

基于空洞空间金字塔池化的U-Net网络在肺部图像分割上的应用

夏文静¹, 周腊珍¹, 陈红池¹, 李坊佐^{1,2}, 吴頔³, 张翔^{1,2}

1. 赣南医学院医学信息工程学院, 江西 赣州 341000; 2. 赣南医学院组织工程材料与生物制造江西省重点实验室, 江西 赣州 341000; 3. 赣南医学院第一附属医院, 江西 赣州 341000

【摘要】目的:胸部X线图像中肺野的自动分割是相关疾病筛查和诊断的关键步骤,为了适应计算机辅助诊断系统的要求,提出一种基于空洞空间金字塔池化的U-Net网络对胸部X线图像中肺野进行自动分割。**方法:**在编码和解码之间引入带有空洞卷积的空间金字塔池化用于扩大接受域;同时,在多个尺度上获取图像上下文信息,用于从胸片中分割肺野,使用Montgomery数据集及深圳数据集进行验证。根据医学图像分割常用指标准确性、Dice相似系数及交并比评价基于空洞空间金字塔池化的U-Net网络分割肺野的性能。**结果:**验证准确性为98.29%,Dice相似系数为96.61%,交并比为93.47%。**结论:**本文提出一种基于空洞空间金字塔池化的U-Net网络用于分割肺野,相较于其他方法学习到更多边缘分割特征,取得更好的分割结果。

【关键词】胸部X线图像;肺野分割;U-Net;空洞空间金字塔池化

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2023)03-0336-06

Lung field segmentation using U-Net based on atrous spatial pyramid pooling

XIA Wenjing¹, ZHOU Lazhen¹, CHEN Hongchi¹, LI Fangzuo^{1,2}, WU Ting³, ZHANG Xiang^{1,2}

1. School of Medical Information Engineering, Gannan Medical University, Ganzhou 341000, China; 2. Key Laboratory of Biomaterials and Biofabrication in Tissue Engineering, Gannan Medical University, Ganzhou 341000, China; 3. The First Affiliated Hospital of Gannan Medical University, Ganzhou 341000, China

Abstract: Objective The auto-segmentation of lung fields in chest X-ray images is a critical step in the screening and diagnosis of related diseases. In order to meet the requirements of computer-aided diagnosis system, a U-Net network based on atrous spatial pyramid pooling is proposed to automatically segment lung fields in chest X-ray images. **Methods** Atrous spatial pyramid pooling was introduced between encoding and decoding to enlarge the receptive field, and the image context was obtained at multiple scales to segment lung fields in chest radiographs. Montgomery chest X-ray set and the Shenzhen chest X-ray set were used for validation. The performance of atrous spatial pyramid pooling based U-Net in lung field segmentation was evaluated by the commonly used evaluation criteria for medical image segmentation (accuracy, Dice similarity coefficient and intersection over union). **Results** The validation showed that the accuracy, Dice similarity coefficient, and intersection over union were 98.29%, 96.61%, and 93.47%, respectively. **Conclusion** Compared with other methods, U-Net based on atrous spatial pyramid pooling for lung field segmentation learns more edge segmentation features, and achieves better segmentation results.

Keywords: chest X-ray image; lung field segmentation; U-Net; atrous spatial pyramid pooling

【收稿日期】2022-09-21

【基金项目】国家自然科学基金(11865003);江西省自然科学基金(20192BAB212008);赣南医学院科研启动基金(QD201805)

【作者简介】夏文静,硕士研究生,研究方向:深度学习和医学图像处理,E-mail: pzxi Xiaxia@163.com

【通信作者】李坊佐,博士,讲师,硕士生导师,研究方向:医学影像技术,E-mail: lfz880920@163.com

前言

胸部X光片(Chest X-rays, CXR)是筛查大量肺部疾病(如结核病、肺癌、肺炎和肺气肿等)最常用的诊断方法之一。与计算机断层扫描(CT)和磁共振成像(MRI)相比,CXR具有易于获取、辐射剂量低和价格低等优点,是肺部成像的首选技术。计算机辅助诊断系统中,CXR中肺野的分割起着重要的作用。例如,解剖结构和肺面积可以为疾病的早期表现提供重要的信息,包括心脏肿大、肺气肿等,肺野分割

