



仪器仪表学报
Chinese Journal of Scientific Instrument
ISSN 0254-3087, CN 11-2179/TH

《仪器仪表学报》网络首发论文

题目: 基于改进 VNet 的肺结节分割方法研究
作者: 钟思华, 王梦璐, 郭兴明, 张瑶, 郑伊能
收稿日期: 2020-6-30
网络首发日期: 2020-09-27
引用格式: 钟思华, 王梦璐, 郭兴明, 张瑶, 郑伊能. 基于改进 VNet 的肺结节分割方法研究. 仪器仪表学报.
<https://kns.cnki.net/kcms/detail/11.2179.TH.20200925.1712.008.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式 (包括网络呈现版式) 排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊 (光盘版)》电子杂志社有限公司签约, 在《中国学术期刊 (网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊 (网络版)》是国家新闻出版广电总局批准的网络连续型出版物 (ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

基于改进 VNet 的肺结节分割方法研究*

钟思华^{1,2}, 王梦璐¹, 郭兴明¹, 张瑶¹, 郑伊能³

(1.重庆大学生物工程学院 重庆市医疗电子工程技术研究中心 重庆 400044; 2.上海联影医疗科技有限公司 中央研究院 上海 201807; 3.重庆医科大学附属第一医院 放射科 重庆 400044)

摘要：目前，肺癌的发病率和致死率高居癌症首位，对其进行早期诊治对于提高患者的生存率和改善预后极其重要。肺结节是肺癌的早期表现，临床上医生通过观察对其分割后的体积和形态等特征来进行良恶性诊断，然而采用人工的方式进行肺结节分割非常低效。本文提出了一种基于 MSVNet 网络的肺结节分割方法，该网络继承了原始 VNet 的结构，同时引入了多尺度特征结构，通过提取肺结节图像的多尺度特征，同时利用深监督策略进行特征优化，能够有效地提升模型的分割性能。本文利用 LIDC-IDRI 肺结节公开数据集对模型的性能进行了评估，结果表明，本文方法所取得的分割结果与金标准相近，具有良好的肺结节分割性能，以及较高的分割鲁棒性，对不同大小的肺结节均能取得较好的分割效果。

关键词：肺结节分割；深度学习；多尺度特征；深监督

中图分类号：TP391.41 TH7 **文献标识码：**A **国家标准学科分类代码：**520.6040

Study on the improved VNet network based pulmonary nodule segmentation method

Zhong Sihua^{1,2}, Wang Menglu¹, Guo Xingming¹, Zhang Yao¹, Zheng Yineng³

(1. Chongqing Engineering Research Center for Medical Electronics Technology, College of Biological Engineering, Chongqing University, Chongqing 400044, China;

2. Corporation Research Center, Shanghai United Imaging Healthcare Limited Company, Shanghai 201807, China;

3. Department of Radiology, the First Affiliated Hospital of Chongqing Medical University, Chongqing 400044, China)

Abstract: At present, the incidence and mortality of lung cancer are the highest among cancers. Early diagnosis and treatment of lung cancer are extremely important to improve the survival rate and prognosis of patients. Pulmonary nodules are the early manifestation of lung cancer, which are clinically diagnosed as benign or malignant by observing the characteristics such as volume and morphology after segmentation. However, manual segmentation of pulmonary nodules is very inefficient. In this study, a segmentation method of pulmonary nodules based on MSVNet network is proposed, which inherits the structure of the original VNet, meanwhile a multi-scale feature structure is introduced. Through extracting the multi-scale feature of pulmonary nodule image and optimizing the feature with deep supervision strategy, the segmentation performance of the model can be effectively improved. In this study, the performance of the model is evaluated using the LIDC-IDRI pulmonary nodule public data set. The results show that the segmentation results obtained with the proposed method are similar to the gold standard. The proposed method has good pulmonary nodule segmentation performance and high segmentation robustness, and can achieve good segmentation results for the pulmonary nodules with different sizes.

Key words: pulmonary nodule segmentation; deep learning; multi-scale feature; deep supervision

0 引言

肺癌是最常见的一种癌症类型，同时其致死率也

收稿日期：2020-6-30

* 基金项目：国家自然科学基金项目（31570003）资助

是最高的,2018年,全球新增210万例肺癌患者,因肺癌死亡的人数达177万例^[1]。因此对肺癌进行早期诊治对于提高患者的生存率及改善预后尤为重要。肺结节是肺癌的早期表现,它是一种由多种系统和器官共同引发的肉芽肿性疾病,在肺部多呈球状结节,直径通常不超过30毫米^[2]。目前国内肺结节诊治标准将肺结节中,直径小于5毫米的界定为微小结节,直径在5毫米至10毫米之间的界定为小结节,直径大于10毫米的界定为大结节^[3]。临床上,医生通过观察患者胸部计算机断层扫描(Computed tomography, CT)图像中肺结节的体积、形态等特征来对其良恶性进行诊断。随着CT成像技术的飞速发展,CT图像的数据量越来越大,以人工阅片的方式检测并分割肺结节给医生带来了沉重的工作负担;另一方面,肺结节的尺寸较小,且其灰度特征和形态特征与肺实质中的血管等组织相近,这容易对医生的判断产生干扰。在这种情况下,对肺结节的自动分割方法进行研究具有十分重要的临床价值。

目前,针对肺结节分割方法的研究主要分为两类:基于手动提取特征的传统方法和基于自动提取特征的深度学习方法。

基于手动提取特征的传统方法主要是根据肺结节图像的特点,构建出相应的特征并用于提取候选肺结节区域,包括阈值法^[4]、形态学法^[5]、模糊聚类法^[6]和模板匹配法^[7]等。虽然上述方法均能达到较好的肺结节分割效果,但是它们依赖于人对于肺结节的先验知识。不同的人所提取的肺结节特征以及分割模型参数的设定存在差异,导致最终的肺结节分割结果受人的主观性影响较大。

