

可分离卷积与注意力机制结合的肝癌靶区CT图像自动分割

张福全,周永康,杨悦,陈芷涵,高宇楠,林小惟
复旦大学附属中山医院放射治疗科,上海 570228

【摘要】目的:开发一种可分离卷积与注意力机制结合的肝癌靶区CT图像分割方法。**方法:**基于U-Net卷积神经网络的肝癌靶区分割方法,为增强分割模型的特征表达能力,将注意力模块与U-Net模型相结合,提高与分割任务相关性更大的特征通道权重;在网络模型编码阶段引入本文提出的可分离卷积,补充下采样过程中损失的特征信息。**结果:**在复旦大学附属中山医院50例肝癌患者数据集上进行靶区分割,实验结果表明,相比于已有方法,本文提出方法的分割平均戴斯相似系数比3D U-Net提高4.54%。**结论:**基于可分离卷积和注意力机制的U-Net卷积神经网络对肝癌靶区达到更好的分割精度,有望在临床应用中提高医生的工作效率。

【关键词】肝癌;深度学习;卷积神经网络;靶区分割;注意力机制

【中图分类号】R318;R811.1

【文献标志码】A

【文章编号】1005-202X(2025)07-0918-05

Automatic liver cancer target area segmentation in CT image using separable convolution and attention mechanism

ZHANG Fuquan, ZHOU Yongkang, YANG Yue, CHEN Zhihan, GAO Yunan, LIN Xiaowei

Department of Radiotherapy, Zhongshan Hospital, Fudan University, Shanghai 570228, China

Abstract: Objective To propose an approach integrating separable convolution and attention mechanism for automatically segmenting target area in radiotherapy for liver cancer from CT images. **Methods** The novel liver cancer segmentation method based on the U-Net convolutional neural network integrated attention block and U-Net model to increase the weights of feature channels with greater relevance to segmentation tasks, thereby enhancing the feature expression ability, and a separable convolution was used in the U-Net encoding stage for supplementing feature information lost during the downsampling process. **Results** Target area segmentation was conducted on a dataset of 50 liver cancer patients at Zhongshan Hospital, Fudan University. Experimental results showed that the proposed method improved the average Dice similarity coefficient by 4.54% as compared with 3D U-Net. **Conclusion** The U-Net convolutional neural network based on separable convolution and attention mechanism achieves better segmentation accuracy for liver cancer target area, which is expected to improve the work efficiency of doctors in clinic.

Keywords: liver cancer; deep learning; convolutional neural network; target area segmentation; attention mechanism

前言

肝癌是中国高发的恶性肿瘤之一^[1]。由于肝癌对放射线敏感性较高,放射治疗成为肝癌的主要治疗手段^[2-3]。通过放射性射线照射肿瘤杀死癌细胞,多数病人经过放疗后能达到根治疾病的目的^[4]。目前,放疗医师通常使用手工勾画靶区,这种方法不仅

繁琐、劳动强度大、耗时长,并且勾画的结果受主观因素影响较大,具有观察者差异性。近年来,基于图谱的传统分割方法将标注好的模板图像与待分割图像进行配准,将模板标签转换到待分割图像得到分割结果^[5]。这种方法对于形状差别不大的器官效果较好,然而肝癌发展侵犯周围器官和组织,患者靶区与模板图像具有差异性,配准精度降低导致分割精确度较差。近年来,深度学习已经成为机器学习中突出的研究领域。卷积神经网络通过卷积操作提取图像特征,能有效提取图像中的信息,在医学图像分割领域得到广泛应用^[6-7]。Gao等^[8]将深度学习算法应用到肝癌靶区CT图像中,进行恶性肿瘤鉴别。Lyu等^[9]通过弱监督纳入深度卷积神经网络和门控循环神经网络算法,基于多期对比增强CT和临床特征区

【收稿日期】2025-02-23

【基金项目】上海市“科技创新行动计划”扬帆计划(23YF1438600)