的性能在这类应用中起着至关重要的作用。由于年龄、性别和健康状况等因素的影响,肺部的准确分割一直是医学图像分析的难点之一。外部物体的存在,如心脏起搏器、心脏支架和手术夹等,也使肺部的自动分割变得更加困难^[1]。肺部分割方法主要有4大类:①基于规则的模型,使用预定义的解剖规则对肺部进行分割^[2];②基于像素的方法,将每个像素标记为肺或肺外组织^[3];③基于可变形的模型,使用对象形状和图像外观来分割^[4];④基于配准的模型,使用分割的肺部数据库匹配肺部组织^[5]。

随着深度学习的发展,深度学习和深度卷积神经网络在肺部分割中的应用日益突出^[6]。深度神经网络(Deep Neural Network, DNN)使用前一层的特征来学习更高级别的特征^[7]。全卷积网络(Fully Convolutional Neural Networks, FCN)是Shelhamer等提出的一种新颖的视觉模型,可以提高语义分割任务的性能。然而,FCN具有较大的接受域和较弱的边缘约束,导致分割精度较低^[8]。U-Net网络的精度高,是Ronneberger等^[9]提出的用于语义分割的端到端编码器-解码器网络。在FCN中,池化层对特征进行降维,减小了下一层输入特征图的尺寸。在U-Net结构中,编码器部分使用最大池化进行下采样,将图像编码为多级特征表示,解码器模块通过反卷积操作进行上采样。多个分割映射和前向特征映射在平行层上连接,通过上采样操作扩展特征大小以满足原始图像大小。在网络收缩路径中,经过两次卷积层之后应用最大池化层,下采样之后的卷积层会使特征映射的数量加倍,这对于获取分割所需的输入图像的信息是非常必要的。Alkhonaini等^[10]将空洞空间金字塔池化(ASPP)模块加入到U-Net网络中,使用不同扩张率的空洞卷积来捕捉多尺度信息从而提高获取全局信息的能力^[11]。

本文提出一种基于ASPP的U-Net肺部分割算法并在公开数据集上对所提出的模型进行评估,提出的肺分割算法由编码器网络、解码器网络和ASPP组成。

1 研究方法

1.1 空洞卷积

图像分割是一个像素级的输出,需要上采样还原到原始图像的大小。这种处理方法有两个关键点:一是通过池化过程减小特征图的大小,二是通过上采样过程恢复原图尺寸。通过池化和上采样过程,会丢失很多原始图像的信息,导致无法达到理想的分割效果。

针对上述问题,本文算法引入了ASPP模块。

ASPP使用多个不同采样率的并行空洞卷积层有效地实现映射。空洞卷积是一种用于密集特征提取和视野放大的卷积算法。普通卷积与空洞卷积的对比如图1所示,标准卷积是扩张率等于1的特例。卷积核大小为 k ,扩张率为 r 的空洞卷积在连续的滤波器值之间引入 $r-1$ 个零,接受野 K 的大小为:

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

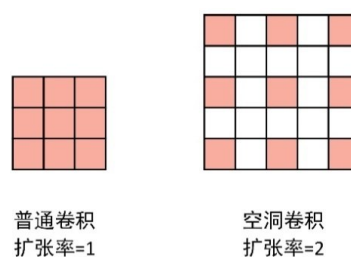


图1 普通卷积和空洞卷积

Figure 1 Conventional convolution and atrous convolution

相比于普通卷积,采用空洞卷积在不损失特征图分辨率的情况下扩大接受域,同时也不会增加参数数量,可以很好地应用空洞卷积获取图像全局信息。

1.2 批归一化层(Batch Normalization, BN)

批归一化层BN在深度学习领域中非常有效。它通过计算 mini-batch 内的均值和方差对特征进行归一化,加快网络的训练收敛速度,提升网络稳定性,并使层数非常深的网络收敛^[12]。在实践中,BN需要处理足够大的 batch size,因为 mini-batch 会导致统计信息的不准确估计。由于硬件的限制,在实际操作中需要减小 batch size。

