

改进全卷积神经网络的甲状腺结节分割方法

张雅婷, 帅仁俊, 黄道宏, 赵 宸, 吴梦麟

(南京工业大学计算机科学与技术学院, 南京 211816)

摘要: 为了更加精确地分割出甲状腺结节, 本文提出了一种改进的全卷积神经网络(Fully convolutional network, FCN)分割模型。相较于FCN, 本文方法加入了空洞空间卷积池化金字塔(Atrous spatial pyramid pooling, ASPP)模块与多层特征传递模块(Feature transfer, FT), 并采用LinkNet模型中Decoder模块进行上采样, VGG16主干网络实现特征提取下采样。实验采用来自斯坦福AIMI(Artificial intelligence in medicine and imaging)共享数据集的17 413张超声甲状腺结节图像分别用于训练、验证和测试。实验结果表明, 相比于其他多种分割模型, 本文模型在平均交并比(mean Intersection over union, mIoU), Dice相似系数, F_1 分数3个分割指标上分别达到了79.7%, 87.6%和98.42%, 实现了更好的分割效果, 有效地提升了甲状腺结节的分割精确度。

关键词: 甲状腺结节; 分割; 特征提取; 空洞空间卷积池化金字塔

中图分类号: TP391 **文献标志码:** A

Improved FCN Segmentation Method for Thyroid Nodules

ZHANG Yating, SHUAI Renjun, HUANG Daohong, ZHAO Chen, WU Menglin

(College of Computer Science and Technology, Nanjing Technology University, Nanjing 211816, China)

Abstract: In order to segment thyroid nodules more accurately, this paper proposes an improved fully convolutional network (FCN) segmentation model. Compared with FCN, the atrous spatial pyramid pooling (ASPP) module and the multi-layer feature transfer (FT) module are added. The decoder module in LinkNet model is used for up-sampling, and the VGG16 backbone network is used for feature extraction down-sampling. The experiment uses 17 413 ultrasound thyroid nodule images from Stanford AIMI shared data set for training, verification and testing, respectively. Experimental results show that compared with other segmentation models, the proposed model achieves 79.7%, 87.6% and 98.42% in mean intersection over union (mIoU), Dice similarity coefficient and F_1 score respectively, achieving better segmentation effect and effectively improving the segmentation accuracy of thyroid nodules.

Key words: thyroid nodules; segmentation; feature extraction; atrous spatial pyramid pooling (ASPP)

引言

甲状腺是内分泌系统的腺体,甲状腺结节是在甲状腺内形成的异常肿块。研究表明,甲状腺结节患者中大约 11% 会患上甲状腺癌^[1-3]。甲状腺癌已成为女性第五大常见癌症^[4],严重影响了人们的健康。甲状腺结节的良恶性分类依赖于结节的轮廓和纵横比等特征。其中,恶性结节通常存在浸润性不规则的边缘,造成结节组织区域与周围环境难以区分^[5]。因此准确地分割出甲状腺结节的轮廓,对其后续良恶性分类及其治疗有着重要的意义。目前,超声检查已成为临床上首选的甲状腺结节检测手段^[6],就超声甲状腺图像来说,甲状腺组织分布于肌肉和支气管之间的组织间隙,因其超声扫描时通常是采用手持式扫描仪对甲状腺进行扫描,导致其在超声影像中大小、形状及周边组织存在明显差异,同时甲状腺组织周边复杂器官干扰使得图像分割不理想^[7]。放射科医师对超声图像中结节的分割较为主观,高度依赖于医生的个人经验,因此需要一种能准确分割甲状腺结节的方法。

传统的甲状腺结节分割方法大致分为基于轮廓分割和基于区域分割两种类型。基于轮廓的方法使用超声图像中的轮廓信息来分割甲状腺结节^[8-10]。Maroulis 等^[8]提出了一种基于无边主动轮廓 (Active contour without edges, ACWE) 模型的可变背景主动轮廓 (Active contour without edges, VBAC) 模型,用于甲状腺超声图像的分割,VBAC 模型对背景分布不均的甲状腺结节实现了较好的分割效果,但对非低回声结节的分割效果欠佳。基于 VBAC 模型的联合回声纹理 (Joint echogenicity-texture, JET) 模型将纹理信息结合到活动轮廓模型中,使模型能够得到一些无法通过像素强度和平均像素强度变化来区分的发现。这种方法的实现过程虽然比较简单,但是需要预先设置初始轮廓,并且效率较低^[5]。基于区域的方法使用区域的统计特性来区分不同的组织^[11-12]。Zhao 等^[11]提出了一种基于归一化模型的甲状腺结节分割方法,但缺乏通用性和普适性。Alrubaidi 等^[12]提出了一种基于径向梯度和方差缩减统计的超声甲状腺结节分割方法,该方法需要专业医生的人工干预且效率低。基于现有研究,基于区域的方法还需要丰富的先验信息才能获得更准确的分割效果^[5, 11, 12]。