目前,深度学习广泛应用于医学图像处理领域并取得了成功^[8-9]。深度神经网络借助大数据进行训练,对于输入网络的图像数据能够自动地学习并提取特征,最终实现肺结节分割等任务。钟思华等人^[10]利用经典分割网络UNet进行肺结节分割,同时引入了密集连接机制以提高网络的分割效果;Huang等人^[11]在完成肺结节检测后,使用改进的FCN(Fully Convolutional Network, FCN)^[12]对肺结节进行分割,该网络通过融合不同尺度的肺结节特征,增强了对特征的利用,有效地提升了网络肺结节分割性能;Qin等人^[13]搭建了一种肺结节分割模型,模型包括两部分,首先,利用生成对抗网络产生模拟肺结节图像,增大训练样本量;随后,通过三维卷积神经网络提取多个肺结节特征用于肺结节分割,与其他方法相比,该模型能够准确地对多种类型的肺结节进行分割;由于医学图像具有目

标尺寸小、边缘模糊的特点,使用传统网络难以实现高效分割,因此,Zhou等人^[14]提出充分利用UNet提取出的多尺度图像特征进行网络优化,能够有效提升网络性能,该方法的有效性在肺结节等医学图像上得到印证。

VNet^[15]是一种三维的全卷积神经网络,首次提出后应用于前列腺图像的分割,并取得了较好的效果。该网络的结构与UNet类似,均为对称的编码器-解码器结构,能对图片中每个像素类别进行预测。在医学图像处理中,诸如肺结节等病灶具有三维结构,使用VNet能够充分利用此类图像的三维特征,提升网络的性能。然而肺结节体积较小,原始VNet提取特征困难,容易出现漏检、错检等情况,无法取得满意的分割结果。

因此,本文结合肺结节图像的特点,在VNet网络结构的基础上,提出了一种多尺度VNet网络(multi-scale VNet, MSVNet)用于肺结节分割。该网络通过提取不同尺度下的肺结节特征,并用于网络训练,能够有效提升肺结节分割的准确性。

1 MSVNet 肺结节分割原理

1.1 肺结节分割方法框架

本文用于肺结节分割的框架如下图所示。

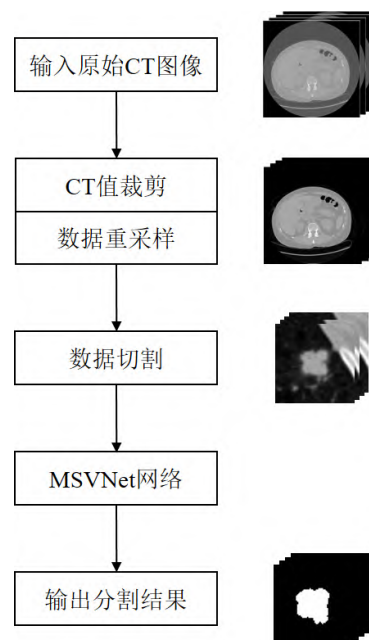


图1 肺结节分割框架

Fig. 1 Framework of pulmonary nodule segmentation

1) CT值裁剪:原始的胸部CT图像的CT值分布范围为-1000Hu~+3000Hu,这给病灶的查找与观察带来了极大的困难。为了凸显病灶区域,依据临床上常用的肺窗CT值范围,将原始输入图像的CT值裁剪到

-1000Hu 到+600Hu 之间;

2) 数据重采样: 为了消除因 CT 图像数据的采样间隔不统一而带来的实验误差, 对 CT 图像数据进行重采样, 分别将沿着 x 轴、y 轴和 z 轴三个方向的数据采样间距调整为 1mm;

3) 数据切割: 由于肺结节区域在 CT 图像上的占比远小于无关背景区域, 因此, 为了减少因训练数据类不平衡对模型训练过程造成的影响, 同时减少模型训练所占用的显存资源, 在肺结节所在位置的周围截取至少包含一个肺结节的三维图像块, 其尺寸为 $64 \times 64 \times 32$ (高度 \times 宽度 \times 深度);

4) 输出分割结果: 将尺寸为 $64 \times 64 \times 32$ 的三维肺结节图像块输入训练好的 MSVNet 模型中, 输出相同尺寸的二值分割结果。

1.2 MSVNet 网络结构

1.2.1 VNet 网络结构

MSVNet 是在 VNet 的结构基础上进行改进, VNet 的结构如图 2 所示。

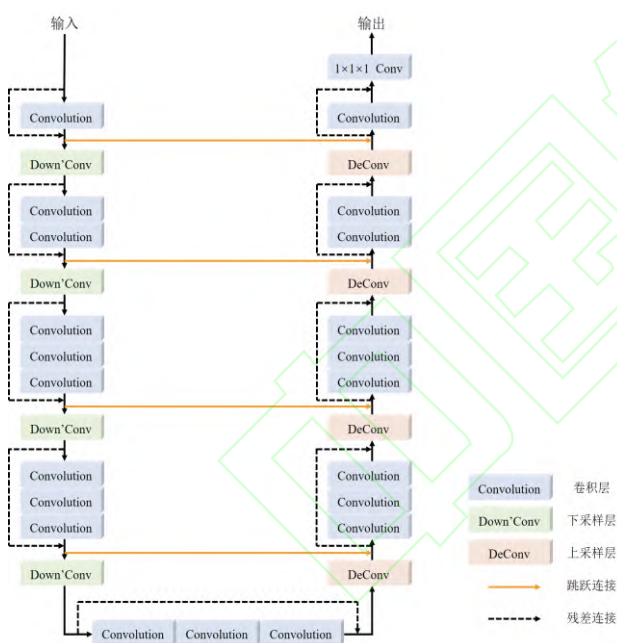


图 2 VNet 网络结构

Fig. 2 VNet network structure

VNet 网络左边为编码器, 负责对输入的图像数据进行编码, 提取图像中的特征。编码器部分包含了四个具有不同特征分辨率的编码模块, 每个编码模块由数量不等的卷积层和下采样层组成。卷积层用于提取图像的语义特征, 浅层的卷积层能够提取简单特征, 而深层的卷积层则能够提取复杂特征。为了提高特征的利用程度, 网络还使用了残差连接 (residual connection) 将模块中的第一个卷积层的输入与最后一

个卷积层的输出相连; 下采样层用于特征图的下采样操作, 通过缩小特征图的分辨率以增大其感受野 (receptive field) 的范围; 同时由于特征尺寸减小, 能够降低网络的计算量以及参数的存储。由于下采样操作只关注数据中的关键信息, 而忽略一些冗余信息, 因此能提高图像特征的鲁棒性。在下采样层中, 网络使用步长为 2 的卷积运算代替传统的池化操作, 目的是为了减小网络在训练过程中占用的内存。