1.3 激活函数

ReLU 函数对 DNN 的成功做出了很大贡献。ReLU 函数取代了 sigmoid 和 tanh 等饱和激活函数,主要是由于它的非饱和特性,可以解决梯度消失问题,同时加快了收敛速度。

非饱和激活函数(如ReLU函数)是实现DNN重要的一步,使用ReLU函数的DNN比使用饱和激活函数的同类网络训练得更快。ReLU函数在输入大于0时,直接输出该值;输入小于0时,输出值0。ReLU函数与饱和激活函数相比具有优越性,其正部分的导数为常数^[13]。因此,ReLU函数不会受到梯度消失的影响。ReLU激活函数的公式如下:

$$f(x)=\max(0, x) \quad (2)$$

1.4 本文算法

本文提出一种基于深度卷积网络的肺分割算法,网络结构如图2所示,主要由编码器网络、解码器网络和ASPP 3部分组成。

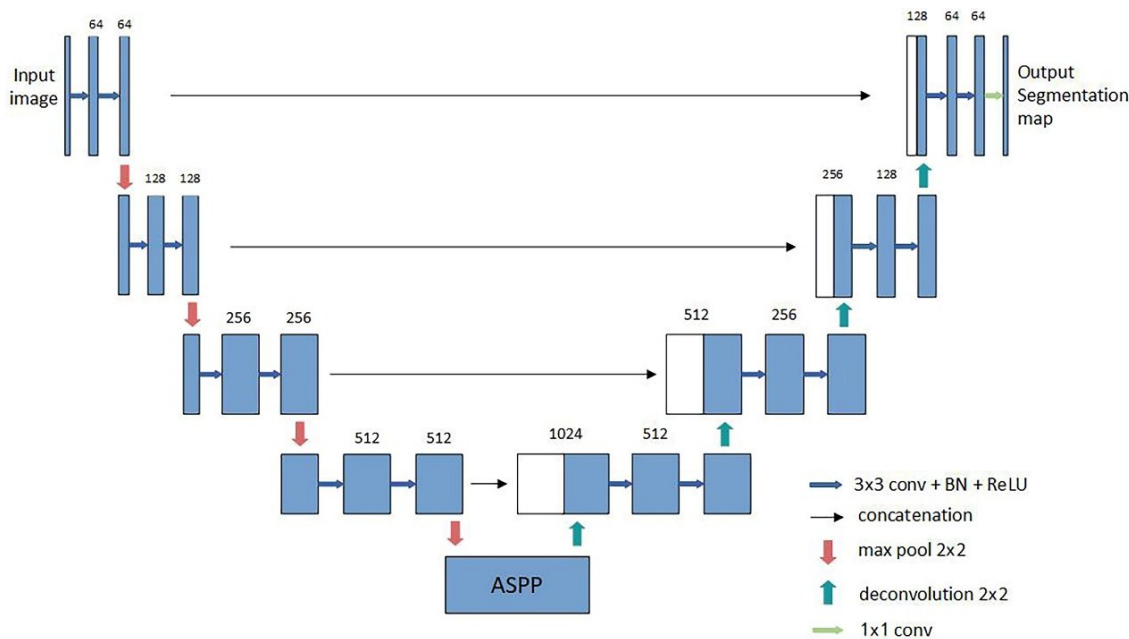


图2 本文提出的网络结构
Figure 2 The proposed network structure

U-Net 结构在生物医学图像处理领域取得了巨大的进展。U-Net 体系结构主要分为编码器和解码器两个部分。将图像尺寸设为 224×224 输入编码器,在编码和解码路径的卷积后面加上 BN 模块加快收敛速度,并使用 ReLU 函数进行激活。编码模块使用池化进行下采样,解码模块通过转置卷积操作对特征图进行上采样,上采样和下采样之间通过跳跃连

接进行连接。ASPP 网络结构以实现高效的密集语义标记性能而闻名,它旨在处理多尺度上的目标分割问题^[14]。该网络结构由多个空洞卷积分支组成,每个分支具有不同的扩张率,ASPP 在不同尺度下的目标检测方面表现出更好的性能。本文算法采用 ASPP 替代 U-Net 网络的底层。ASPP 网络结构如图 3 所示。