卷积神经网络在计算机图像识别领域取得了巨大成功^[13-17]。与传统的图像分割方法相比,深度学习算法大大提高了分割的准确性和算法的自动化程度。文献[18-19]等使用卷积神经网络 (Convolutional neural network, CNN),通过甲状腺边缘像素点对甲状腺和甲状腺结节进行分类,根据分类结果生成分割概率图,但分割性能不理想。文献[19-20]使用全卷积神经网络 (Fully convolutional networks, FCN) 对结节超声图像进行分割,FCN 仅关注于每个像素间的分类,忽略了空间结构和一致性。同时,使用上采样处理后的图像会变得模糊平滑,对于图像中的细节不敏感,导致分割误差较大。Zhou 等^[21]提出了一种基于 U-Net 的标记引导分割算法;Ying 等^[22]使用基于 U-Net 的改进网络对甲状腺结节超声图像进行分割,其仅在单一尺度上进行预测,无法完全解决尺度变化问题,此外,在分割一些不明显的甲状腺结节时,仍会产生较大的误差。迟剑宇等^[23]在研究中将深度网络和浅层纹理特征相结合,用于诊断甲状腺结节癌变情况,其采用编码器-解码器结构,具有结构简单、计算效率高等特点,编码器由预训练过的分类模型组成,例如 ResNet (Residual network) 和 VGG (Visual geometry group network) 能够将多个不同语义级别的抽象特征提取出来,在解码器中将这些特征组合在一起,生成图像中的目标响应区域^[7, 23]。

本文受编码器-解码器结构启发,针对甲状腺超声图像中甲状腺组织大小和形态多样性以及周边

组织的复杂性问题^[24],为探索更精确的深度学习技术以实现超声甲状腺结节的分割,提出了一种改进的FCN甲状腺结节分割模型。本文提出的模型加入多层特征传递模块 (Feature transfer, FT) 和空洞空间卷积池化金字塔 (Atrous spatial pyramid pooling, ASPP) 模块,并采用LinkNet模型中Decoder模块实现特征融合与特征图上采样。本文模型对比其他10种分割模型,在平均交并比 (mean Intersection over union, mIoU), Dice相似系数, F_1 分数3个性能指标上均实现了最优值,本文模型实现了更加精确的分割精度。

1 改进的FCN

改进的FCN由VGG16下采样模块、ASPP模块、多层特征传递模块和Decoder上采样模块4部分组成,模型结构如图1所示,模型细节如表1所示。首先,超声甲状腺结节图像作为特征提取模块VGG16的输入提取特征,同时特征图分辨率下采样32倍;其次经过ASPP模块,ASPP使用不同的膨胀率获取不同尺度的感受野,提取多尺度信息;然后经过Decoder模块实现特征图上采样,还原输入图像分辨率大小。与此同时,将来自VGG16不同感受野的特征图作为多层特征传递模块的输入再次下采样,增大感受野,然后传递给Decoder模块。一方面,Decoder模块完成特征图上采样;另一方面,Decoder模块融合多层次特征图,保留了更丰富的局部细节和抽象的语义信息。