【作者简介】张福全,研究方向:人工智能、放射治疗、深度学习,E-mail:
Zhang.fuquan@zs-hospital.sh.cn

【通信作者】林小惟,研究方向:人工智能、放射治疗、深度学习,E-mail:
lin.xiaowei@zs-hospital.sh.cn

分恶性肿瘤。然而,CT图像平面内和平面间分辨率存在差异,平面内分辨率较高,平面间分辨率较低,切片间距过大导致各向异性间距问题。之前的研究基于2D或3D卷积神经网络,忽视了CT图像存在的各向异性间距问题。其次,肝靶区的体积较小,只出现在少量CT图像切片中,并且不同患者的靶区形状变化较大,与周围器官或组织对比度差导致靶区边界不清晰,给分割带来很大的难度。现有的研究工作采用归一化等医学图像预处理方式增加靶区与周围组织的对比度,没有考虑靶区及其边界的特征信息^[10]。

为了解决肝癌靶区CT图像分割中存在的问题,本文设计一种新颖的卷积神经网络分割模型。首先,为解决CT图像中存在的各向异性间距问题,本文提出一种可分离卷积,将3D卷积拆分为 $3\times 3\times 1$ 平面内卷积和 $1\times 1\times 3$ 平面间卷积,输入图像首先经过 $3\times 3\times 1$ 平面内卷积提取图像中的空间信息,然后采用 $1\times 1\times 3$ 平面间卷积处理CT图像平面间分辨率较差、切片间距过大的问题,有效提取切片间的上下文信息。其次,为解决肝癌靶区在CT图像中占比过小、形态多样和边界不清晰问题,本文提出一种空间注意力模块,该模块使得模型能够充分利用CT图像的像素级信息,更好地关注肝癌靶区和边界的特征。最后,利用残差机制将可分离卷积和空间注意力模块进行结合,形成残差注意力分离卷积,替换3D U-Net中的3D卷积。通过残差机制能够最大限度地减少语义间隙,探究更深层次的网络模型,实现特征增强。实验结果表明,本文提出的肝癌靶区CT图像分割模型能有效地提高肝癌靶区分割性能。

1 方法

1.1 可分离卷积

由于CT图像的成像机制,CT图像在平面内分辨率和平面间分辨率存在较大差异,平面内分辨率较高,平面间间距较大导致分辨率差,即CT图像的各向异性间距问题。本研究中肝癌患者靶区CT图像平面分辨率约为1 mm,CT图像切片的间距较大,平面间分辨率约为3 mm,导致CT图像中横断面显示清晰,而矢状面、冠状面显示较差。受深度可分离卷积对完整卷积操作进行拆分,以及对空间和通道图像分别处理方式的启发^[11],针对CT图像存在的平面内分辨率和平面间分辨率差异导致的各向异性间距问题,对深度可分离的卷积进行改进,提出一种可分离卷积(Seperable Convolution, SC)。SC结构如图1所示,X、Y轴代表CT平面内图像,Z轴为CT切片间图像,SC将 $3\times 3\times 3$ 的卷积核拆分为X、Y、Z轴为 $3\times 3\times 1$

的平面内卷积和 $1\times 1\times 3$ 的平面间卷积。由于本文实验采用的数据集CT图像平面内分辨率较高,约为1 mm,CT图像切片间距较大,约为3 mm,因此将平面内卷积设定为 $3\times 3\times 1$ 、平面间卷积设定为 $1\times 1\times 3$ 的方式。首先对输入图像进行平面内卷积,处理CT平面图像,再进行平面间卷积处理切片间CT图像,平面间卷积由原来的Pointwise convolution处理通道改进为处理CT图像的切片。

