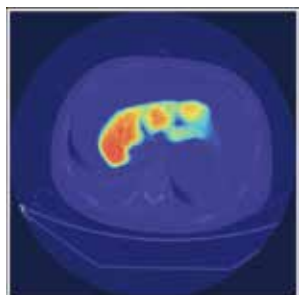
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

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

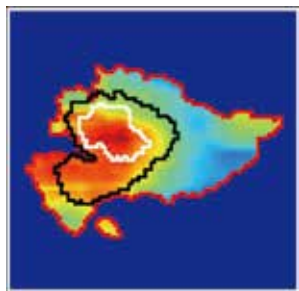
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



肺部影像解剖结构分割数据集及应用(第2111页)



LFSCA-UNet: 基于空间与通道注意力机制的肝纤维化区域分割网络(第2121页)



语义拉普拉斯金字塔多中心乳腺肿瘤分割网络(第2193页)

编者按 I

综述

深度医学图像配准研究进展: 迈向无监督学习

马露凡, 罗凤, 严江鹏, 徐哲, 罗捷, 李秀 2037

U-Net网络医学图像分割应用综述

周涛, 董雅丽, 霍兵强, 刘珊, 马宗军 2058

医学图像深度学习技术: 从卷积到图卷积的发展

唐朝生, 胡超超, 孙君顶, 司马海峰 2078

多尺度变换像素级医学图像融合: 研究进展、应用和挑战

周涛, 刘珊, 董雅丽, 霍兵强, 马宗军 2094

数据集论文

肺部影像解剖结构分割数据集及应用

覃文军, 李小硕, 周庆华, 刘盼, 杨金柱 2111

计算机断层扫描图像

LFSCA-UNet: 基于空间与通道注意力机制的肝纤维化区域分割网络

陈弘扬, 高敬阳, 赵地, 吴忌, 陈金军, 全显跃, 李欣明, 薛峰, 周沐瑶, 柏冰冰 2121

融合上下文和多尺度特征的胸部多器官分割

吉淑滢, 肖志勇 2135

注意力机制下密集空洞卷积的肺部图像分割

郭宁, 柏正尧 2146

融合注意力机制和特征金字塔网络的CT图像肺结节检测

张福玲, 张少敏, 支力佳, 周涛 2156

改进Faster R-CNN模型的CT图磨玻璃密度影目标检测

杨淑莹, 邓东升, 郑清春 2171

融合注意力机制与可变形卷积的多尺度骨病变检测

方成, 柏正尧 2181

磁共振图像

语义拉普拉斯金字塔多中心乳腺肿瘤分割网络

王黎, 曹颖, 郭顺超, 唐雷, 鄧子翔, 王荣品, 王丽会 2193

引入注意力机制和多视角融合的脑肿瘤MR图像U-Net分割模型

罗恺锴, 王婷, 叶芳芳 2208

集成注意力增强和双重相似性引导的多模态脑部图像配准

田梨梨, 程欣宇, 唐堃, 张健, 王丽会 2219

LRUNet: 轻量级脑肿瘤快速语义分割网络

何康辉, 肖志勇 2233

自适应多模态特征融合胶质瘤分级网络

王黎, 曹颖, 田梨梨, 陈祈剑, 郭顺超, 张健, 王丽会 2243

研究应用

面向运动想象脑电图识别的镜卷积神经网络

罗靖, 王耀杰, 刘光明, 王晓帆, 鲁晓峰, 黑新宏 2257

先验权重共享码本下内窥镜图像大肠病变分类

朱霆威, 李胜, 何熊熊 2270

融合空洞卷积与注意力的胃癌组织切片分割

陈颖锲, 李晗, 周雪婷, 万程 2281

视觉显著性的眼底图像视盘检测

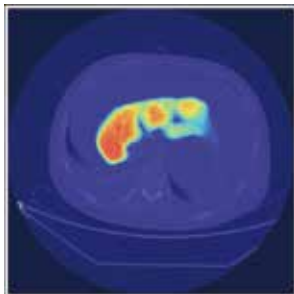
吕鹏飞, 王莹, 王思齐, 于晓升, 吴成东 2293

CONTENTS

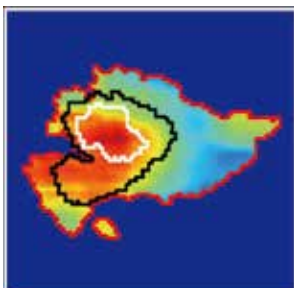
JOURNAL OF IMAGE AND GRAPHICS



Pulmonary image anatomical structure segmentation dataset and applications(P2111)



LFSCA-UNet: liver fibrosis region segmentation network based on spatial and channel attention mechanisms(P2121)



Semantic Laplacian pyramids network for multicenter breast tumor segmentation(P2193)

Review

- Deep-learning based medical image registration pathway: towards unsupervised learning
Ma Lufan, Luo Feng, Yan Jiangpeng, Xu Zhe, Luo Jie, Li Xiu 2037
- U-Net and its applications in medical image segmentation:a review
Zhou Tao, Dong Yali, Huo Bingqiang, Liu Shan, Ma Zongjun 2058
- Deep learning-based medical images analysis evolved from convolution to graph convolution
Tang Chaosheng, Hu Chaochao, Sun Junding, Sima Haifeng 2078
- Research on pixel-level image fusion based on multi-scale transformation: progress application and challenges
Zhou Tao, Liu Shan, Dong Yali, Huo Bingqiang, Ma Zongjun 2094

Dataset

- Pulmonary image anatomical structure segmentation dataset and applications
Tan Wenjun, Li Xiaoshuo, Zhou Qinghua, Liu Pan, Yang Jinzhu 2111

Computerd Tomography Image

- LFSCA-UNet: liver fibrosis region segmentation network based on spatial and channel attention mechanisms
Chen Hongyang, Gao Jingyang, Zhao Di, Wu Ji, Chen Jinjun, Quan Xian Yue, Li Xinming, Xue Feng, Zhou Muyao, Bai Bingbing 2121
- Integrated context and multi-scale features in thoracic organs segmentation
Ji Shuying, Xiao Zhiyong 2135
- The integration of attention mechanism and dense atrous convolution for lung image segmentation
Guo Ning, Bai Zhengyao 2146
- Detection of pulmonary nodules in CT images by combining an attention mechanism and a feature pyramid network
Zhang Fuling, Zhang Shaomin, Zhi Lijia, Zhou Tao 2156
- Ground-glass opacity target detection in CT scans based on improved Faster R-CNN model
Yang Shuying, Deng Dongsheng, Zheng Qingchun 2171
- Multi-scale bone lesion detection based on attention mechanism and deformable convolution
Fang Cheng, Bai Zhengyao 2181