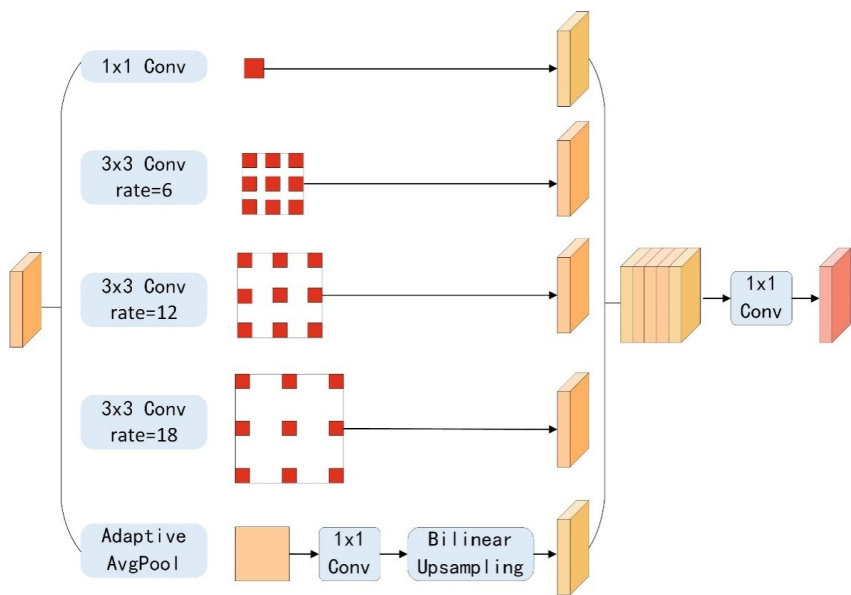


图3 ASPP 结构
Figure 3 Atrous spatial pyramid pooling structure

本文的网络中,ASPP 对上一层特征图进行了 5 次采样操作:采用普通的 1×1 卷积核提高非线性表达

能力,扩张率分别为 6、12、18 的 3×3 空洞卷积增加感受野。最后一次卷积操作,引入平均池化,将特征图

送入1×1卷积核,利用双线性插值上采样到图像原始大小。

网络各层空洞卷积与多尺度特征提取相关联,不同的扩张率对应不同尺度的特征图^[15]。每个分支都是相互独立的,较小的扩张率用于关联小范围信息,而较大的扩张率用于关联大范围信息。在网络的最后阶段,将金字塔模型得到的不同尺度特征图进行拼接,将融合了5个多尺度的特征图送入1×1卷积层中,最后将输出的特征图送入到解码模块中。

2 结果与分析

2.1 实验数据集

本研究使用两个公共数据集进行训练和验证。(1)Montgomery数据集包含138张胸部X光片,其中80张正常图像,58张结核病图像。(2)深圳数据集由深圳市第三人民医院收集,包含662张图像,其中326张正常图像,336张结核病图像^[16]。验证集占20%,训练集占80%。

2.2 模型超参数

Dice损失函数被广泛应用于语义分割中,但有时它的微分在数值上不稳定,可能导致训练过程中的振荡。结合Dice损失函数和二分类交叉熵损失函数(Binary Cross-entropy, BCE)可以避免此问题^[17]。因此,结合Dice损失函数 L_{dice} 和二分类交叉熵损失函数 L_{BCE} 作为最终损失函数 L_{seg} :

$$L_{seg}=L_{dice}+\alpha L_{BCE}$$
(3)

本实验中, α 值为1。

2.3 实验设置

实验基本工作平台为一块NVIDIA GeForce RTX 3060 Ti GPU显卡,GPU的内存为8 GB。采用PyTorch搭建模型,模型训练20个epochs。学习率设置为0.000 1,选择Adam作为训练过程优化器^[18],batch size设置为3。

2.4 评价指标

为了更好地评价和分析实验中肺分割效果,需明确实验中的评价指标和评价方法^[19]。医学图像中常用的分割指标包括:Dice相似系数(DSC)、准确性(ACC)反映模型正确区分所有样本的能力,交并比(IOUS)度量两个集合间的相似性。计算公式如下:

$$DSC = \frac{2TP}{2TP + FP + FN}$$
(4)

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$
(5)

$$IOU = \frac{TP}{TP + FP + FN}$$
(6)