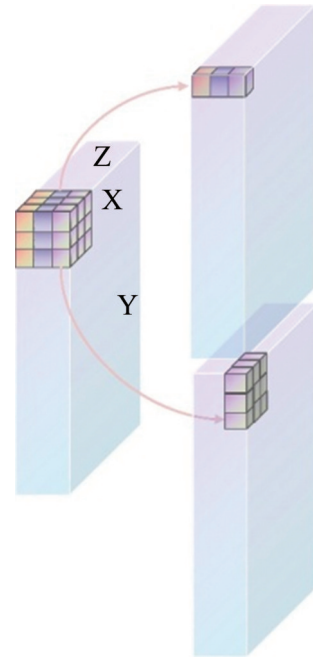


图1 可分离卷积模块

Figure 1 Seperable convolution module

1.2 残差注意力分离卷积模块

在卷积神经网络中,低层次的网络感受野较小,几何细节信息表征能力强。随着网络层次加深,模型的表达能力越强,感受野越大,语义信息表征能力越强。但不断增加网络深度会引发梯度爆炸、梯度弥散等问题。为解决该问题,He等^[12]提出残差机制,能够实现更好的特征融合,防止网络在训练过程中出现梯度爆炸、退化的情况,进一步提升网络的表征能力。残差结构如图2a所示,当输入 F 通过卷积、ReLU操作后,产生的结果为 $C(F)$,将其与原始输入 F 相加,得到最后的结果 $C(F)+F$ 。经过梯度传播后,现在传到前一层的梯度就多了个 F 的梯度,来自深层的梯度能直接去到上一层,使浅层的网络层参数得到有效训练,使得残差网络在反向传播过程中梯度信息更容易传播。SE通道注意力模块已被证明在图像分类领域中具有很好的效果,但在图像分割中,像素级别的空间信息更具有重要性^[13]。为探究空间注

注意力对肝癌靶区CT图像分割的有效性,本文对SE通道注意力模块进行改进,针对鼻咽癌靶区分割存在的问题,提出空间注意力模块(Spatial Attention Model, SAM),SAM结构如图2所示,由两个可分离卷积SC、ReLU激活函数和Sigmoid函数组成。相比于SE注意力,舍去了全局平均池化压缩通道信息,采用可分离卷积层替换全连接层,减少参数量的同时能更好地提取空间信息,并解决CT图像的各向异性间距问题。

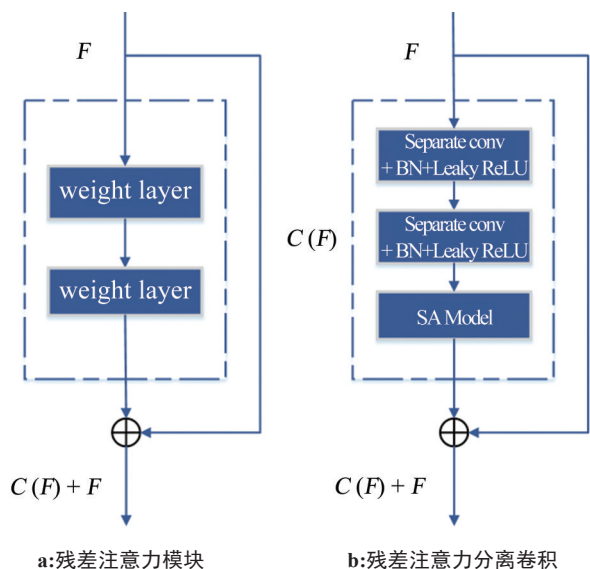


图2 残差注意力模块

Figure 2 Residual attention modules

本文基于残差机制将可分离卷积和空间注意力模块进行结合,定义为残差注意力分离卷积(Residual Attention Separable Convolution, RAS convolution),结构如图2所示,每个RAS卷积由两条

路径组成,第一条路径对输入特征图进行两次可分离卷积SC、归一化、Leaky ReLU激活函数来进行一系列特征学习,输出的特征图通过SAM重新分配特征图像素权重。第二条路径是输入特征图F,将这两条路径相加作为输出,以用于重新校准空间的上下文信息,并且可以在非常深的网络中更新参数,能够探究更深层次的卷积神经网络模型。

1.3 RAS U-Net 模型

基于医学图像分割中,3D U-Net只需要较少的训练图像就可以达到精确的分割效果,适用于本文和大多数实际医疗应用的数据集。本文提出的网络(RAS U-Net)基于3D U-Net作为主要网络模型方法,并做进一步改进提升分割性能。3D U-Net采用各向同性的3D卷积处理输入图像,忽略CT图像存在的各向异性间距问题。为解决该问题,将3D U-Net中的3D卷积替换为本文提出的残差注意力分离卷积(RAS convolution)。RAS U-Net网络结构如图3所示,由编码器、解码器结构组成,左半部分为编码器结构,右半部分为解码器结构。RAS U-Net共包含9个RAS convolution。特征图通过RAS convolution进行特征提取,之后进行最大池化层下采样,缩小特征图维度并将通道数翻倍。为避免过度拟合,提高RAS U-Net模型的泛化性,在RAS U-Net模型中加入正则化操作,即在编码器中的每个RAS convolution之后增添Dropout层,参数设置为0.5。解码器阶段特征图通过上采样恢复维度并减少通道数量。编码器中产生的特征图通过Skip connection拼接到对应层次的解码器通道维度中,以补充特征信息。残差机制能探究更深的网络模型,有助于提取更深层次的靶区特征信息,提高分割精度。传统的3D U-Net只进行3次最大池化下采样操作。RAS U-Net模型最后通过Softmax层输出分割结果的概率图。