Magnetic Resonance Image

- Semantic Laplacian pyramids network for multicenter breast tumor segmentation
WANG Li, CAO Ying, Guo Shunchao, Tang Lei, Kuai Zixiang, Wang Rongpin, Wang Lihui 2193
- U-Net segmentation model of brain tumor MR image based on attention mechanism and multi-view fusion
Luo Kaikai, Wang Ting, Ye Fangfang 2208
- Multimodal brain image registration with integrated attention augmentation and dual similarity guidance
Tian Lili, Cheng Xinyu, Tang Kun, Zhang Jian, Wang Lihui 2219
- LRUNet: a lightweight rapid semantic segmentation network for brain tumors
He Kanghui, Xiao Zhiyong 2233
- Adaptive multi-modality fusion network for glioma grading
Wang Li, Cao Ying, Tian Lili, Chen Qijian, Guo Shunchao, Zhang Jian, Wang Lihui 2243

Research and Application

- Mirror convolutional neural network for motor imagery electroencephalogram recognition
Luo Jing, Wang Yaojie, Liu Guangming, Wang Xiaofan, Lu Xiaofeng, Hei Xinhong 2257
- Colorectal lesion classification of endoscopic images based on priori weight shared codebook application
Zhu Tingwei, Li Sheng, He Xiongxiang 2270
- The fusing of dilated convolution and attention for segmentation of gastric cancer tissue sections
Chen Yingsi, Li Han, Zhou Xueting, Wan Cheng 2281
- Optic disc detection based on visual saliency in fundus image
Lyu Pengfei, Wang Ying, Wang Siqi, Yu Xiaosheng, Wu Chengdong 2293

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2021)09-2146-10

论文引用格式: Guo N and Bai Z Y. 2021. The integration of attention mechanism and dense atrous convolution for lung image segmentation. Journal of Image and Graphics, 26(09): 2146-2155 (郭宁, 柏正尧. 2021. 注意力机制下密集空洞卷积的肺部图像分割. 中国图象图形学报, 26(09): 2146-2155) [DOI:10.11834/jig.200429]

注意力机制下密集空洞卷积的肺部图像分割

郭宁, 柏正尧*

云南大学信息学院, 昆明 650500

摘要: **目的** 卷积神经网络 (convolutional neural network, CNN) 在计算机辅助诊断 (computer-aided diagnosis, CAD) 肺部疾病方面具有广泛的应用, 其主要工作在于肺部实质的分割、肺结节检测以及病变分析, 而肺实质的精确分割是肺结节检出和肺部疾病诊断的关键。因此, 为了更好地适应计算机辅助诊断系统要求, 提出一种融合注意力机制和密集空洞卷积的具有编码—解码模式的卷积神经网络, 进行肺部分割。**方法** 将注意力机制引入网络的解码部分, 通过增大关键信息权重以突出目标区域抑制背景像素干扰。为了获取更广更深的语义信息, 将密集空洞卷积模块部署在网络中间, 该模块集合了 Inception、残差结构以及多尺度空洞卷积的优点, 在不引起梯度爆炸和梯度消失的情况下, 获得了更深层次的特征信息。针对分割网络常见的特征丢失等问题, 对网络中的上/下采样模块进行改进, 利用多个不同尺度的卷积核级联加宽网络, 有效避免了特征丢失。**结果** 在 LUNA (lung nodule analysis) 数据集上与现有 5 种主流分割网络进行比较实验和消融实验, 结果表明, 本文模型得到的预测图更接近于标签图像。Dice 相似系数、交并比 (intersection over union, IoU)、准确度 (accuracy, ACC) 以及敏感度 (sensitivity, SE) 等评价指标均优于对比方法, 相比于性能第 2 的模型, 分别提高了 0.443%、0.272%、0.512% 以及 0.374%。**结论** 本文提出了一种融合注意力机制与密集空洞卷积的肺部分割网络, 相对于其他分割网络取得了更好的分割效果。

关键词: 肺分割; 卷积神经网络 (CNN); 计算机辅助诊断 (CAD); 注意力机制; 密集空洞卷积 (DAC)

The integration of attention mechanism and dense atrous convolution for lung image segmentation

Guo Ning, Bai Zhengyao*

School of Information Science and Engineering, Yunnan University, Kunming 650500, China

Abstract: Objective As an important criterion for the diagnosis of early-stage lung cancer, chest computed tomography (CT) images-based pulmonary nodules detection have been implemented via location observation, scope and shape of the lesions. The CT image has been analyzed lung organizational structures like the lung parenchyma and the contextual part, such as hydrops, trachea, bronchus, and ribs. CT images-based lung parenchyma has been hard to interpret automatically and precisely. The precise extraction of lung parenchyma has played a vital role in lung-based diseases analyses. Most of lung segmentation have been conducted based on regular image processing algorithms like threshold or morphological operation. The convolutional neural networks (CNNs) have been used in computerized pulmonary disease analysis. CNN-driven lung segmentation algorithms have been adopted in computer-aided diagnosis (CAD). The U-shape structure has been

收稿日期: 2020-08-05; 修回日期: 2020-09-22; 预印本日期: 2020-09-29

* 通信作者: 柏正尧 baizhy@ynu.edu.cn

designed for medical image segmentation based on end-to-end fully convolutional network (FCN) structure. The credibility for biomedical image segmentations have been realized based on the encoding and decoding symmetric network structure. A novel convolutional neural network based on U-Net architecture has been illustrated via integrating attention mechanism and dense atrous convolution (DAC) . **Method** The network has contained an encoder and a decoder. The encoder has consisted of convolution and down sampling. The deductible spatial dimension of feature maps have been used to learn more semantic information. And the attention mechanism decoder has been implemented for de-convolution and up-sampling to re-configure the spatial dimension of the feature maps. The decoding mode using attention mechanism has been manipulated to make the target area output more effectively. Meanwhile, the algorithm of lung image segmentation has been used to identify the target-oriented neural network's attention using transmitted skip-connection to improve the weight of the salient feature. The feature resolution capability has been enhanced to the requirements for intensive spatial prediction via pooling consecutive operations and convolution striding. The DAC block has been deployed between the encoder and the decoder to extract multi-scale information of the context sufficiently. The advantages of Inception, ResNet and atrous convolution for the block have been inherited to capture multi-sized features consequently. The max-pooling and up-sampling operators have been utilized to reduce and increase the resolution of feature maps intensively based on the classic U-Net framework, which could lead to feature loss and accuracy reduced problems during training. The original max-pooling and up-sampling operators have been replaced via down-sample and up-sample block with inception structure to widen the multi-filters network and avoid feature loss. The Dice coefficient loss function has been used instead of the cross entropy loss to identify the gap between prediction and ground-truth. The deep learning framework Pytorch have been used on a server with two NVIDIA GeForce RTX 2080Ti graphics cards and each GPU has 11 Gigabyte memory. At the experimental stage, the original images have been resized to 256×256 pixels and 80% of these for training besides the test remaining. The proposed model has been trained for 120 epochs. Based on an initial learning rate of 0.000 1, the Adam has been opted as the optimization algorithm. **Result** In order to verify the efficiency of the proposed method, we conduct multi-compatible verifications called FCN-8 s, U-Net, UNet ++, ResU-Net and CE-Net (context encoder network) have been conducted. Four segmentation metrics have been adopted to assess the segmentation. These metrics has evolved the Dice similarity coefficient (DSC), the intersection over union (IoU), sensitivity (SE) and accuracy (ACC) . The experimental results on the LUNA16 dataset have demonstrated the priorities in terms of all metrics results. The average Dice similarity coefficient has reached 0.985 9, which has 0.443% higher than the segmentation results of the second-performing CE-Net. The model consequence has achieved 0.972 2, 0.993 8, and 0.982 2 each in terms of IoU, ACC and SE. This second qualified segmentation performance has reached; 0.272%, 0.512% and 0.374% each (more better). Compared with other algorithms, the predictable results of modeling has closer to the label made. The adhesive difficulties on the left and right lung cohesion issue have been resolved well. **Conclusion** An encoded/decoded structure in novel convolutional neural network has been integrated via attention mechanism and dense atrous convolution for lungs segmentation. The experiment results have illustrated that the qualified and effective framework for segmenting the lung parenchyma area have its own priority.