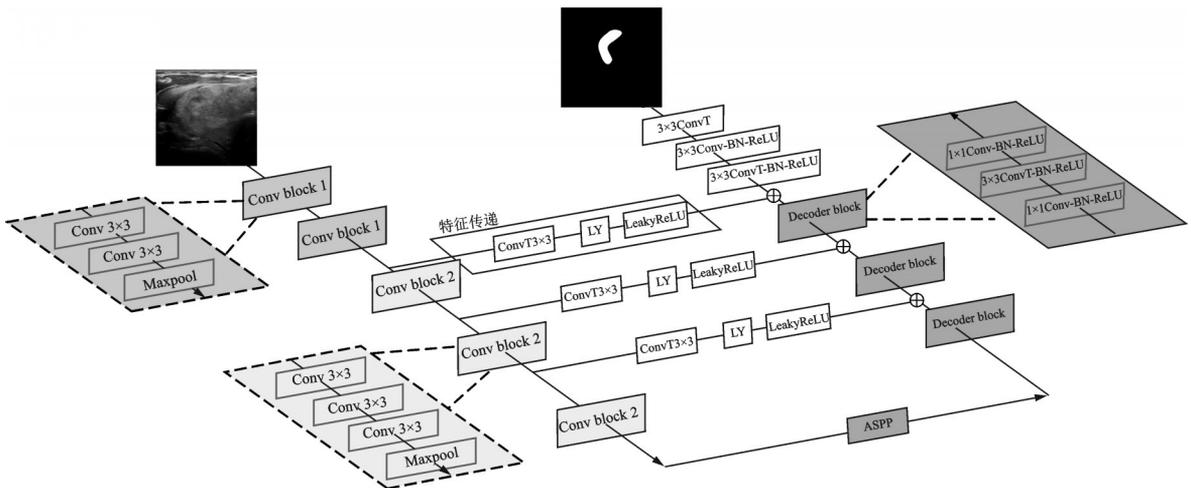


图1 改进的FCN模型结构图

Fig.1 Structure diagram of improved FCN model

1.1 VGG16

VGG16网络结构简洁,常用于主干网络提取图像特征^[25],采用VGG16作为特征提取模块可以实现更高的分割精度。假设超声甲状腺结节图像尺寸为224像素 \times 224像素 \times 3像素,经过两个卷积核大小为3,步长为1,填充为1的卷积核后,输出特征图分辨率保持不变,通道数增加为64。随后,经过步长为2的最大池化下采样层后,特征图长宽减小为1/2,通道数保持不变。输入超声图像经过13个卷积层,4个最大池化层后,特征图长宽减小为输入图像尺寸的1/32,通道数增加为512。模型细节见表1。

表 1 改进 FCN 的模型参数

Table 1 Model parameters of improved FCN

VGG16		多层特征传递模块		Decoder	
模块配置	输出分辨率/像素	模块配置	输出分辨率/像素	模块配置	输出分辨率/像素
Conv3, s1	224×224, 64				
Conv3, s1				ConvT3, s2	224×224, 2
Maxpool, s2	112×112, 64			Conv3, s1	112×112, 32
Conv3, s1	112×112, 128			ConvT3, s2	112×112, 32
Conv3, s1		Conv3, s2	56×56, 96		
Maxpool, s1	56×56, 128			Conv1, s1	56×56, 96
Conv3, s1	56×56, 256			ConvT3, s2	56×56, 192/4
Conv3, s1				Conv1, s1	28×28, 192/4
Conv3, s1		Conv3, s2	28×28, 192		
Maxpool, s2	28×28, 256			Conv1, s1	28×28, 192
Conv3, s1	28×28, 512			ConvT3, s2	28×28, 384/4
Conv3, s1				Conv1, s1	14×14, 384/4
Conv3, s1		Conv3, s2	14×14, 384		
Maxpool, s2	14×14, 512			Conv1, s1	14×14, 384
Conv3, s1				ConvT3, s2	14×14, 768/4
Conv3, s1				Conv1, s1	7×7, 768/4
Conv3, s1					
Maxpool, s2	7×7, 512				
ASPP	7×7, 768				

1.2 多层特征传递

多层特征传递包含 3 个特征 FT, 特征传递模块如图 2 所示, 特征传递模块是 1 个短连接, 由 1 个 3×3 卷积、LayerNorm 和 LeakyReLU 激活函数组成^[26-27]。一方面, 3 个特征传递模块对来自 VGG16 的不同层次特征图再次下采样, 进一步增大感受野; 另一方面, 将 VGG16 下采样得到的特征图再次下采样后传递给 Decoder 模块, 辅助不同感受野的特征信息进行融合。

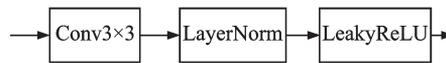


图 2 特征传递模块

Fig.2 Feature transfer module

1.3 空洞空间卷积池化金字塔

ASPP 对输入特征图以不同膨胀率的空洞卷积进行并行采样, 不同膨胀率的空洞卷积可以获取不同尺度的感受野, 提取多尺度信息^[28-29]。ASPP 结构图如图 3 所示, ASPP 由 1 个 1×1 卷积, 3 个膨胀卷积 (r 为膨胀率), 1 个池化操作组成。其中, 池化操作包含 1 个全局平均池化层, 1×1 卷积层和 1 个上采样层 (双线性插值)。