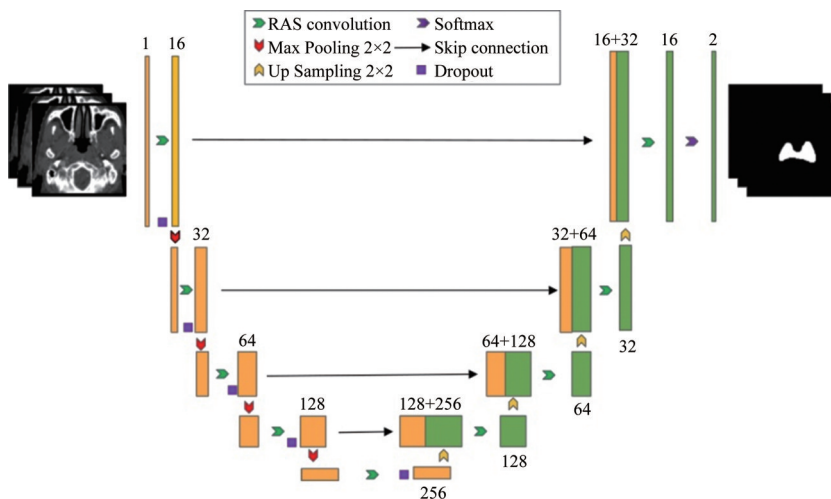


图3 RAS U-Net网络结构图

Figure 3 RAS U-Net structure

2 实验与结果分析

2.1 实验数据集

本文实验采用的数据集由复旦大学附属中山医院提供,共包含50例头部至胸部区域的肝癌患者的CT图像。CT图像平面尺寸为 272×272 像素,空间分辨率分布为 $2.9 \sim 3.0$ mm,切片数量分布在130层,平均层厚为3 mm。数据集中肝癌靶区均由放疗专家进行手工标注,作为金标准,用于本文的模型训练和模型测试过程。

2.2 数据预处理

不同医院CT设备的采集参数和类型不同,导致获取到的CT图像在像素尺寸、空间分辨率等均存在差异。如果分辨率存在差异的原始CT图像不经处理直接输入到模型中进行训练,将产生数据分布变化较大的特征图,导致模型无法较好地进行学习,降低模型收敛速度。为解决上述问题,在CT图像输入网络训练之前,将原始CT图像进行重采样,以确保输入的图像具有相同的空间分辨率和位置信息。由于最近邻插值计算量小,并且CT图像具有像素值连续性,因此本文使用最近邻插值重采样方法。本文肝癌患者靶区CT图像平面分辨率约为 $2.9 \sim 3.0$ mm,切片数130,将图像强制插值到较小的间距会产生伪影,因此将所有肝癌患者靶区CT图像重采样为平面内分辨率为1 mm,平面间分辨率为3 mm。本文对CT图像采取截断归一化方式,即对感兴趣器官的窗宽减去窗位后,选择 $[-150, 150]$ 软组织窗截断灰度范围,然后归一化到 $[-1, 1]$ 内,消除过大或过小HU组织的影响,增加肝癌靶区的对比度。本文肝癌靶区CT图像平面尺寸为 272×272 像素。因此,根据先验信息去除大部分与任务无关的背景区域,对肝癌患者的CT图像进行裁剪处理,能够缓解肝癌靶区在CT图像中占比过小的问题,大幅度减少模型的计算量,加快训练时间。