网络右边为解码器, 负责对所提取的特征图进行解码, 恢复特征图的分辨率。解码器部分也包含了四个具有不同特征分辨率的解码模块, 每个解码模块由数量不等的卷积层和上采样层组成。卷积层的结构设计 with 编码器部分类似, 同样使用了残差连接; 上采样层利用反卷积运算对特征图进行上采样操作, 恢复特征图的分辨率。特征图每经过一次反卷积运算, 其尺寸都增大一倍, 最终可以得到与输入图像大小相同的特征图。编码器部分与解码器部分之间通过三个卷积层进行连接。

此外, 网络的分类器由 $1 \times 1 \times 1$ 卷积层和 sigmoid 激活层组成, 其中, $1 \times 1 \times 1$ 卷积层用于减少特征图的数量, sigmoid 激活层用于计算最终的特征图中每个像素的类别, 输出网络的分割概率图。跳跃连接 (skip connection) 将网络中的浅层简单特征与深层抽象特征融合起来, 可得到更为精细的分割结果。

1.2.2 多尺度特征结构

在 VNet 网络的特征提取阶段, 特征分辨率不同的编码模块具有大小不同的感受野, 所提取的目标对象的特征也不同。特征分辨率大的模块可以提取肺结节的一些简单特征, 如边缘、颜色等; 而特征分辨率小的模块因为感受野变大, 而且经过了多次卷积运算, 可以提取肺结节的复杂特征。经过四个模块提取到复杂特征后, VNet 网络才进行上采样操作将所提取的特征恢复到与输入图像相同的尺寸大小。在这一过程中, 网络只关注了单一尺度下的复杂特征, 忽略了其他尺度的特征对肺结节检测任务的贡献。肺结节常因其尺寸较小而容易出现漏检的问题, 为了提高网络的分割性能, 需要充分利用不同尺度下的肺结节特征。

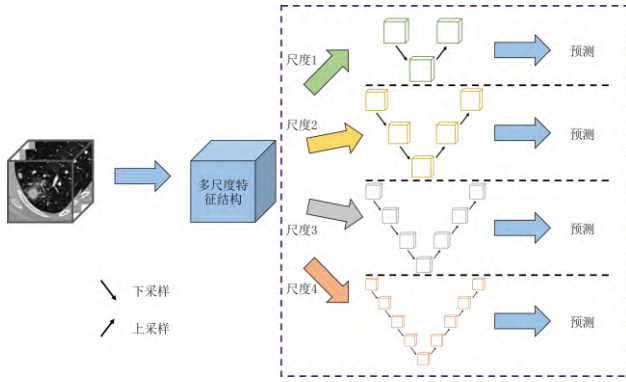


图3 多尺度特征结构

Fig. 3 Multi-scale feature structure

多尺度特征的结构如图3所示,在网络每提取一个尺度的特征后,便将该特征进行上采样操作直至恢复到与输入图像相同的尺寸。由多尺度特征网络所得到的多尺度特征,既有助于像素定位的简单特征又有用于像素分类的复杂特征,利用这些多尺度特征可以有效解决肺结节漏检的问题。

1.2.3 MSVNet 网络结构

本文在原始 VNet 网络的基础上,结合多尺度特征的结构特点,提出了一种改进的 MSVNet 网络来进行肺结节分割,以此提高分割的准确率,该网络的结构如图4所示。

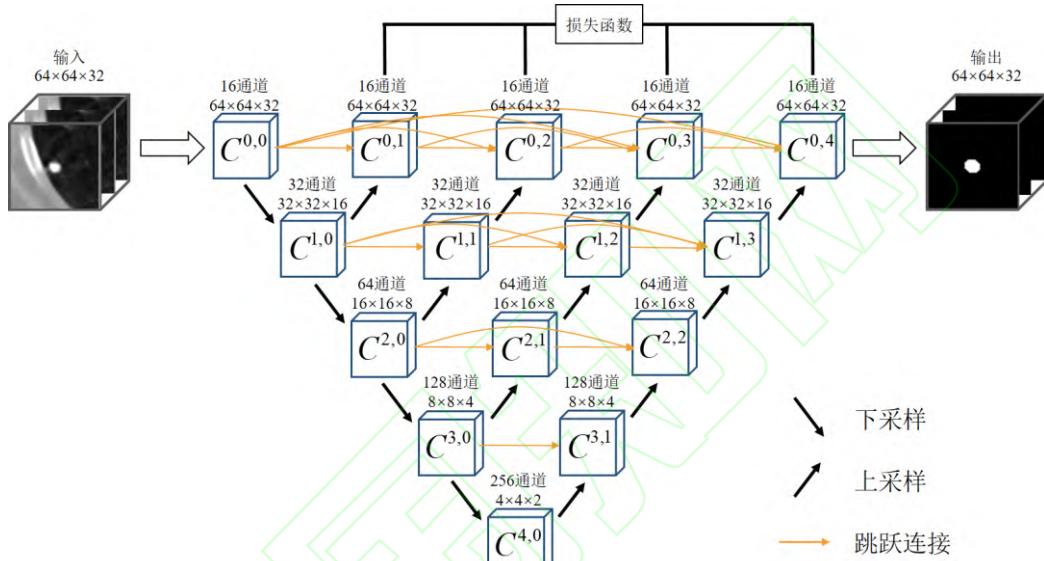


图4 MSVNet 网络结构

Fig. 4 MSVNet network structure

图中, $C^{i,j}$ 表示网络中的节点,用来代替 VNet 网络中的卷积层及其残差连接结构, i 表示编码器中下采样层的序号, j 表示具有相同特征分辨率的节点的序号。网络节点均使用 $3 \times 3 \times 3$, 步长为 1 的卷积核对特征图进行卷积运算,下采样由 $3 \times 3 \times 3$, 步长为 2 的卷积核进行卷积运算完成,上采样则由 $3 \times 3 \times 3$, 步长为 2 的卷积核进行反卷积运算完成。MSVNet 网络由四个不同尺度的子网络组合而成,每个子网络都具有对称的编码器—解码器结构。当图像输入进网络时,首先经由编码器分别提取四个不同尺度的特征,随后通过四个不同的解码器得到四个预测结果,并将第四个尺度的预测结果作为网络最终的输出结果。