其中,TP为真阳性,表示正确分类为正样本的正样本

数(即肺部区域分类为肺部);TN为真阴性,表示正确分类为负样本的负样本数(即背景区域分类为背景);FP为假阳性,表示错误分类为正样本的负样本数(即背景区域分类为肺部);FN为假阴性,表示错误分类为负样本的正样本数(即肺部区域分类为背景)。

2.5 分割效能分析

本文实验在公开数据集上进行实验,实验参数包含Dice(训练相似系数)、Val_dice(验证相似系数)、Accuracy(训练准确性)、Val_Accuracy(验证准确性)、IoU(训练交并比)、Val_IoU(验证交并比)、Loss(训练损失)及Val_Loss(验证损失)。损失函数曲线如图4所示,准确性曲线如图5所示,Dice相似系数曲线如图6所示,交并比曲线如图7所示。

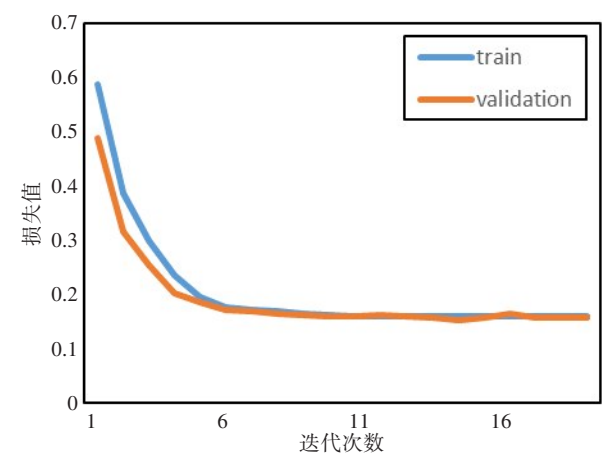


图4 损失函数曲线
Figure 4 Loss function curve

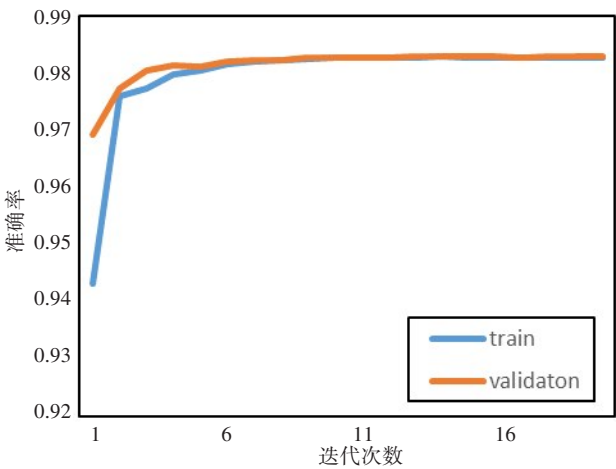


图5 准确性曲线
Figure 5 Accuracy curve

在训练的20个epochs内,经历5个epochs后,模型的准确性达到98%以上。Dice相似系数、准确性

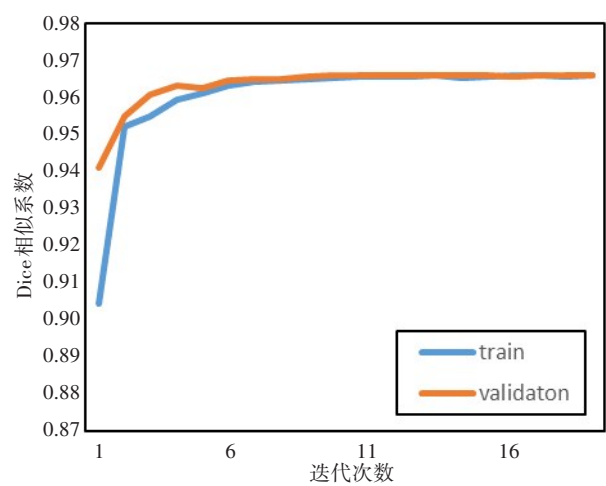


图 6 Dice 相似系数曲线
Figure 6 Dice similarity coefficient curve

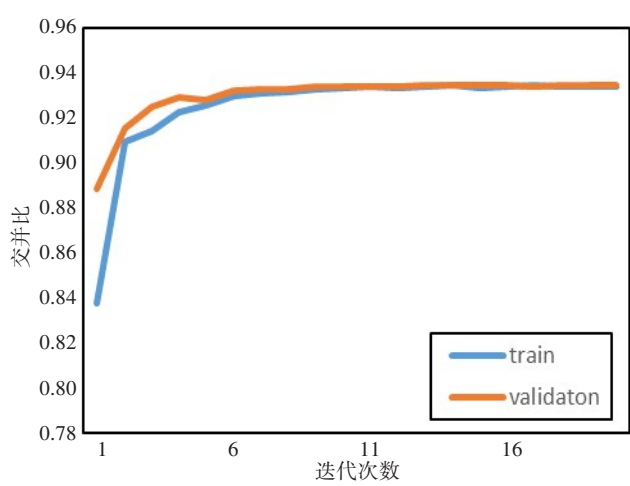


图 7 交并比曲线
Figure 7 IoU curve