Key words: lung segmentation; convolutional neural networks (CNN); computer-aided diagnosis (CAD); attention mechanism; dense atrous convolution (DAC)

0 引 言

肺癌是威胁人类生命的常见恶性肿瘤之一,每年有数百万人被确诊或死于肺癌。据统计,每年死于肺癌的患者约占全球癌症死亡人数的五分之一,超过乳腺癌、结肠癌和前列腺癌死亡人数的总和。根据中国国家癌症中心提供的数据,到 2025 年,中国的肺癌患者将会达到 100 万人,成为世界上最大

的“肺癌大国”。早期肺癌一般没有明显临床症状,但可以通过有经验的医生观察胸部计算机断层扫描 (computer tomography, CT) 中肺结节的位置、范围和形状来进行诊断和治疗。在肺分割过程中,去除会干扰病变分析的区域,例如气管、胸骨、心包和脂肪,仅保留肺实质和病变区域,可以更好地帮助找到肺结节。因此,肺实质的精确提取在肺相关疾病的分析中起着至关重要的作用。

肺分割领域的大多数方法是基于传统的图像处

理算法。其中,阈值法是实现图像分割的最常用方法之一。然而,肺中有许多部位的灰度值与体外相似,仅通过阈值法不可能获得更好的效果,因此出现了许多结合阈值法和其他方法的改进方法。袁克虹和向兰茜(2011)提出了一种自动分割肺实质的方法,包括阈值法和边界追踪法,同时还利用基于计算局部2维凸包的算法对边缘进行修正。Shojaii 等人(2005)使用分水岭变换找到肺部边界,然后使用滚球滤波器平滑轮廓,在处理形态不规则的肺部图像方面表现良好。

与传统的肺分割方法不同,基于深度卷积网络的算法也已达到了最优性能,但不会依赖于手工设计的特征提取,还可以广泛应用于各种 CAD (computer-aided diagnosis) 系统。随着端到端的全卷积网络 (fully convolution networks, FCN) (Long 等, 2015) 结构的出现, Ronneberger 等人(2015)提出了用于生物医学图像分割的 U 型网络 (U-Net) 架构。该网络是具有编码路径和解码路径的对称结构。在两个路径之间,有一些跳跃层可以连接相应的特征,从而可以完成像素级的分割。Skourt 等人(2018)将此网络应用于肺分割并获得较好的结果。Oktay 等人(2018)在 U-Net 网络中添加注意力机制,增强了对特定任务有用的前景像素的敏感度,得到了令人满意的结果。Zhang 等人(2020)将 Inception 模块和密集连接加入到 U-Net 网络。尽管分割结果表现好,但是由于网络结构太过于复杂,导致计算量负担大,训练缓慢。

受以上研究工作的启发,本文提出了一种基于深度卷积网络的肺分割网络,该网络是具有编码—解码模式的算法,结构与 U-Net 网络相似。其工作主要有以下几个方面的贡献:

1) 通过编码路径中引入注意力机制增加关键区域的权重,以突出显著特征,抑制输入图像中不相关的区域,从而提高分割精度。

2) 基于 U-Net 深度网络架构,在网络中间添加了一个密集空洞卷积 (dense atrous convolution, DAC) 模块,该模块集合了 Inception 模块和残差网络与空洞卷积的优点,在避免梯度爆炸和消失的同时获得了多尺度的特征信息。

3) 为了解决采样过程中的特征丢失问题,重新设计改进了上/下采样模块,利用 Inception 结构加宽网络。

1 相关工作

1.1 空洞卷积

在卷积神经网络中,随着卷积核的尺寸增大,相应的感受野会变得更大。此外,学习参数数目也将增加,造成训练过程中的数据过拟合。为了解决这些问题,Yu 和 Koltun(2016)提出了空洞卷积 (atrous convolution),可以在不损失分辨率的情况下扩大特征图的感受野。这是一种针对密集预测任务的特殊设计 (Chen 等, 2018)。

空洞卷积也称为扩张卷积。与常规卷积不同,空洞卷积是具有扩张率 (dilation rate) 的卷积。当扩张率等于 1 时,空洞卷积可以视为常规卷积;但是当扩张率大于 1 时,卷积核会对特征图以扩张率减 1 进行间隔采样。感受野大小 F 计算为

$$F = (r - 1) \times (k - 1) + k \quad (1)$$

式中, r 和 k 分别代表扩张率和卷积核大小。当卷积核大小为 3 时,不同扩张率的空间卷积对比如图 1 所示。

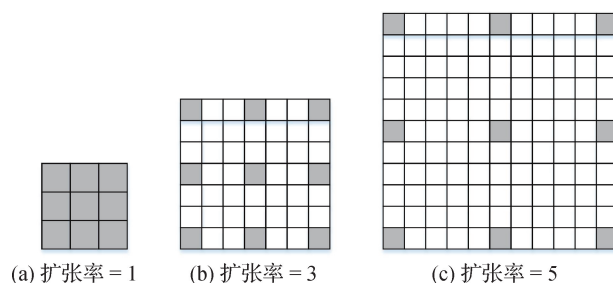


图 1 具有不同扩张率的空间卷积

Fig. 1 Atrous convolution with different dilation rates
((a) dilation rate = 1; (b) dilation rate = 3; (c) dilation rate = 5)