该网络在所有卷积层后面均添加了群组归一化层^[16]和 ReLu 非线性激活层^[17],以此提高网络的性能。其中,群组归一化层能够准确地估算批量较小的输入数据的统计信息,并进行归一化操作,从而解决了因内存限制无法输入大批量数据的问题。对于格式为

$[N, C, H, W]$ 的输入数据 x_i , 其中 N 为批量大小, C 为通道数, H 和 W 分别为数据的高度和宽度,则群组归一化层分别通过式(1)、(2)实现对 x_i 的标准归一化和转换重构操作。

$$\hat{x}_i = \frac{x_i - \mu_i}{\sqrt{\sigma_i^2 + \varepsilon}} \quad (1)$$

$$y_i = \gamma \hat{x}_i + \beta \quad (2)$$

式中, μ_i 为均值, σ_i^2 为方差, ε 为平滑因子以避免分母为零, γ 和 β 为可学习重构参数, y_i 即为网络对输入数据 x_i 进行归一化操作后的输出值。 μ_i 和 σ_i^2 的计算方式分别如下式所示。

$$\mu_i = \frac{1}{(C/G)HW} \sum_{c=gC/G}^{(g+1)C/G} \sum_{h=1}^H \sum_{w=1}^W x_i \quad (3)$$

$$\sigma_i^2 = \frac{1}{(C/G)HW} \sum_{c=gC/G}^{(g+1)C/G} \sum_{h=1}^H \sum_{w=1}^W (x_i - \mu_i)^2 \quad (4)$$

式中, C/G 代表将输入数据的 C 个通道分成 G 组。

而 ReLu 激活层则负责为网络引入了非线性因素, 由此提高了网络的非线性表达能力。其计算方式如式(5)所示。

$$ReLU = \max(0, x) \quad (5)$$

同时, 网络中还使用了 Dropout 操作^[18], 以此避免网络训练过程中出现过拟合问题。

为了减少网络在上采样过程中带来的特征信息损失, 同时加快网络的收敛速度, 利用跳跃连接将网络中具有相同特征分辨率的节点连接起来, 具体连接方式如下图所示。

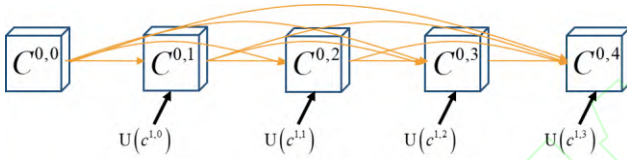


图5 网络节点的连接方式

Fig. 5 Connection mode of network nodes

因此, 网络中每个节点的输出结果可由式(6)计算得到。

$$c^{i,j} = \begin{cases} F(c^{i-1,j}), & j=0 \\ F\left(\left[c^{i,k}\right]_{k=0}^{j-1}, U(c^{i+1,j-1})\right), & j>0 \end{cases} \quad (6)$$

式中, $c^{i,j}$ 为节点 $C^{i,j}$ 的输出; $F(\cdot)$ 表示对输入进行卷积运算, 并利用非线性激活函数对卷积结果进行非线性变换; $U(\cdot)$ 表示对输入进行上采样操作; $[\cdot]$ 表示对输入进行叠加操作。

最后, 为了让四个子网络都能够得到充分的训练, 获得更加精确的预测结果, 本文采用深监督(deep supervision)机制对网络进行训练, 即将损失函数同时引入所有子网络当中, 分别计算每个网络的损失值并同时优化。

1.3 损失函数

本文采用 Focal Tversky Loss^[19]作为损失函数进行网络训练。由于 Tversky 相似系数可以表示为:

$$Tversky = \frac{|X \cap Y|}{|X \cap Y| + \alpha |X - Y| + (1 - \alpha) |Y - X|} \quad (7)$$

因此, Focal Tversky Loss 的计算方式如下所示:

$$L_{FT} = \sum (1 - Tversky)^{1/\gamma} \quad (8)$$

式中, X 为网络对目标区域的预测结果; Y 为目标区域的真实结果, $|X - Y|$ 代表网络预测出的假阳性区域, $|Y - X|$ 则表示网络预测出的假阴性区域; α 和 γ

为权重因子, 通过调节网络预测区域的权重, 使训练样本中数量少及难以学习的类别得到充分训练。特别地, 当 $\alpha = 0.5$, $\gamma = 1$ 时, Focal Tversky Loss 为 Dice Similarity Coefficient (DSC) Loss, 该损失函数常应用于医学图像分割中。DSC 的表达式为:

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|} \quad (9)$$

2 实验与分析

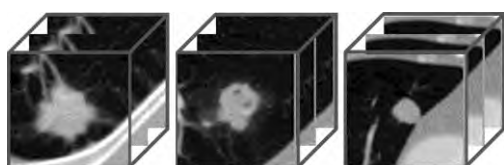
2.1 实验数据

本实验的数据来自于 LIDC-IDRI(Lung Image Database Consortium and Image Database Resource Initiative)胸部影像数据集^[20]。对于该数据集中, 直径大于或等于 3mm 的肺结节, 共有四名经验丰富的医生对其轮廓等信息分别进行了标注, 并将综合后的轮廓标注结果保存至 XML 文件中。因此, 纳入本实验的数据需满足两个条件, 一是 CT 影像的切片厚度不大于 2.5mm, 二是 CT 影像所包含的肺结节的直径大于或等于 3mm。对符合实验条件的数据进行预处理操作, 包括 CT 值裁剪、数据重采样和切割, 然后根据 XML 中医生对肺结节轮廓的标注, 生成与肺结节图像尺寸相同的肺结节轮廓标签。最终可以得到 728 例尺寸为 $64 \times 64 \times 32$ 的三维肺结节样本及其相对应的轮廓标签, 并按照 8: 1: 1 的比例将其随机划分为训练集、验证集和测试集。实验数据的分布统计如表 1 所示, 肺结节图像与轮廓标签如图 6 所示。

表 1 实验数据的分布统计

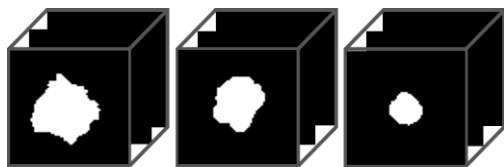
Table 1 The distribution statics of experiment data