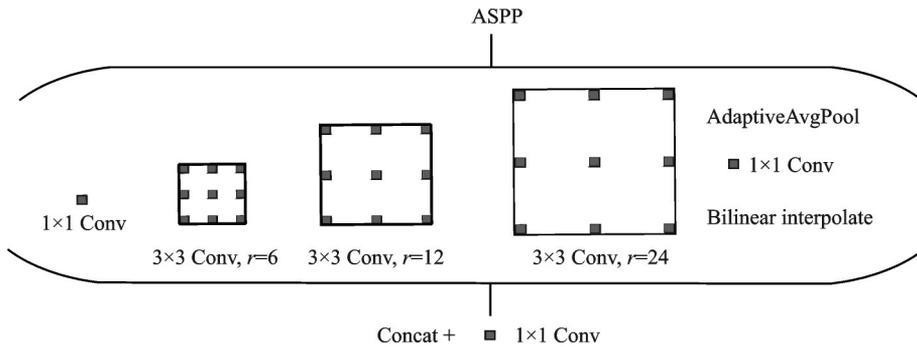


图3 ASPP结构图

Fig.3 Structure diagram of ASPP

1.4 Decoder

本文采取 LinkNet 网络中 Decoder 模块实现特征图上采样^[30], Decoder Block 是 Decoder 模块的主要组成部分。Decoder Block 如图 1 所示, Decoder Block 由 1 个 1x1 卷积块, 1 个 3x3 转置卷积块和 1 个 1x1 卷积块组成。1x1 卷积块如图 4 所示, 1x1 卷积块由 1 个 1x1 卷积、BatchNorm 和 ReLU 激活函数组成。3x3 转置卷积块如图 5 所示, 3x3 转置卷积块由 1 个 3x3 转置卷积、BatchNorm 和 ReLU 激活函数组成。经过 1 个 Decoder Block 后特征图长宽增加为原来的 2 倍, 通道数减小为 1/2。

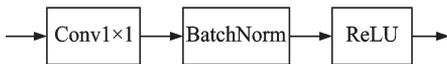


图4 1x1 卷积块

Fig.4 1x1 convolution block



图5 3x3 转置卷积块

Fig.5 3x3 transposed convolution block

首先, 特征图经过 3 个 Decoder Blocks 上采样后, 长宽增加为原来的 8 倍, 通道数减小为 1/8, 然后依次经过 1 个 3x3 转置卷积块, 1 个 3x3 卷积块和 1 个 3x3 转置卷积后, 长宽增加为原来的 32 倍, 通道数减小为 2, 实现甲状腺结节分割。一方面, Decoder 模块完成了特征图上采样; 另一方面, Decoder 模块通过融合多层次特征信息, 保留了更丰富的局部细节和抽象的语义信息。

1.5 损失函数

本文使用两种损失函数: BCE Loss 和 Dice Loss^[31-32]。BCE Loss 计算预测分割图和真实分割图之间每个像素的误差, 公式为

$$L_{BCE} = - \sum_{i=1}^H \sum_{j=1}^W [G(i,j)P(i,j) + (1 - G(i,j)) \ln(1 - P(i,j))] \tag{1}$$

式中: $P(i,j)$ 和 $G(i,j)$ 表示图像中位置 (i,j) 的预测值和真实值的像素; W 和 H 为图像的宽度和高度。

Dice Loss 本质上是对预测分割图像和真实分割图像重叠部分的度量, 公式为

$$L_{Dice} = 1 - \frac{2 \sum_{i=1}^H \sum_{j=1}^W G(i,j)P(i,j)}{\sum_{i=1}^H \sum_{j=1}^W [G(i,j)^2 + P(i,j)^2]} \tag{2}$$

本文使用 BCE Loss 和 Dice Loss 共同监督预测分割图像, 公式为

$$\text{Loss} = \alpha L_{\text{BCE}} + \beta L_{\text{Dice}} \quad (3)$$

式中 α 、 β 为超参数,用于平衡两个损失函数之间的权重,使网络能够获得更好的性能。通过实验对比,当设置超参数 α 为0.2, β 为0.8时,实现了更高的分割精度。

2 实验分析与结果

2.1 数据集

本文使用的数据集是斯坦福AIMI共享数据集,包含斯坦福大学医学中心167名经活检证实的甲状腺结节患者的记录,年龄在19岁至84岁之间(平均年龄56岁)。该数据集由17 412张超声甲状腺结节图像和放射科医生注释的17 412张分割图像组成,训练、验证和测试图像分别占比60%、20%和20%。

2.2 实验设置

本文中实验基于PyTorch1.9.0框架实现,所有的运算都在1块内存为11 GB的NVIDIA GeForce GTX 1080Ti的显卡上来加速网络训练。在训练期间,网络的最大学习率设置为0.000 1,采用学习率Warmup更新策略^[33],AdamW优化器进行模型优化^[34]。学习率变化曲线如图6所示,其中横坐标Step计算公式为

$$\text{Step} = \frac{17\,412 \times 60\%}{\text{batch_size}} \times \text{epoch} \quad (4)$$

式中:batch_size为一次训练所取的图像数量;epoch为训练迭代次数。

本文使用mIoU、Dice相似系数、 F_1 分数、FLOPs和Params作为超声甲状腺结节分割任务的评价指标^[35-37],公式分别为

$$\text{mIoU} = \frac{1}{k+1} \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k p_{ij} + \sum_{i=0}^k p_{ji} - p_{ii}} \quad (5)$$