在卷积核大小不变的情况下,相比于常规卷积,空洞卷积不需要学习更多的参数且不会造成信息损失,就能获得更大的感受野,这样有利于增强图像分割像素点分类的准确性。

1.2 Inception

Szegedy 等人(2015)将 Inception 模块应用于 GoogLeNet 网络,并在分类和检测竞争中获得了最佳分数。此后,从 Inception-v2、v3 (Ioffe 和 Szegedy, 2015; Szegedy 等, 2016) 到 v4 (Szegedy 等, 2017), Inception 网络不断得到改进和创新,以获得更好的性能。

Inception 网络的目的是解决卷积层堆叠、避免冗余计算问题,同时还能使网络更深、更广。其中不同尺度的卷积核在增强网络的泛化能力和结构表达能力的同时,也可以为网络模型增加更多的非线性,极大地提高了卷积神经网络学习特征的能力。通常,一个 Inception 网络是由多个 Inception 模块串联组成的。

如图2所示,单独的 Inception 模块包含3个不同大小的卷积核和一个最大池化层,并在每一层使用一个 1×1 卷积进行降维,提高计算效率。将这4个分支得到的特征图拼接起来,传送至下一个 Inception 模块,可使该网络获得不同的感受野以增加网络宽度。

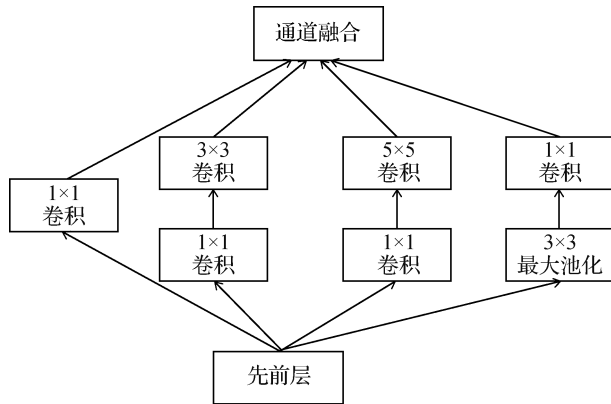


图2 单一 Inception 结构

Fig. 2 Single Inception structure

1.3 残差网络

对于卷积神经网络,网络的层次越深,训练难度越大。因为这不仅会导致网络的退化,而且还容易造成梯度消失和梯度爆炸。针对此问题,He 等人(2016)提出了残差网络,通过在卷积神经网络中引入快捷连接(shortcut connection)加强特征的传递。残差块是由常规卷积每两层增加一个捷径构成的,而数个残差块连在一起便组成残差网络。如图3所示, x 是网络的输入, $F(x)$ 表示经过两个卷积层操作的输出,在送入下一层之前将通过快捷连接的映射与原始输出进行叠加 $F(x) + x$ 。以这样的方式减轻深层网络训练的难度,从而兼顾了网络性能的提升。

2 本文方法

本文网络架构如图4所示,其中包含一条编码

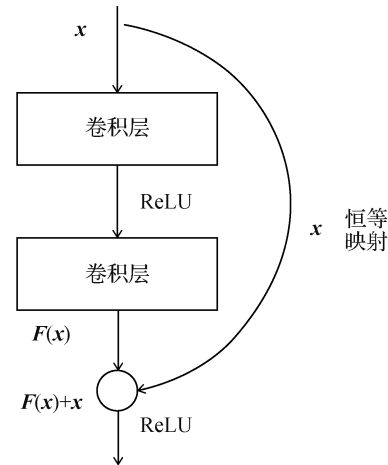


图3 残差块结构

Fig. 3 The structure of the residual block

路径(左侧)和一条解码路径(右侧)。编码路径由8个卷积层和4个下采样层组成,旨在通过逐渐减小特征图的空间尺寸来学习更多的语义信息。相比之下,具有注意力机制的解码路径是通过使用反卷积和上采样操作来恢复特征图的空间尺寸。同时,在两条路径之间添加密集空洞卷积模块,充分提取上下文的有效信息。

2.1 注意力机制模块

深度学习领域中的注意力机制(attention mechanism)(Luong 等,2015)本质上类似于人类的视觉选择性注意力机制,往往关注于已收集的所有信息的某些特定部分。考虑到该机制在辨别和聚焦方面的优势,注意力机制已广泛应用于各种人工智能领域。

在编码—解码模式的标准 U-Net 网络中,尽管在浅层网络有更多的空间信息,但是能学习到的特征并没有深层的丰富。此外,在利用级联操作进行特征融合时相似的浅层特征不应该被重复提取,否则将会导致计算负担和模型参数冗余。因此,本文将注意力机制应用于解码路径,获取更高层次的特征信息,增加目标区域的权重避免背景像素的干扰,从而提高模型的学习能力。注意力机制模块内部结构如图5所示。

该模块有两个输入,一个是上采样特征 $g(F_g \times H_g \times W_g \times D_g)$,另一个是通过跳跃连接传输的具有相同分辨率的编码特征 $x^l(F_l \times H_x \times W_x \times D_x)$,其中上采样特征 g 可以视为增强 x^l 学习能力的门控信号(gating signal)。两个输入首先经过 $1 \times 1 \times 1$ 的卷