	训练集/个	验证集/个	测试集/个	总计/个
肺结节 [3,5]	150	20	21	191
直径 [5,10]	343	44	43	430
/mm (10,30]	90	9	8	107
总计/个	583	73	72	728



(a) $64 \times 64 \times 32$ 肺结节图像

(a) Pulmonary nodule images with size of $64 \times 64 \times 32$



(b) $64 \times 64 \times 32$ 肺结节轮廓标签图像

(b) Pulmonary nodule label images with size of $64 \times 64 \times 32$

图 6 实验数据

Fig. 6 Experiment data

2.2 实验环境

本文的实验环境如下, 处理器: Intel(R) Core(TM) i7-8700K@3.70GHz; 内存: 64GB; 显卡: RTX 2080Ti, 11GB; 操作系统: Windows 10; 开发语言: Python 3.5; 深度学习框架: Tensorflow 1.13.1。

2.3 评价指标

实验选取 Dice 相似系数值(DSC)、敏感度(Sensitivity)和准确度(Precision)作为肺结节分割的评价指标。其中, Dice 相似系数的计算方式如式(9)所示, 敏感度和准确度的计算方式分别如式(10)、(11)所示。

$$Sensitivity = \frac{|X \cap Y|}{|X \cap Y| + |Y - X|} \quad (10)$$

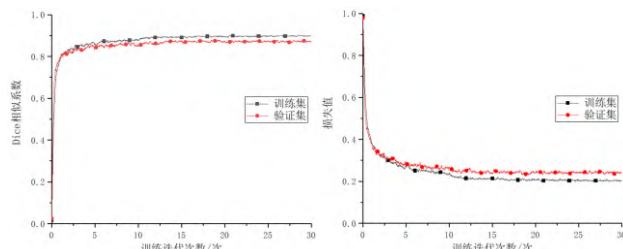
$$Precision = \frac{|X \cap Y|}{|X \cap Y| + |X - Y|} \quad (11)$$

2.4 参数设置

为了确定模型的训练参数, 实验采用 MSVNet 作为基本模型进行研究, 并利用验证集调整模型参数。

2.4.1 训练迭代次数

训练迭代次数对于网络训练至关重要, 本文通过观察网络训练和验证曲线的变化趋势, 采用早停策略确定该参数。由图 7 可知, 当训练迭代次数达到 30 次时, 网络在验证集上的 Dice 相似系数值和损失值曲线的变化趋于稳定, 因此, 将训练迭代次数设置为 30 次。此外, 为了让网络得到充分训练, 将初始学习率设置为 0.01, 每经过 10 次迭代训练, 学习率衰减为原来的 0.1 倍。



(a) Dice 相似系数值曲线

(b) 损失值曲线

(a) Dice similarity coefficient value curve

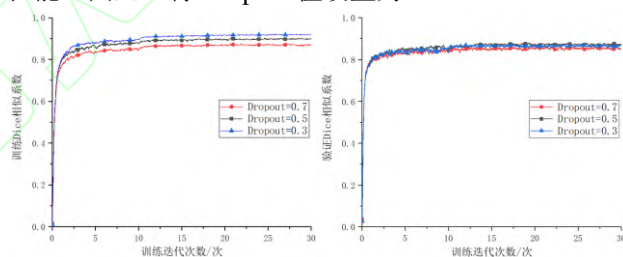
(b) Loss value curve

图 7 网络训练和验证曲线

Fig. 7 Network training and verification curves

2.4.2 Dropout 值

Dropout 是指在网络训练过程中, 以一定的概率随机丢弃网络中的神经元, 可以有效避免网络训练出现过拟合的问题。本文分别探究了当 Dropout 为 0.3、0.5 和 0.7 时, 网络的训练和验证 Dice 相似系数值的变化情况。由图 8(a)可知, 随着 Dropout 值的升高, 网络在训练集上的 Dice 相似系数值不断降低。图 8(b)表明, 当 Dropout 为 0.5 时, 网络在验证集上具有最好的 Dice 相似系数值。而 Dropout 值过高或过低均会导致网络训练过程中出现过拟合或欠拟合的情况, 影响网络的性能。因此, 将 Dropout 值设置为 0.5。



(a) 训练 Dice 相似系数值

(b) 验证 Dice 相似系数值

(a) Dice similarity coefficient value for training

(b) Dice similarity coefficient value for verification

图 8 网络训练和验证 Dice 相似系数值

Fig. 8 Dice similarity coefficient values for network training and verification

2.4.3 权重因子

本文对损失函数中的权重因子 α 的取值进行了探究, 结果如表 2 所示。从表中可知, α 的取值决定了训练样本中所有类别的收敛程度, α 越小, 模型预测的准确度越高, α 越大, 则模型预测的敏感度越高。当 α 为 0.5 时, 模型的整体性能达到最优。因此, 将 α 设置为 0.5。此外, 根据文献[18], 将 γ 设置为 0.75。

表 2 权重因子对模型性能的影响

Table 2 The influence of weight factors on the model performance

权重因子	DSC(%)	Sensitivity(%)	Precision(%)
0.3	85.47	81.42	92.39
0.5	87.21	88.13	88.17
0.7	85.47	91.40	82.10

3 实验结果

根据 2.4 节中的实验结果设置网络参数,利用测试集对 3D UNet^[21]、VNet、未使用深监督策略的 MSVNet (MSVNet-od)以及本文的使用深监督策略的 MSVNet (MSVNet-wd)四种网络结构进行对比分析,结果如表 3 所示。

表 3 不同模型的性能对比

Table 3 Performance comparison of different models

损失函数	模型	DSC(%)	Sensitivity(%)	Precision(%)
DSC Loss	3D UNet	86.81	89.91	85.09
	VNet	86.21	88.51	86.64
	MSVNet-od	86.81	86.74	89.13
	MSVNet-wd	87.40	88.89	88.37
FT Loss	3D UNet	78.12	89.46	85.89
	VNet	78.99	89.93	86.34
	MSVNet-od	86.29	90.12	85.45
	MSVNet-wd	88.18	89.87	88.67

注: DSC Loss 为 Dice Similarity Coefficient (DSC) Loss; FT Loss 为 Focal Tversky Loss