和交并比随 epoch 的增加而增加。训练过程中 Dice 相似系数、准确性和交并比分别达到 96.59%、98.28%、93.44%，验证过程中相似系数、准确性和交并比分别达到 96.61% 和 98.29%、93.47%。Loss 随 epochs 的增加而减小，训练过程最小 loss 为 15.84%，

验证过程最小 Loss 为 15.69%。随机选取了 3 个样本在本文模型上进行验证，分割结果如图 8 所示。Label Image 指的是金标准，Predicted Image 指的是本文模型分割结果。

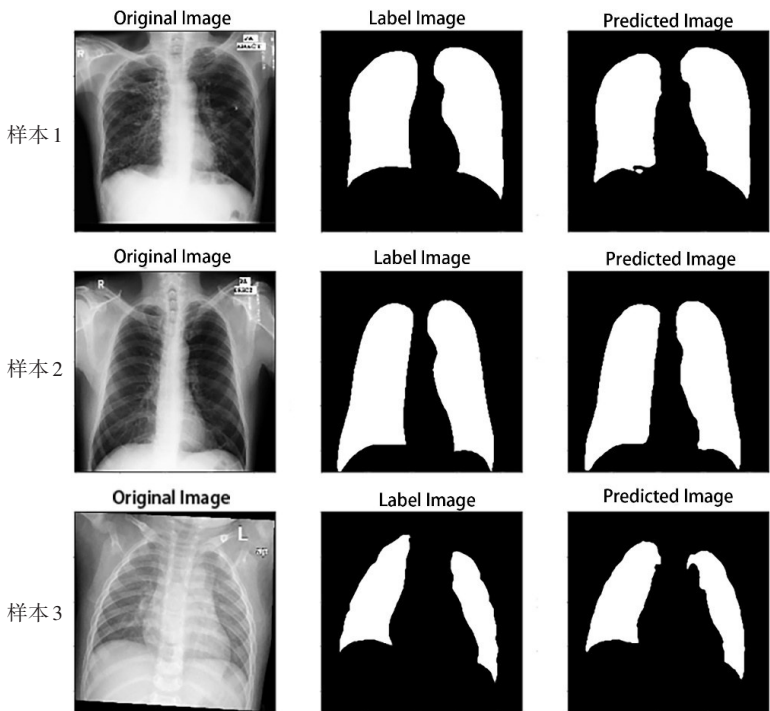


图 8 分割结果图
Figure 8 Segmentation results

2.6 与其他算法的比较

表 1 显示本文算法与经典算法 FCN、SegNet、U-Net 进行比较^[20]。结果显示，本文算法的分割评价指标准确性、Dice 相似系数和交并比均优于其余常见的分割算法，说明本文提出的分割模型用于胸部 X

光片的肺野分割性能更好。

3 讨论

基于经典 U-Net 模型在医学图像处理领域的优势，本文提出一种改进的 U-Net 的肺分割算法。利用

表1 不同分割算法评价指标的比较

Table 1 Comparison among different segmentation algorithms

算法	ACC/%	DSC/%	IOU/%
FCN	78.32	69.62	53.43
SegNet	84.09	79.14	65.58
U-Net	95.55	92.17	85.72
本文算法	98.29	96.61	93.47