2.3 结果与分析

本文模型在12.6 GB显存NVIDIA RTX 3060 GPU进行训练、测试。模型实现采用PyTorch框架,模型训练的初始学习率设置为 10^{-4} ,批量大小设置为8,优化器为Adam^[14]算法,损失函数为Dice Loss。在50例肝癌靶区CT图像中随机选取30例患者的CT图像作为训练集,10例作为验证集,10例作为测试集。本文实验中,采取随机裁剪、随机翻转的数据增强方法,能防止数据集过少导致的过拟合问题。为了对肝癌靶区CT图像的分割结果进行客观、全面的评估,本文采用戴斯相似系数(Dice Similarity Coefficient, DSC)、平均对称表面距离(Average Symmetric Surface Distance, ASSD)和相对体积误差

(Relative Volume Error, RVE)3种评估指标对比不同分割模型的分割表现。本文对3D U-Net^[15]、V-Net^[16]和SE注意力模块结合3D U-Net的SE U-Net^[17],以及本文提出的RAS U-Net共4种分割模型进行性能测试,所有模型编码全部使用PyTorch框架实现,进行统一的CT图像预处理。实验结果如表1所示,展示不同模型的分割结果。原始的3D U-Net分割效果较差,DSC值为69.54%,ASSD为7.45 mm,RVE为56.8%,尽管3D U-Net广泛应用在医学图像分割领域,但没有经过优化的3D U-Net无法准确识别肝癌靶区的边界,导致对肝癌靶区分割结果较差。V-Net相对3D U-Net的分割结果有较少提升,DSC值提高1.16%。SE U-Net在3D U-Net基础上添加通道注意力模块,能弥补网络局部性过强导致的全局性较差的问题,从而补充全局的上下文信息,增加感受野大小,分割结果较3D U-Net模型有大幅度提升,DSC值提高2.09%。本文提出的RAS U-Net在DSC、ASSD和RVE指标中均取得最好的效果,DSC值为74.08%,ASSD为5.46 mm,RVE为34.5%,DSC值比3D U-Net提高4.54%,具有更好的分割性能。

表1 不同模型的分割结果

Table 1 Segmentation results of different models

模型	DSC/%	ASSD/mm	RVE/%
3D U-Net	69.54	7.45	56.8
SE U-Net	71.63	6.21	38.9
V-Net	70.70	6.76	42.6
RAS U-Net	74.08	5.46	34.5

3 讨论

肝癌靶区手工勾画不仅耗时,并且过于依赖放疗医师的临床能力。由于肝和周围的组织对比度较低,导致分割效果差。近年来,许多学者提出不同的解决方案,邓金城等^[18]提出带孔全卷积神经网络。门阔等^[19]采用反卷积神经网络将CT图像输入到模型中进行训练。张倩雯等^[20]单纯采用ResUnet获取图像的多尺度信息,但这些方法无法提高较小靶区的分割精度。本研究主要提出基于可分离卷积和注意力机制的肝癌靶区CT图像分割模型。首先,为解决CT图像中存在的各向异性间距问题,本文提出一种可分离卷积,将普通的3D卷积拆分为平面内卷积和平面间卷积,更好地处理CT图像中平面内和平面间分辨率不一致的情况。其次,提出一种空间注意力模块,使得模型更加注重肝癌靶区和边界的特征。基于残差机制设计残差注意力分离卷积,替换3D U-Net

模型中的普通3D卷积。将提出的模型与其他经典的分割模型进行比较,实验结果表明,本文提出的方法可以有效分割肝癌靶区CT图像,取得最精确的分割效果。综上所述,本文提出的方法能提升现有方法的肝癌靶区自动分割精度,且耗时很短,有助于提升医师临床实践中的效率。但由于本文模型架构相比于对比方法模型更为复杂,导致模型参数计算量较大,进一步使网络模型更加简洁和少参数量是接下来的研究重点。