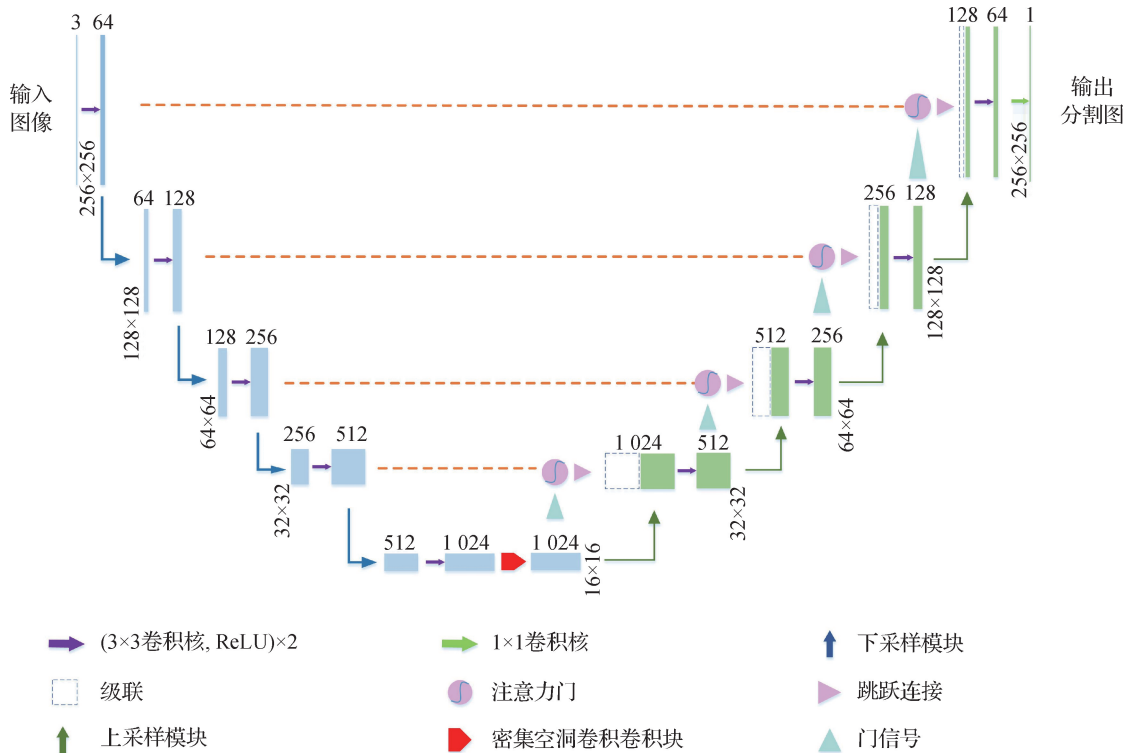


图4 本文网络结构

Fig. 4 Proposed network structure

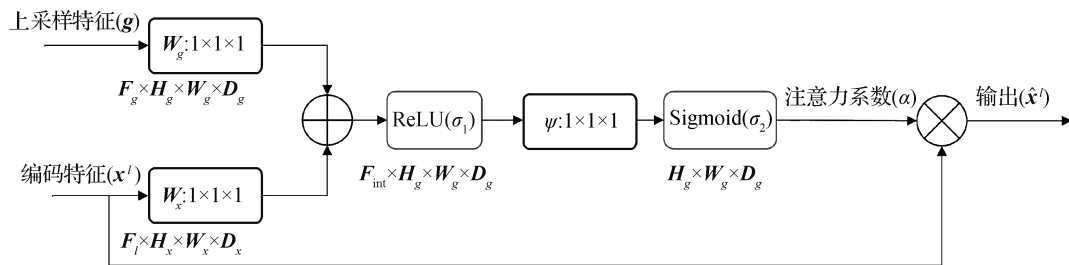


图5 注意力机制内部结构

Fig. 5 Internal structure of the attention mechanism module

积运算后,分别获得 $W_g^T g_i$ 和 $W_x^T x_i^l$ 。随后,对两个结果相加再进行 $\text{ReLU}(\text{rectified linear unit}) (\sigma_1(x_i^l) = \max(0, x_i^l))$ 激活。融合后的特征将再次进行卷积运算,并经过 s 形激活函数 $\text{Sigmoid}(\sigma_2)$, 得到注意力系数 α , 即

$$q_{\text{att}}^l = \psi^T(\sigma_1(W_x^T x_i^l + W_g^T g_i + b_g)) + b_\psi \quad (2)$$

$$\alpha = \sigma_2(q_{\text{att}}^l(x_i^l, g_i)) \quad (3)$$

式中, $b_g \in \mathbf{R}$, $b_\psi = \frac{1}{1 + \exp(-x_i)}$ 代表偏差项, ψ 代表 $1 \times 1 \times 1$ 大小的卷积核。最后,用编码特征乘以注意力系数 α , 逐像素合并输出 x^l 。

2.2 密集空洞卷积模块

U-Net 网络及其变体具有一些共同的局限性:

一方面,通过连续的池化和卷积步长操作来降低特征分辨率,往往会影响到需要非常详细空间信息的预测任务,例如本文中的肺部图像分割;另一方面,如果通过扩大卷积核的大小来增大感受野,模型参数会相应增加,这样也会不利于模型的训练。考虑到以上这些情况,本文在网络中间添加了一个密集空洞卷积(DAC)(Gu 等,2019)模块。

尽管看起来像一个密集连接块(densely connected block),实际上该模块的结构集结了 Inception、残差网络和空洞卷积,并将这些方法的优点融合起来。如图6所示,DAC 模块有4个级联分支,并且每个分支中空洞卷积的数量逐个增加,对应的感受野分别是3,7,9,19。同时每个分支

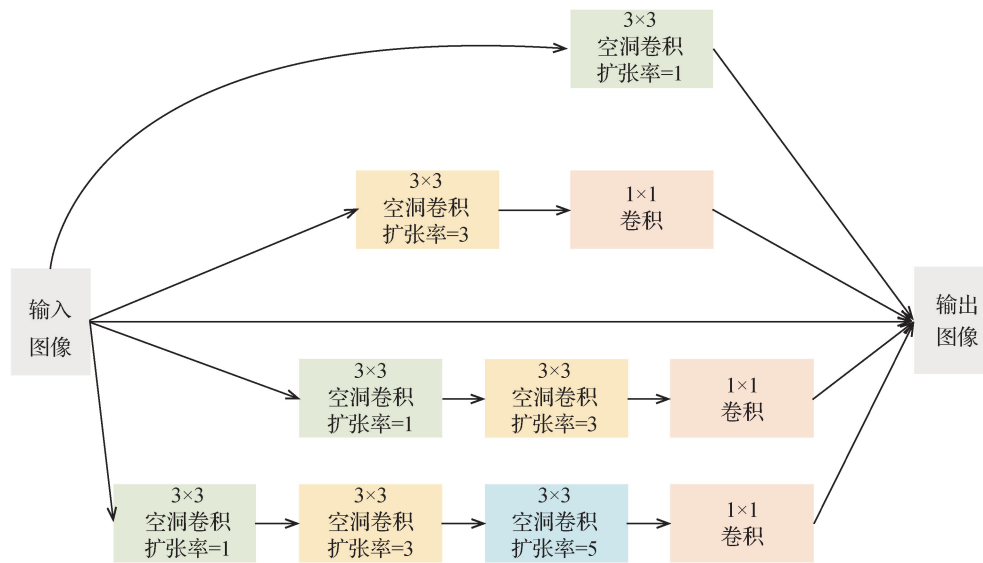


图6 密集空洞卷积

Fig.6 Dense atrous convolution block

最后都使用一个 1×1 卷积用于 ReLU (Glorot 等, 2011) 激活。此外, 该模块也引入残差网络中的快捷连接将原始特征与其他特征融合在一起, 避免了梯度的爆炸和消失。通过结合不同膨胀率的空洞卷积, DAC 块能够充分提取不同尺度的特征。