空洞卷积提取高分辨率的密集图像特征图,并利用ASPP模块扩大接受域,同时在多个尺度上获取图像上下文信息。该方法在公开数据集上验证Dice相似系数为96.61%,与其余常见方法相比分割效果更好。然而,改进的U-Net网络参数非常庞大,训练非常耗时。下一阶段的任务是进一步提升肺分割精度,提出科学合理的策略提高训练时效。

【参考文献】

[1] Islam J, Zhang Y. Towards Robust lung segmentation in chest radiographs with deep learning[J]. arXiv e-prints, 2018, arXiv: 1811.12638.

[2] Armato SG, Giger ML, Macmahon H. Automated lung segmentation in digitized posteroanterior chest radiographs[J]. Acad Radiol, 1998, 5(4): 245-255.

[3] Ginnekne BV, Stegman MB, Loog M. Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database[J]. Med Image Anal, 2006, 10(1): 19-40.

[4] Shao YQ, Gao YZ, Guo YR, et al. Hierarchical lung field segmentation with joint shape and appearance sparse learning[J]. IEEE Trans Med Imaging, 2014, 33(9): 1761-1780.

[5] Candemir S, Jaeger S, Palaniappan K, et al. Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration[J]. IEEE Trans Med Imaging, 2014, 33(2): 577-590.

[6] 俞益洲, 石德君, 马杰超, 等. 人工智能在医学影像分析中的应用进展[J]. 中国医学影像技术, 2019, 35(12): 1808-1812.

Yu YZ, Shi DJ, Ma JC, et al. Advances in application of artificial intelligence in medical image analysis[J]. Chinese Journal of Medical Imaging Technology, 2019, 35(12): 1808-1812.

[7] Mittal A, Hooda R, Sofat S. LF-SegNet: a fully convolutional encoder-decoder network for segmenting lung fields from chest radiographs[J]. Wireless Pers Commun, 2018, 101(1): 511-529.

[8] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Trans Pattern Anal Mach Intell, 2017, 39(4): 640-651.

[9] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-assisted Intervention. Cham: Springer, 2015: 234-241.

[10] Alkhonaini MA, Hassine SB, Obayya M, et al. Detection of lung tumor using ASPP-Unet with whale optimization algorithm [J]. Comput Mater Con, 2022, 72(2): 3511-3527.

[11] Chen LC, Papandreou G, Kokkinos I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Trans Pattern Anal Mach Intell, 2018, 40(4): 834-848.

[12] Ge WY, Lin Y, Wang ZT, et al. An improved U-Net architecture for image Dehazing[J]. IEICE Trans Inf Syst, 2021, 104(12): 2218-2225.

[13] Jiang XH, Pang YW, Li XL, et al. Deep neural networks with Elastic Rectified Linear Units for object recognition[J]. Neurocomputing, 2018, 275: 1132-1139.

[14] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions [J]. arXiv e-prints, 2015, arXiv: 1511.07122.

[15] Chen J, Wang C, Tong Y. AtICNet: semantic segmentation with atrous spatial pyramid pooling in image cascade network[J]. Eurasip J Wirel Comm, 2019(1): 146.

[16] Jaeger S, Candemir S, Antani S, et al. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases [J]. Quant Imaging Med Surg, 2014, 4(6): 475-477.

[17] Zhao X, Zhang P, Song F, et al. D2A U-Net: automatic segmentation of COVID-19 CT slices based on dual attention and hybrid dilated convolution[J]. Comput Biol Med, 2021, 135: 104526.

[18] Kingma D, Ba J. Adam: a method for stochastic optimization[J]. arXiv preprint arXiv: 1412.6980, 2014.

[19] Mittal A, Hooda R, Sofat S. Lung field segmentation in chest radiographs: a historical review, current status, and expectations from deep learning[J]. IET Image Process, 2017, 11(11): 937-952.

[20] Gites S, Mishra A, Kotecha K. Enhanced lung image segmentation using deep learning[J]. Neural Comput Applic, 2022. DOI: 10.1007/s00521-021-06719-8. Online ahead of print.

(编辑:黄开颜)