首先,对比了不同损失函数对模型性能的影响。由表可知,对于三种网络:3D UNet、VNet 和 MSVNet-od,使用 Focal Tversky Loss 进行优化训练后所取得的分割性能,相比 DSC Loss 均出现了不同程度的下降。但是,对于本文的方法 MSVNet-wd,使用 Focal Tversky Loss 进行优化训练后所取得的性能最优,这是因为该网络具有四个尺度的子网络,而每个尺度网络所提取的肺结节特征的侧重点不同,因此,不同尺度网络的预测结果中会存在不同数量的假阳性或假阴性区域。利用 Focal Tversky Loss 对四个尺度的子网络进行深监督,可以有效地优化子网络的分割性能,减少子网络的预测结果中无关区域的数量,使得 MSVNet 网络整体性能达到最优。反观另外三种网络,在训练过程中均无法充分利用多尺度特征,使用 Focal Tversky Loss 进行网络训练反而会影响其优化过程,最终影响网络的分割性能。

其次,对比了不同网络结构的肺结节分割性能。由表可知,本文的 MSVNet-wd 能够取得最优的肺结节分割结果,其中,DSC 为 88.18%,Sensitivity 为 89.87%,Precision 为 88.67%。由于 3D UNet 和 VNet 两种网络

结构只提取了一个尺度的肺结节特征,对于肺结节等小目标的分割任务具有较大的难度,难以取得准确的分割结果。相比 MSVNet 结构,两者所取得的 Dice 相似系数值最低,分别为 78.12%和 78.99%,这说明多尺度特征结构能够提取更多不同尺度的肺结节特征,从而提升网络的分割性能。而对于 MSVNet-od,因其在网络训练过程中没有使用深监督策略,通过不同子网络所提取的多尺度特征无法得到充分地优化,从而影响该网络的分割性能。本文所提出的 MSVNet-wd 网络,利用多尺度特征结构提取肺结节图像的多尺度特征,同时结合深监督策略优化所提取的多尺度特征,最终实现了最优的肺结节分割结果。

图 9 展示了以上四种网络在三种肺结节直径([3, 5)、[5, 10]、(10, 30]mm)范围下,对测试集数据预测所取得的平均 Dice 相似系数值。从图中可知,在三种肺结节直径范围下,多尺度网络相较单一尺度网络均能显著提高网络的分割性能,而 MSVNet-wd 的分割性能最优,并且对于不同大小肺结节的分割效果基本保持一致。

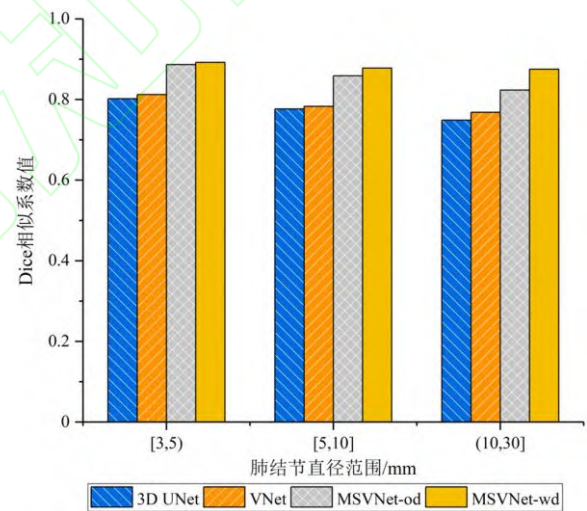


图 9 不同大小肺结节的分割结果

Fig. 9 Segmentation results of the pulmonary nodules with different sizes

此外,本文还统计了四种网络对测试集数据预测所取得的 Dice 相似系数值的分布情况,结果如图 10 所示。从图中可知,本文提出的方法所取得的 Dice 相似系数值主要集中于 0.8 至 1 之间,且绝大部分分布在 0.9 至 1 之间。

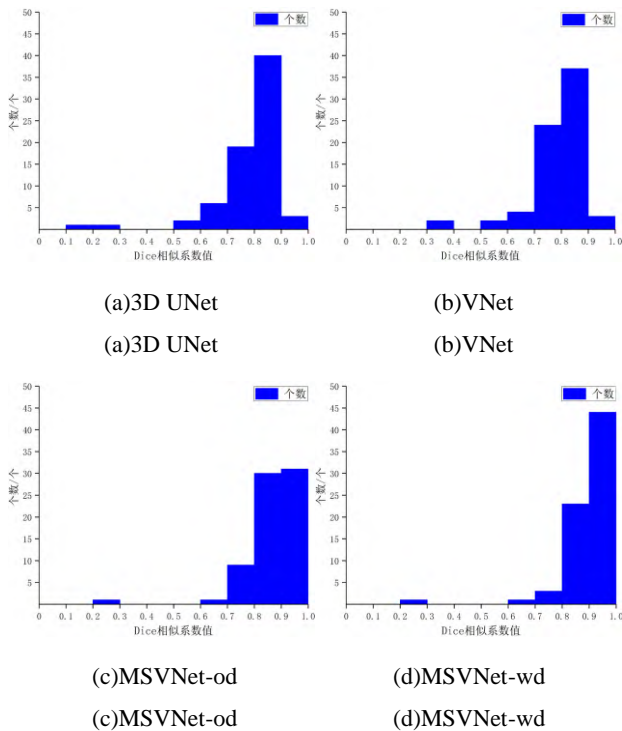


图 10 预测结果中 Dice 相似系数值的分布情况

Fig. 10 The distribution of Dice similarity coefficient values in the prediction results

综合以上实验结果,证明本文方法具有良好的肺结节分割性能,以及较高的分割鲁棒性,对不同大小的肺结节均能取得较好的分割效果。

图 11 展示了四种网络对五个肺结节图像的三维分割结果,其中第一行代表医生对肺结节轮廓标注的金标准,肺结节直径从左到右依次增大。从图中可以直观地看出,虽然 3D UNet、VNet 和 MSVNet-od 等三种网络能够将肺结节区域分割出来,但分割结果中存在大量的假阳性区域;而本文方法通过提取肺结节的多尺度特征,并利用深监督策略优化网络,能够有效区分肺结节区域与其他背景区域,因此其分割准确率较高。