2.3 改进的上/下采样模块

在标准的 U-Net 框架中, 最大池化和上采样操作分别用于减少和增加特征图的分辨率, 但训练过

程中可能会导致特征丢失和降低精度的问题。因此, 本文对上/下采样模块进行改进, 避免特征的丢失。

如图7所示, 两个模块具有相同的结构, 只是下采样模块中的最大池化层和卷积层在上采样模块中被替换成了上采样层和转置卷积层。该结构由多个不同尺度的卷积核并行级联构成, 这可以看做是简单的 Inception 结构, 可以感知不同范围和尺度的局部特征, 从而提高网络的学习能力。

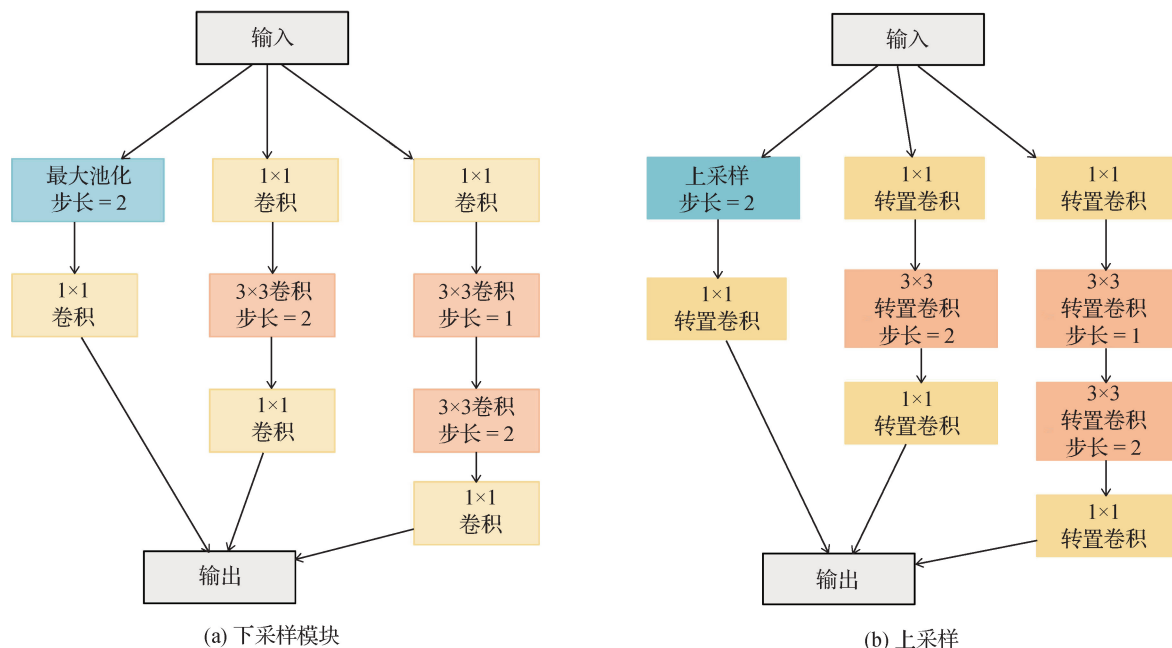


图7 下采样与上采样模块

Fig.7 Down-sample and up-sample block ((a) down-sample block; (b) up-sample block)

2.4 损失函数

损失函数可以用于评估预测结果与真实结果之间的差异。当损失函数较小时,对应模型的鲁棒性相对较强。由于肺实质分割问题属于像素分割,因此,将常用的二分类交叉熵函数(binary cross entropy, BCE)损失函数替换为 Dice 损失函数(Milletari 等, 2016)。

Dice 相似性系数(C_{dice})是计算两个样本之间重叠区域的度量,其范围在 0~1 之间,1 表示完全重叠,0 代表没有重叠区域。Dice 损失函数计算为

$$\text{Loss}_{\text{dice}} = 1 - C_{\text{dice}} = 1 - \frac{2 \sum_i p_i g_i}{\sum_i p_i^2 + \sum_i g_i^2} \quad (4)$$

式中, N 表示像素个数, $g_i \in \{0, 1\}$ 表示像素 i 是否为前景的标签,而 $p_i \in (0, 1)$ 则表示 softmax 层输出像素 i 的预测结果。

3 实验结果与分析

3.1 数据集

本文所使用的数据集来自 2016 年的肺结节分析(lung nodule analysis, LUNA)竞赛。该竞赛的目的是检测肺部 CT 图像中的结节并降低假阳性。为了方便参赛者更好地找到肺部区域,竞赛同时提供了分割数据集(<https://www.kaggle.com/kmader/finding-lungs-in-ct-data/data>),该数据集包含了手动分割肺部的 2 维和 3 维图像。本文只使用 2 维的 CT 图像,其中包含 267 幅图像及其相应的标签图像。同时,将原始图像的尺寸调整为 256×256 像素,其中 80% 的样本用于训练和验证,其余的用于测试。

3.2 参数设置

实验使用两块 NVIDIA GeForce RTX 2080Ti GPU 显卡,其中每个 GPU 均有 11 GB 的内存。在实验阶段,深度学习框架采用 PyTorch,模型训练 120 个批次。此外,学习率设置为 0.000 1 并选择 Adam(Kingma 和 Ba, 2015)作为优化算法。同时训练和测试的批次大小设置为 8,这意味着在每个批次中将会有 8 个样本作为输入被送入到网络中。

3.3 评价指标

采用 4 个常用分割评价指标评估该方法的分割结果。这些指标包括 Dice 相似性系数(DSC)、交并

比(IoU)、敏感度(SE)和准确性(ACC)。计算公式分别为

$$DSC = \frac{2 |GT \cap SR|}{|GT| + |SR|} \quad (5)$$

$$IoU = \frac{|GT \cap SR|}{|GT \cup SR|} \quad (6)$$

$$SE = \frac{TP}{TP + FN} \quad (7)$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

式中, GT 代表由专家标注的标签, SR 代表本文模型得到的预测结果。 TP 是被正确分类为肺实质的肺部像素数量; TN 是被正确分类为背景的背景像素数量; FP 是被错误分类为肺实质的背景像素数量; FN 是被错误分类为背景的肺部像素数量。

对于医学图像分割, DSC 和 IoU 被广泛认为是其重要评价指标。 ACC 和 SE 可以帮助确保评估模型的鲁棒性。

3.4 实验结果分析

3.4.1 消融实验

为充分探究、分析本文提出的多个模块对肺分割任务的有效性,对注意力机制、DAC 和改进的上/下采样模块进行了消融实验,对比改进前后的模型分割性能。消融实验结果如表 1 所示,从实验的比较结果中可以看出,加入注意力机制与 DAC 模块对模型性能提升最大。本文将 DAC 模块应用到该模型,能够进一步提取全局信息,保证了分割的完整性。而采用注意力机制的解码模式可以达到增强目标特征并且抑制背景的目的,有助于提高分割质量。尽管改进后的上/下采样模块对模型的贡献比不上注意力机制和 DAC,但也在一定程度上促进了模型的分割效果。