【参考文献】

- [1] 中华医学会放射肿瘤学分会, 中国生物医学工程学会精确放疗分会肝癌学组与消化系统肿瘤专家委员会, 中国研究型医院学会放射肿瘤学分会肝癌学组. 2016年原发性肝癌放疗共识[J]. 中华放射肿瘤学杂志, 2016, 25(11): 1141-1150.
Radiation Oncology Branch of the Chinese Medical Association, Expert Committee on Liver Cancer and Digestive System of China Institute of Biomedical Engineering, Liver Cancer Research Group of Radiation Oncology Branch of China Research Hospital. Consensus on radiation therapy for primary liver cancer in 2016[J]. Chinese Journal of Radiation Oncology, 2016, 25(11): 1141-1150.
- [2] 曾昭冲. 肝细胞肝癌放射治疗现状与展望[J]. 中华肿瘤防治杂志, 2009, 16(13): 961-965.
Zeng ZC. Radiation therapy for hepatocellular carcinoma-current status and perspectives[J]. Chinese Journal of Cancer Prevention and Treatment, 2009, 16(13): 961-965.
- [3] Chen WQ, Chiang CL, Dawson LA. Efficacy and safety of radiotherapy for primary liver cancer[J]. Chin Clin Oncol, 2021, 10(1): 9.
- [4] Eccles C, Brock KK, Bissonnette JP, et al. Reproducibility of liver position using active breathing coordinator for liver cancer radiotherapy[J]. Int J Radiat Oncol Biol Phys, 2006, 64(3): 751-759.
- [5] Shin HC, Roth HR, Gao MC, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning[J]. IEEE Trans Med Imaging, 2016, 35(5): 1285-1298.
- [6] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2015: 3431-3440.
- [7] Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks[J]. Med Phys, 2017, 44(2): 547-557.
- [8] Gao RT, Zhao S, Aishanjiang KD, et al. Deep learning for differential diagnosis of malignant hepatic tumors based on multi-phase contrast-enhanced CT and clinical data[J]. J Hematol Oncol, 2021, 14(1): 154.
- [9] Lyu F, Ma AJ, Yip TC, et al. Weakly supervised liver tumor segmentation using couinaud segment annotation[J]. IEEE Trans Med Imaging, 2022, 41(5): 1138-1149.
- [10] Wong KC, Moradi M, Tang H, et al. 3D segmentation with exponential logarithmic loss for highly unbalanced object sizes[C]// Medical Image Computing and Computer Assisted Intervention-MICCAI 2018. Cham: Springer International Publishing, 2018: 612-619.
- [11] Howard AG, Zhu ML, Chen B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17). <https://arxiv.org/abs/1704.04861>.
- [12] He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 770-778.
- [13] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 7132-7141.
- [14] Kingma DP, Ba J. Adam: a method for stochastic optimization[EB/OL]. (2017-01-30). <https://arxiv.org/abs/1412.6980>.
- [15] Ronneberger O, Fischer P, Brox T, et al. U-Net: convolutional networks for biomedical image segmentation[EB/OL]. (2015-05-18) [2021-12-01]. <https://arxiv.org/abs/1505.04597>.
- [16] Milletari F, Navab N, Ahmadi SA. V-Net: fully convolutional neural networks for volumetric medical image segmentation[C]//2016 Fourth International Conference on 3D Vision (3DV). Piscataway, NJ, USA: IEEE, 2016: 565-571.
- [17] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. IEEE Trans Pattern Anal Mach Intell, 2020, 42(8): 2011-2023.
- [18] 邓金城, 彭应林, 刘常春, 等. 深度卷积神经网络在放射治疗计划图像分割中的应用[J]. 中国医学物理学杂志, 2018, 35(6): 621-627.
Deng JC, Peng YL, Liu CC, et al. Application of deep convolution neural network in radiotherapy planning image segmentation[J]. Chinese Journal of Medical Physics, 2018, 35(6): 621-627.
- [19] 门阔, 戴建荣. 利用深度反卷积神经网络自动勾画放疗危及器官[J]. 中国医学物理学杂志, 2018, 35(3): 256-259.
Men K, Dai JR. Automatic segmentation of organs at risk in radiotherapy using deep deconvolutional neural network[J]. Chinese Journal of Medical Physics, 2018, 35(3): 256-259.
- [20] 张倩雯, 陈明, 秦玉芳, 等. 基于3D ResUnet网络的肺结节分割[J]. 中国医学物理学杂志, 2019, 36(11): 1356-1361.
Zhang QW, Chen M, Qin YF, et al. Lung nodule segmentation based on 3D ResUnet network[J]. Chinese Journal of Medical Physics, 2019, 36(11): 1356-1361.

(编辑:陈丽霞)