最后,为了进一步验证本文方法的有效性,将本文所提出的 MSVNet 网络与其他肺结节分割方法在 LIDC-IDRI 数据集上的实验结果进行对比,结果如表 4

所示。从表中可知,本文方法在 Dice 相似系数值、敏感度和准确度等三个评价指标上的表现均优于其他方法,表明本文所提出的基于 MSVNet 网络的肺结节分割方法能够较准确地实现肺结节分割,分割效果与目前先进的分割算法相当。

表 4 不同分割方法的结果对比

Table 4 Result comparison of different segmentation

methods				
类型	作者	DSC(%)	Sensitivity(%)	Precision(%)
传统算法	Shakir ^[22]	85.69	—	—
	Shakibapour ^[23]	82.35	87.11	85.59
	Hesamian ^[24]	81.24	—	79.75
深度学习	Liu ^[25]	81.58	87.3	79.71
	Qin ^[13]	84.83	88.95	85.11
	本文	88.18	89.87	88.67

4 结论

本文提出了一种基于 MSVNet 网络的肺结节分割方法,该网络继承了原始 VNet 网络的结构特点,同时引入了多尺度特征结构。在网络训练阶段,MSVNet 能够提取肺结节图像的多尺度特征,并利用深监督策略对所提取的多尺度特征进行优化,因此网络得以充分利用肺结节的特征,准确地区分肺结节区域与无关背景区域,最终实现准确的肺结节分割结果。本文选取 Dice 相似系数值,敏感度和准确度作为评价指标,在 LIDC-IDRI 肺结节公开数据集上对模型性能进行评估,取得的分割结果为 DSC: 88.18%, Sensitivity: 89.87%, Precision: 88.67%。实验结果表明,本文方法能够准确地对不同大小的肺结节进行分割,分割性能与目前先进的分割算法相当。但是,本文用于模型训练的样本量还不够充足,在下一步研究中可通过数据增强的方法来增加训练样本量;此外,还可进一步研究不同类型的肺结节对模型性能的影响。

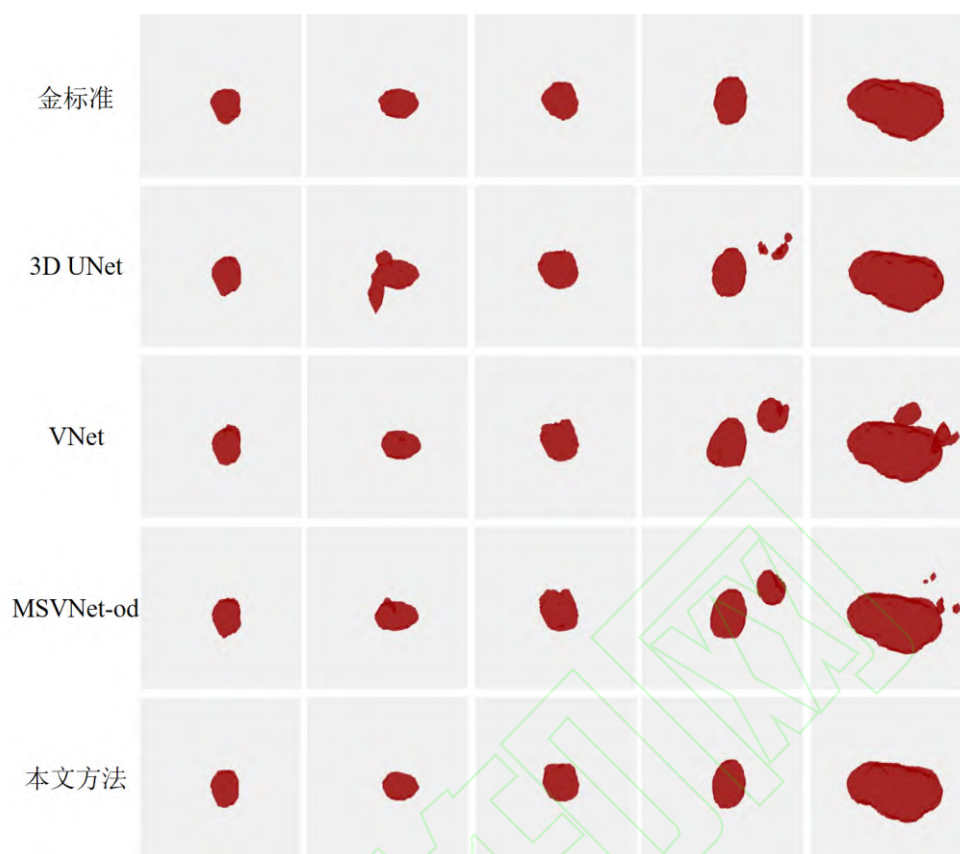


图 11 不同网络的三维分割结果

Fig. 11 Three dimensional segmentation results for different networks

参考文献

- [1] BRAY F, FERLAY J, SOERJOMATARAM I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA: A Cancer Journal for Clinicians, 2018, 68(6): 394-424.
- [2] 周清华, 范亚光, 王颖, 等. 中国肺结节分类、诊断与治疗指南(2016 年版)[J]. 中国肺癌杂志, 2016, 19(12): 793-798.
ZHOU Q H, FAN Y G, WANG Y, et al. China national guideline of classification, diagnosis and treatment for lung nodules [J]. Chin J Lung Cancer, 2016, 19(12): 793-798.
- [3] 张晓菊, 白莉, 金发光, 等. 肺结节诊治中国专家共识(2018 年版)[J]. 中华结核和呼吸杂志, 2018, 41(10): 763-771.
ZHANG X J, BAI L, JIN F G, et al. China experts consensus on the diagnosis and treatment of lung nodules (2018 Version)[J]. Chinese Journal of Tuberculosis and Respiratory Diseases, 2018, 41(10): 763-771.
- [4] ARMATO S G, GIGER M L, MORAN C J, et al. Computerized detection of pulmonary nodules on CT scans[J]. Radiographics, 1999, 19(5): 1303-1311.
- [5] 魏颖, 徐心和, 贾同, 等. 基于多尺度形态学滤波的 CT 图像疑似肺结节提取[J]. 东北大学学报 (自然科学版), 2008, 29(3): 312-315.
WEI Y, XU X H, JIA T, et al. Extraction of suspected nodular lung lesions based on multi-scale morphology filtering in thoracic CT images[J]. Journal of Northeastern University, 2008, 29(3): 312-315.
- [6] 刘慧, 张彩明, 邓凯, 等. 改进局部自适应的快速 FCM 肺结节分割方法[J]. 计算机辅助设计与图形学学报, 2014, 26(10): 1727-1736.
LIU H, ZHANG C M, DENG K, et al. Research on fast FCM pulmonary nodule segmentation algorithm using improved self-adaption[J]. Journal of Computer-Aided Design & Computer Graphics, 2014, 26(10): 1727-1736.
- [7] 陈业航, 李智, 冯宝, 等. 基于改进的活动轮廓模型的胸膜接触型肺结节分割[J]. 仪器仪表学报, 2019, 40(11): 107-116.
CHEN Y H, LI ZH, FENG B, et al. Pleural contact