式中: i 代表真实值; j 代表预测值; p_{ij} 表示将 i 预测为 j 的像素数量; k 表示类别; $k+1$ 表示加上背景类别。

$$\text{Dice} = \frac{2\text{TP}}{\text{FP} + 2\text{TP} + \text{FN}} \quad (6)$$

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (7)$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (8)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (9)$$

式中:TP(True positive)定义为正确分割为甲状腺结节的区域(真阳性);FP(False positive)定义为将非结节区域错误分割为结节区域(假阳性);FN(False negative)定义为错误漏分割甲状腺结节区域(假阴性)。

2.3 实验结果

为研究本文所提模型对超声甲状腺结节的分割性能,将本文模型与其他10种分割模型进行指标对比,表2为不同模型在甲状腺结节测试集图像上的分割性能指标对比结果。由表2可见,本文模型的

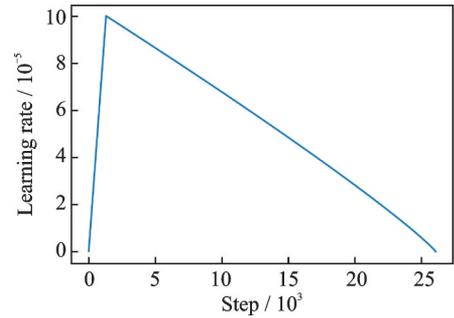


图6 学习率变化曲线

Fig.6 Learning Rate curve

表2 不同模型分割性能指标对比

Table 2 Comparison of segmentation performance indicators of different models

网络	mIoU/%	Dice/%	F_1 /%	FLOPs/GMAC	Params/MB
FCN_VGG16_8s	77.2	85.8	98.23	19.58	19.17
FCN_ResNet50_8s	73.3	82.6	97.94	26.55	32.95
FCN_ResNet101_8s	70.4	80.1	97.64	41.46	51.94
U-Net	75.3	84.3	98.01	50.1	34.51
Swin-Unet	70.8	80.7	97.07	8.65	41.38
LinkNet	71.7	80.7	97.51	1.33	12.95
Segnet	74.2	83.3	98.12	30.73	29.44
PSPNet+MobileNetV2	67.8	78.1	96.30	1.88	2.38
PSPNet+ResNet50	71.6	81.2	97.63	35.37	46.71
DeepLabv3+MobileNetV2	66.7	76.7	97.01	5.05	5.81
改进的FCN	79.7	87.6	98.42	17.84	32.17

mIoU、Dice 相似系数、 F_1 分数、FLOPs 和 Params 分别为 79.7%、87.6%、98.42%、17.84 GMAC 和 32.17 MB。相比于 FCN_VGG16_8S,改进的 FCN 在 mIoU、Dice 和 F_1 指标上分别提升了 2.5%、1.8% 和 0.19%,在 FLOPs 指标上降低了 1.74 GMAC。综合实验结果,对比其他多种分割模型,本文模型在 mIoU, Dice 和 F_1 指标上都取得了最优值,显著提升了甲状腺结节的分割性能。

不同模型在甲状腺结节测试集图像上的分割结果如图 7 所示。图 7A 列为 9 张输入甲状腺图像,分辨率为 224 像素 \times 224 像素。图 7B 例为放射科医生注释的分割图像;图 7C 例为本文模型的分割图像;图 7D~L 例依次为 FCN_VGG16_8s、FCN_ResNet50_8s、FCN_ResNet101_8s、U-Net、Swin-Unet、LinkNet、Segnet、PSPNet_MobileNetV2 和 PSPNet+ResNet50 九种模型的分割图像。从图 7 可以发现,本文所提出的模型对不同大小结节的整体分割效果明显优于其他模型。

为了更好地体现本文模型总体性能的稳定性,给出验证集图像在验证过程中的 mIoU 和 Dice 相似

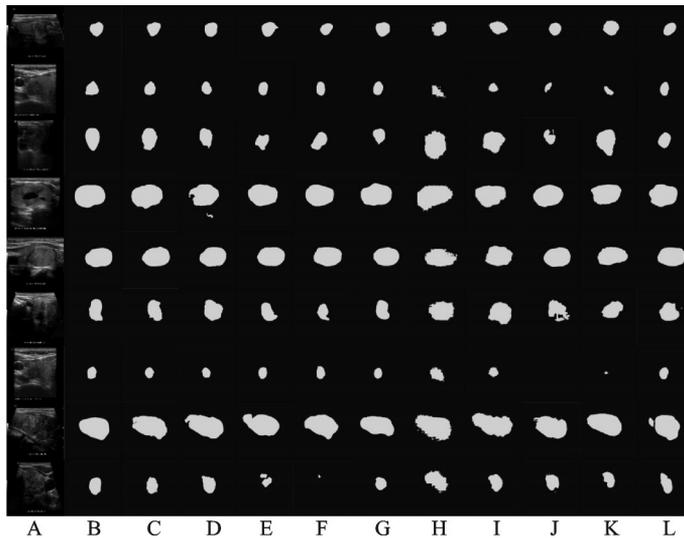


图7 不同模型分割结果

Fig.7 Segmentation results of different models

系数曲线,不同分割模型的 mIoU 曲线如图 8 所示,Dice 相似系数曲线如图 9 所示。由图 8、9 可见,在第 11 次验证中,本文方法的 mIoU 和 Dice 相似系数取得了最高值,且整体的 mIoU 和 Dice 相似系数显著高于其他模型。