表 1 消融实验结果

Table 1 Results of ablation Study

方法	ACC	SE	DSC	IoU
U-Net + DAC + D/U	0.983 9	0.972 1	0.974 7	0.956 6
U-Net + Attn + D/U	0.987 2	0.972 9	0.976 6	0.960 2
U-Net + Attn + DAC	0.993 4	0.980 4	0.983 3	0.967 3
U-Net + Attn + DAC + D/U	0.993 8	0.982 2	0.985 9	0.972 2

注:Attn 为注意力机制,DAC 为密集空洞卷积模块,D/U 为改进后的上/下采样模块,加粗字体为每列最优结果。

3.4.2 与其他分割算法的对比实验

为了验证本文网络的有效性,将所提出的模型与其他最新分割网络进行了一些对比研究。为使实验更有说服力,FCN-8 s, U-Net, UNet++ (Zhou 等, 2018), ResU-Net 和 CE-Net (context encoder network) (Gu 等, 2019) 的实验采用与本文模型相同的优化算法、损失函数和其他初始实验参数。比较结果如表 2 所示。

表 2 不同分割算法的比较

Table 2 Comparison of different segmentation algorithms

方法	ACC	SE	DSC	IoU
FCN-8 s	0.985 3	0.975 9	0.973 6	0.954 5
U-Net	0.986 1	0.974 1	0.975 6	0.958 1
UNet++	0.987 1	0.970 2	0.976 4	0.959 7
ResU-Net	0.987 6	0.972 5	0.976 6	0.963 2
CE-Net	0.988 7	0.978 5	0.981 4	0.969 5
本文	0.993 8	0.982 2	0.985 9	0.972 2

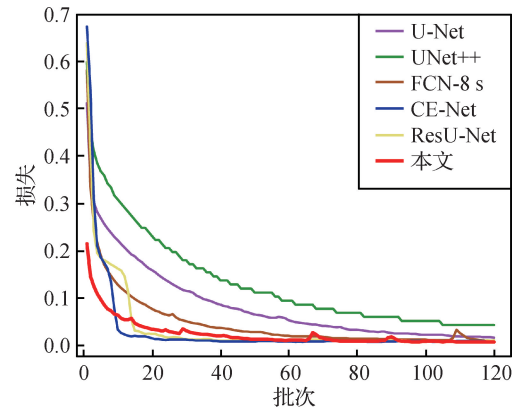
注:加粗字体为每列最优结果。

实验中,本文算法的 Dice 相似性系数(DSC)为 0.985 9,比表 2 中性能第 2 的网络—CE-Net 提高了 0.443%;而在 IoU、ACC 和 SE 方面,本文算法的结果均高于 CE-Net 网络,分别提高了 0.272%, 0.512% 和 0.374%。由此可以得出结论,本文网络的分割有效性优于其他分割网络。

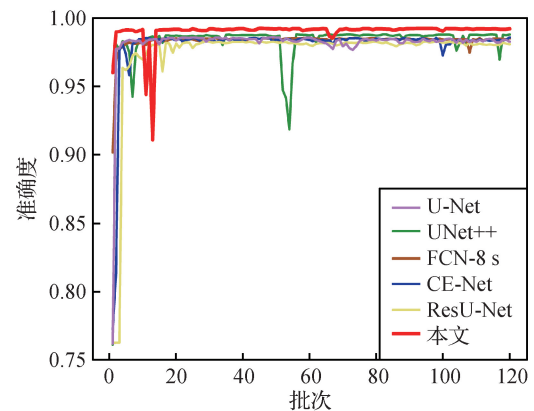
各模型的训练损失曲线及其对应的验证精度曲线如图 8 所示。在确保实验一致性的情况下,即使收敛速度不是最快,但该网络的鲁棒性也相对较强。除此之外,为了直观地说明不同算法的准确性,图 9 给出了 4 种分割结果。

图 9 第 1 行,FCN-8 s 网络的分割效果是最差的。而对于 U-Net, UNet++ 和 ResU-Net 网络,尽管网络中具有跳跃连接,但仍有一些像素被错误地分类。与表现较好的 CE-Net 网络相比,本文模型的预测结果更接近专家的标签。这里可以说明注意力机制是一个有助于模型增强目标区域学习能力的促成因素,从而提高了分割效率。

图 9 第 2 行中,除了本文算法和 CE-Net 网络,其他网络对于支气管和肺实质的许多像素未能进行有效分类。这意味着通过使用 DAC 模块的网络,可以捕获更多高级的特征,学习到更多丰富的信息。



(a) 训练损失



(b) 验证精度

图 8 不同模型的训练损失与验证精度曲线

Fig. 8 Training loss and validation accuracy of different models
((a) training loss; (b) validation accuracy)

肺部分割的棘手问题是当左右肺之间的距离太近时,分割会出现粘连现象,如图 9 中第 3 行所示。而本文的网络很好地解决了这一问题,表明本文方法在捕获更广泛和更深的语义特征方面比其他方法具有更好的性能。

而令人失望的是,本文模型与实验中比较的所有模型一样,在模糊边界上的分割效果不佳,结果都不尽人意,如图 9 中第 4 行所示。

从分割结果来看,本文方法具有有效分割肺实质区域的稳定网络架构,其性能优于其他分割网络且鲁棒性较强。

4 结 论

本文提出一种融合注意力机制与密集空洞卷积的肺部分割网络。该网络具有编码—解码模式,通过在网络的解码路径部分引入注意力机制使得目标区域能够得到有效的输出,从而提高了分割精度,同