- pulmonary nodule segmentation based on improved active contour model[J]. Chinese Journal of Scientific Instrument, 2019, 40(11): 107-116.
- [8] ARESTA G, JACOBS C, ARAÚJO T, et al. iW-Net: an automatic and minimalistic interactive lung nodule segmentation deep network[J]. Scientific reports, 2019, 9(1): 1-9.
- [9] ROY R, CHAKRABORTI T, CHOWDHURY A S. A deep learning-shape driven level set synergism for pulmonary nodule segmentation[J]. Pattern Recognition Letters, 2019, 123: 31-38.
- [10] 钟思华, 郭兴明, 郑伊能. 改进 U-Net 网络的肺结节分割方法[J/OL]. 计算机工程与应用:2020(17) 203-209 [2020-06-30].
<http://kns.cnki.net/kcms/detail/11.2127.TP.20200422.1137.006.html>.
- ZHONG S H, GUO X M, ZHENG Y N. Improved U-Net network for lung nodule segmentation[J/OL]. Computer Engineering and Applications: 2020(17) 203-209 [2020-06-30].
<http://kns.cnki.net/kcms/detail/11.2127.TP.20200422.1137.006.html>.
- [11] HUANG X, SUN W, TSENG T L B, et al. Fast and fully-automated detection and segmentation of pulmonary nodules in thoracic CT scans using deep convolutional neural networks[J]. Computerized Medical Imaging and Graphics, 2019, 74: 25-36.
- [12] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2015: 3431-3440.
- [13] QIN Y, ZHENG H, HUANG X, et al. Pulmonary nodule segmentation with CT sample synthesis using adversarial networks[J]. Medical physics, 2019, 46(3): 1218-1229.
- [14] ZHOU Z, SIDDIQUEE M M R, TAJBAKHS N, et al. Unet++: A nested u-net architecture for medical image segmentation[C]. 4th International Workshop on Deep Learning in Medical Image Analysis, DLMIA 2018 and 8th International Workshop on Multimodal Learning for Clinical Decision Support, ML-CDS 2018 Held in Conjunction with MICCAI 2018. Springer Verlag, 2018: 3-11.
- [15] MILLETARI F, NAVAB N, AHMADI S A. V-net: fully convolutional neural networks for volumetric medical image segmentation[C]. 2016 Fourth International Conference on 3D Vision (3DV). IEEE, 2016: 565-571.
- [16] WU Y, HE K. Group normalization[C]. 15th European Conference on Computer Vision (ECCV). Springer, 2018: 3-19.
- [17] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines[C]. Proceedings of the 27th International Conference on Machine Learning (ICML-10). Omnipress, 2010: 807-814.
- [18] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. The journal of machine learning research, 2014, 15(1): 1929-1958.
- [19] ABRAHAM N, KHAN N M. A novel focal tversky loss function with improved attention u-net for lesion segmentation[C]. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). IEEE, 2019: 683-687.
- [20] ARMATO III S G, MCLENNAN G, BIDAUT L, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans[J]. Medical physics, 2011, 38(2): 915-931.
- [21] ÇİÇEK Ö, ABDULKADIR A, LIENKAMP S S, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2016: 424-432.
- [22] SHAKIR H, KHAN T M R, RASHEED H. 3-D segmentation of lung nodules using hybrid level sets[J]. Computers in Biology and Medicine, 2018, 96: 214-226.
- [23] SHAKIBAPOUR E, CUNHA A, ARESTA G, et al. An unsupervised metaheuristic search approach for segmentation and volume measurement of pulmonary nodules in lung CT scans[J]. Expert Systems with Applications, 2019, 119: 415-428.
- [24] HESAMIAN M H, JIA W, HE X, et al. Atrous convolution for binary semantic segmentation of lung nodule[C]. 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019: 1015-1019.
- [25] LIU H, CAO H, SONG E, et al. A cascaded dual-pathway residual network for lung nodule segmentation in CT images[J]. Physica Medica, 2019, 63: 112-121.

作者简介



钟思华, 分别在 2017 年、2020 年于重庆大学获得学士、硕士学位, 现为上海联影医疗科技有限公司科研合作专家, 主要研究方向为医学图像处理。

E-mail: zhongsh@cqu.edu.cn

Zhong Sihua received his B. Sc. and M. Sc. degrees both from Chongqing University in 2017 and 2020, respectively. Now, he is a research collaboration specialist in Shanghai United Imaging Healthcare Co., Ltd. His main research interest is medical imaging processing.



郭兴明 (通讯作者), 分别在 1984 年、1991 年和 1994 年于重庆大学获得学士、硕士和博士学位, 现为重庆大学教授、博士生导师, 主要研究方向为生物医学信号检测及仪器、远程医疗及医学图像处理。

E-mail: guoxm@cqu.edu.cn

Guo Xingming (corresponding author) received his B. Sc., M. Sc. and Ph. D. degrees all from Chongqing University in 1984, 1991 and 1994, respectively. Now, he is a professor and Ph. D. supervisor in Chongqing University. His main research field covers biomedical signal detection and instrument, telemedicine and, medical image processing.



郑伊能, 分别在 2011 年和 2017 年于重庆大学获得学士和博士学位, 现为重庆医科大学附属第一医院助理研究员, 主要研究方向为医学信号处理与医学图像处理。

E-mail: yinengzheng@cqmu.edu.cn

Zheng Yineng received his B. Sc. and Ph. D. degrees both from Chongqing University in 2011 and 2017, respectively. Now, he is a research assistant in First Affiliated Hospital, Chongqing Medical University. His main research field covers medical signal processing and medical imaging processing.