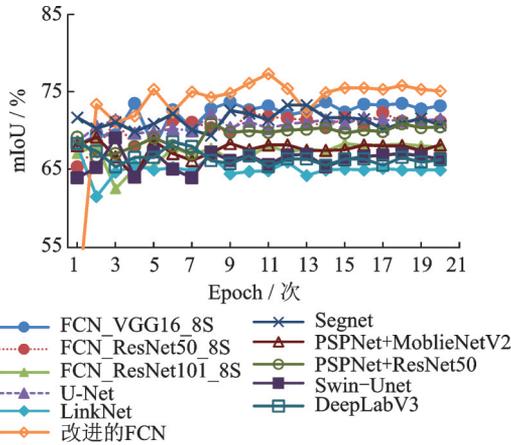


图 8 不同分割模型在验证集图像上的 mIoU 曲线

Fig.8 mIoU curves of different segmentation models on validation set images

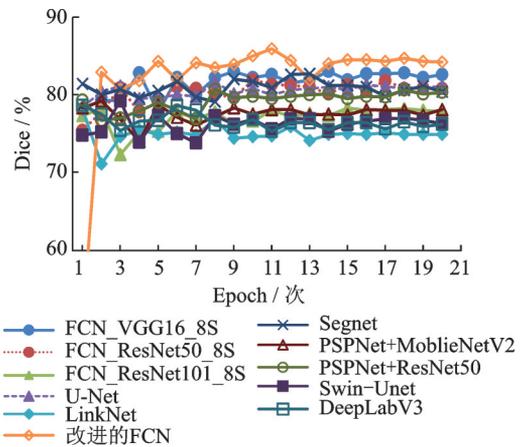


图 9 不同分割模型在验证集图像上的 Dice 曲线

Fig.9 Dice curves of different segmentation models on validation set images

为进一步验证本文所提模型对超声甲状腺结节的分割性能,将本文模型与文献[18,20-23]所提出的甲状腺结节分割方法进行指标对比,表 3 为在甲状腺结节测试集图像上的分割性能指标对比结果。由表 3 可见,本文模型的 mIoU、Precision、Recall 和 F_1 分数分别为 79.7%、99.82%、97.06% 和 98.42%。相比于文献[20],改进的 FCN 在 mIoU、Precision 和 F_1 分数指标上分别提升了 1.8%、1.35% 和 0.12%,在 Recall 指标上降低了 1.09%。相比于文献[21],在 Recall 指标相差 1.12% 的情况下,改进的 FCN 在 mIoU、Precision 和 F_1 分数指标上分别提升了 6.0%、2.85% 和 0.85%。相比于文献[22],改进的 FCN 在 mIoU、Precision、Recall 和 F_1 分数指标上分别提升了 4.4%、0.8%、0.04 和 0.42%。综合实验结果,对比文献[18,20-23],本文模型在 mIoU、Precision 和 F_1 分数指标上取得了最优值,实现了更高精度的超声甲状腺结节分割性能。

2.4 消融实验

为研究主干网络对分割网络性能的影响,比较使用多种主干特征提取网络进行实验对比,其余网络结构保持不变。表 4 为不同主干网络在甲状腺结节测试集图像上的分割性能指标对比。由表 4 可见,超声甲状腺结节测试集图像采用 VGG16 作为主干网络进行特征提取,分割精度更高。其中,mIoU、Dice 相似系数、 F_1 分数、FLOPs 和 Params 分别达到 79.7%、87.6%、98.42%、17.84 GMAC 和 32.17 MB。相比于采用 ResNet50,采用 VGG16 作为主干网络, mIoU、Dice 和 F_1 指标分别提升了 5.3%、4.0% 和 0.43%, Params 指标降低了 45.32 MB。综合实验结果,相比于采用其他特征提取网络,采用 VGG16 作为主干网

表 3 与现有甲状腺结节分割方法性能指标对比

Table 3 Comparison of segmentation performance indicators of existing thyroid nodule segmentation methods