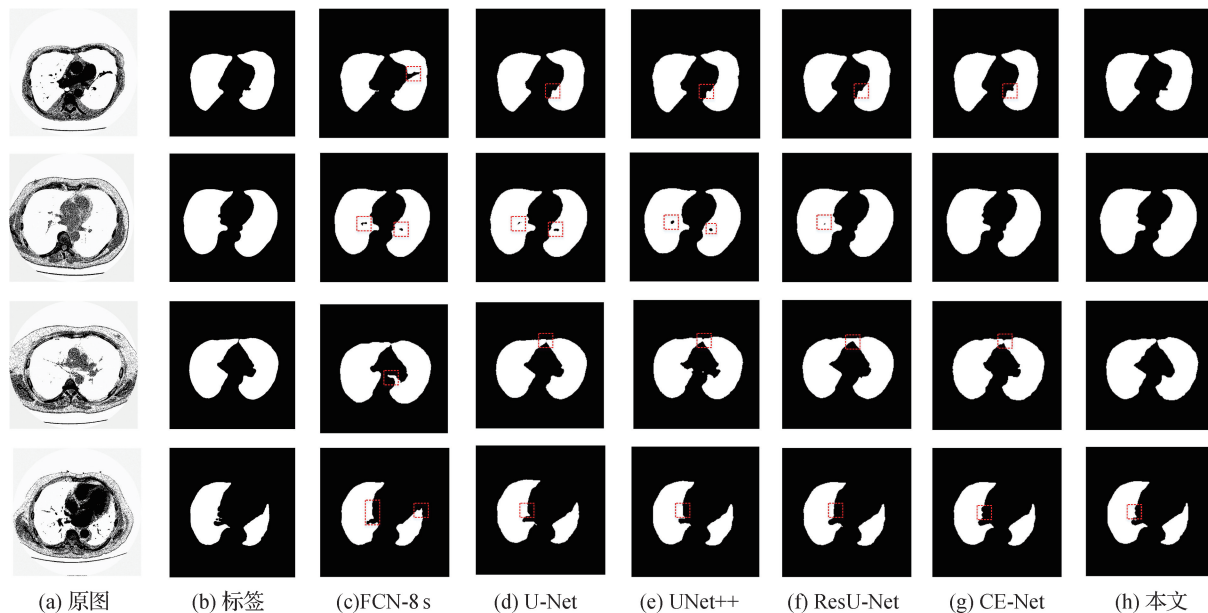


图9 不同模型的分割结果

Fig. 9 Segmentation results of different methods

((a) origin; (b) label; (c) FCN-8 s; (d) U-Net; (e) UNet++ ; (f) ResU-Net; (g) CE-Net; (h) ours)

时在网络中间添加密集空洞卷积模块使得网络可以充分提取上下文的多尺度信息。此外,改进后的上/下采样模块也可以在避免特征丢失的情况下有效地加深网络的深度,使该网络能够学习到更多的图像特征。一系列的对比实验结果表明,本文提出的模型比其他常见的分割网络更适合进行肺部分割的应用。在未来的研究中,该网络可以进一步得到改进和优化,使其同样可以应用于3维CT图像。除此之外,该模型仅在肺分割任务上得到验证,尚未应用于其他2维医学图像分割任务,后续工作可能会将其扩展到不同的医学图像分割任务,使该网络更具通用性。

参考文献 (References)

- Chen L C, Papandreou G, Kokkinos I, Murphy K and Yuille A L. 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): 834-848 [DOI: 10.1109/TPAMI.2017.2699184]
- Glorot X, Bordes A and Bengio Y. 2011. Deep sparse rectifier neural networks [EB/OL]. [2020-7-29]. <http://proceedings.mlr.press/v15/glorot11a/glorot11a.pdf>
- Gu Z W, Cheng J, Fu H Z, Zhou K, Hao H Y, Zhao Y T, Zhang T Y, Gao S H and Liu J. 2019. CE-Net: context encoder network for 2D medical image segmentation. *IEEE Transactions on Medical Imaging*, 38(10): 2281-2292 [DOI: 10.1109/TMI.2019.2903562]
- 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]
- Ioffe S and Szegedy C. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift [EB/OL]. [2020-07-29]. <https://arxiv.org/pdf/1502.03167.pdf>
- Kingma D P and Ba J L. 2015. Adam: a method for stochastic optimization [EB/OL]. [2020-07-29]. <https://arxiv.org/pdf/1412.6980.pdf>
- Long J, Shelhamer E and Darrell T. 2015. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4): 640-651 [DOI: 10.1109/CVPR.2015.7298965]
- Luong M, Pham H and Manning C D. 2015. Effective approaches to attention-based neural machine translation//*Proceedings of 2015 Conference on Empirical Methods in Natural Language*. Lisbon, Portugal: ACL: 1412-1421 [DOI: 10.18653/v1/d15-1166]
- Milletari F, Navab N and Ahmadi S A. 2016. V-Net: fully convolutional neural networks for volumetric medical image segmentation//*Proceedings of the 4th International Conference on 3D Vision*. Stanford, USA: IEEE: 565-571 [DOI: 10.1109/3DV.2016.79]
- Oktay O, Schlempe J, Le Folgoc L, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla N Y, Kainz B, Glocker B and Rueckert D. 2018. Attention U-Net: learning where to look for the pancreas [EB/OL]. [2020-07-29]. <https://arxiv.org/pdf/1804>

03999. pdf

- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//Navab N, Hornegger J, Wells W and Frangi A, eds. Medical Image Computing and Computer-Assisted Intervention. Lecture Notes in Computer Science, Vol. 9351. Munich, Germany: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Shojaii R, Alirezaie J and Babyn P. 2005. Automatic lung segmentation in CT images using watershed transform//Proceedings of 2005 IEEE International Conference on Image Processing. Genova, Italy: IEEE: 1270-1273 [DOI: 10.1109/ICIP.2005.1530294]
- Skourt B A, El Hassani A and Majda A. 2018. Lung CT image segmentation using deep neural networks. Procedia Computer Science, 127: 109-113 [DOI: 10.1016/j.procs.2018.01.104]
- Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A. 2015. Going deeper with convolutions//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE: 1-9 [DOI: 10.1109/CVPR.2015.7298594]
- 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]
- Szegedy C, Ioffe S, Vanhoucke V and Alemi A. 2017. Inception-v4, inception-ResNet and the impact of residual connections on learning [EB/OL]. [2020-7-29]. <https://arxiv.org/pdf/1602.07261.pdf>
- Yu F and Koltun V. 2016. Multi-scale context aggregation by dilated convolutions [EB/OL]. [2020-7-29]. <https://arxiv.org/pdf/1511.07122.pdf>
- Yuan K H and Xiang L X. 2011. Automated lung segmentation for chest CT images used for computer aided diagnostics. Journal of Tsinghua University (Science and Technology), 51(1): 90-95 (袁克虹, 向兰茜. 2011. 用于计算机辅助诊断的肺实质自动分割方法. 清华大学学报(自然科学版), 51(1): 90-95) [DOI: 10.16511/j.cnki.qhdx.2011.01.018]

Zhang Z, Wu C D, Coleman S and Kerr D. 2020. DENSE-INception U-net for medical image segmentation. Computer Methods and Programs in Biomedicine, 192: #105395 [DOI: 10.1016/j.cmpb.2020.105395]

Zhou Z W, Rahman Siddiquee M, Tajbakhsh N and Liang J M. 2018. UNet++: a nested U-net architecture for medical image segmentation//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Lecture Notes in Computer Science, vol. 11045. Switzerland: Springer: 3-11 [DOI: 10.1007/978-3-030-00889-5_1]

作者简介



郭宁, 1995年生, 女, 硕士研究生, 主要研究方向为基于深度学习的医学影像处理。

E-mail: guoning95@qq.com



柏正尧, 通信作者, 男, 教授, 主要研究方向为信号处理、图像处理、模式识别。

E-mail: baizhy@ynu.edu.cn