方法	mIoU	Precision	Recall	F_1	%
文献[18]	65.8	97.78	94.63	96.18	
文献[20]	77.9	98.47	98.15	98.30	
文献[21]	73.7	96.97	98.18	97.57	
文献[22]	75.3	99.02	97.02	98.00	
文献[23]	73.8	98.25	97.05	97.64	
改进的 FCN	79.7	99.82	97.06	98.42	

表4 不同主干网络分割性能指标对比

Table 4 Comparison of segmentation performance indicators of different backbone

主干网络	mIoU/%	Dice/%	F1/%	FLOPs/GMAC	Params/MB
GoogLeNet	76.4	85.2	97.98	4.32	42.49
MobileNetV2	69.2	79.0	97.38	2.67	35.34
ResNet50	74.4	83.6	97.99	9.25	77.49
ResNet101	68.0	78.1	96.85	12.98	96.49
EfficientNetB0	69.2	78.9	97.51	2.72	36.98
EfficientNetB7	70.9	80.6	97.60	9.01	126.01
改进的FCN	79.7	87.6	98.42	17.84	32.17

络, mIoU、Dice 和 F1 指标均取得了最优值, 使用 VGG16 进行特征提取进一步提升了甲状腺结节的分割性能。

使用不同主干网络在甲状腺结节测试集图像上的分割结果如图 10 所示。图 10A 列为 9 张甲状腺结节图像, 分辨率为 224 像素 \times 224 像素。图 10B 列为放射科医生注释的分割图像, 图 10C 列为本文模型分割图像。图 10 D~I 列依次为 B-GoogLeNet、B-MobileNetV2、B-ResNet50、B-ResNet101、B-EfficientNetB0 和 B-EfficientNetB7 的分割图像, 即采用 GoogLeNet、MobileNetV2、ResNet50、ResNet101、EfficientNetB0 和 EfficientNetB7 作为主干网络的分割结果, 其中 B 代表 Based。由图 10 可见, 采用 VGG16 作为主干网络进行分割的整体分割效果明显优于采用其他主干网络。

为了更好地体现采用 VGG16 作为主干网络进行分割的稳定性, 给出甲状腺结节验证集图像在验证过程中的 mIoU 和 Dice 相似系数曲线, 采用不同主干进行分割的平均交并比曲线如图 11 所示,

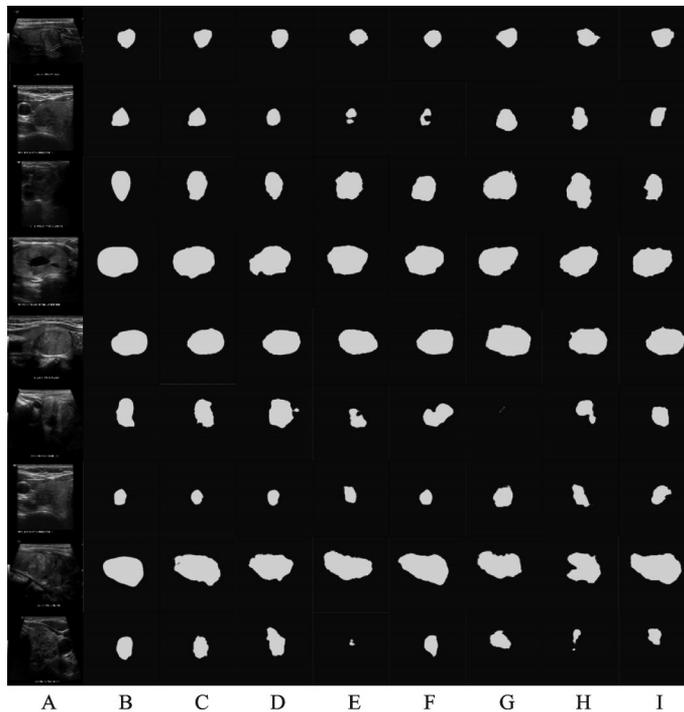


图 10 不同主干网络分割结果

Fig.10 Segmentation results of different backbone networks

Dice 相似系数曲线如图 12 所示。由图 11、12 可见,在第 11 次验证中,采用 VGG16 作为主干网络进行分割的 mIoU 和 Dice 相似系数取得了最优值,且整体的 mIoU 和 Dice 相似系数显著高于采用其他主干网络。

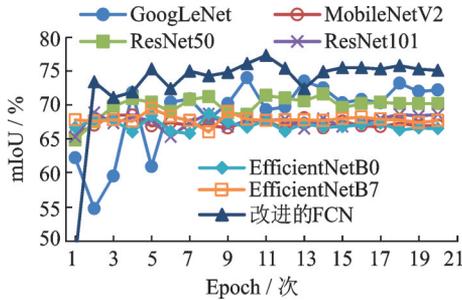


图 11 不同主干网络在验证集图像上的 mIoU 曲线

Fig.11 mIoU curves of different backbone on validation set images

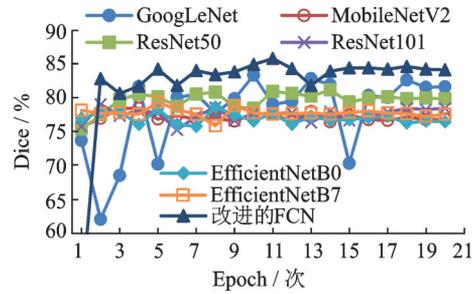


图 12 不同主干网络在验证集图像上的 Dice 曲线

Fig.12 Dice curves of different backbone on validation set images

为了进一步验证本文模型的性能,将超声甲状腺数据集分别应用于 VGG+Decoder 基础模型、VGG+Decoder 基础模型+FT 和 VGG+Decoder 基础模型+ASPP 模块,基础模型同时加入特征传递模块,ASPP 模块模型进行训练,并使用测试集进行测试。表 5 为 4 种模型在甲状腺结节测试集上的分割性能指标对比。由表 5 可见,相比于 VGG+Decoder 基础模型,加入特征传递模块后,在 mIoU, Dice 和 F_1 指标上分别提高了 1.32%、1.1%、0.16%;加入 ASPP 模块后,在 mIoU, Dice 和 F_1 指标上分别提高了 1%、0.8%、0.25%;同时加入特征传递模块和 ASPP 模块后,在 mIoU, Dice 和 F_1 指标上分别提高了 3.8%、2.8%、0.56%。由此可见,同时加入多层特征传递和 ASPP 模块后,模型分割性能得到了显著的提升。为研究输入图像分辨率对分割模型性能的影响,将超声甲状腺结节数据集图像重新调整为 3 种分辨率进行训练,分别为 224 像素 \times 224 像素,384 像素 \times 384 像素,448 像素 \times 448 像素,并使用测试集图像进行测试。表 6 为 3 种分辨率甲状腺结节图像的分割性能指标对比。由表 6 可知,随着输入甲状腺结节图像分辨率的增大分割精度降低,因此本文模型通过增大图像分辨率不能提高分割精度。

表 5 4 种模型分割性能指标对比

Table 5 Comparison of segmentation performance indicators of four models

分割模型	mIoU/%	Dice/%	F_1 /%	FLOPs/GMAC	Params/MB
VGG+Decoder	75.9	84.8	97.86	16.06	15.19
VGG+Decoder+FT	77.22	85.9	98.02	17.11	17.52
VGG+Decoder+ASPP	76.9	85.6	98.11	16.79	29.84
改进的 FCN	79.7	87.6	98.42	17.84	32.17

表 6 3 种分辨率甲状腺图像分割性能对比

Table 6 Comparison of segmentation performance of thyroid images with three resolutions

分辨率/(像素 \times 像素)	mIoU/%	Dice/%	F_1 /%	FLOPs/GMAC	Params/MB
224 \times 224	79.7	87.6	98.42	17.84	32.17
384 \times 384	78.7	86.9	98.32	52.42	32.17
448 \times 448	77.9	86.3	98.23	71.34	32.17

本文采用 BCE Loss 与 Dice Loss 的共同监督预测分割图像,如式(3)所示。为了进一步提升本文模型的分割精度,分别对两个损失函数 α 和 β 采用不同的权重进行训练,并使用测试集图像进行测试。表 7 为不同权重的分割性能指标对比。由表 7 可知,参数 α 设置为 0.2、参数 β 设置为 0.8 时,改进的 FCN 在 mIoU、Dice 和 F_1 指标上取得了最优值。相比于 α 设置为 1.0,参数 β 设置为 0.0 时,改进的 FCN 在 mIoU、Dice 和 F_1 指标上分别提升了 2.3%、1.6% 和 0.35%。相比于 α 设置为 0.0,参数 β 设置为 1.0 时,改进的 FCN 在 mIoU、Dice 和 F_1 指标上分别提升了 1.0%、0.6% 和 0.21%。综合实验结果,当超参数 α 为 0.2、 β 为 0.8 时改进的 FCN 模型可以实现更好的整体分割性能。

3 结束语

本文提出了一种改进 FCN 的超声甲状腺结节分割方法,主要加入多层特征传递模块和空洞空间卷积池化金字塔模块。其中,多层特征传递模块对来自 VGG16 的特征图进一步下采样,并传递多层次特征图辅助特征融合,空洞空间卷积池化金字塔模块提取多尺度的感受野信息。另外,采用 Decoder 模块完成特征图上采样,融合多层次特征信息。本文实验是在斯坦福大学 AIMI 共享数据集上进行,利用 mIoU、Dice 相似系数、 F_1 分数、FLOPs 和 Params 五个指标进行分析。实验结果表明,相比于其他模型,本文模型在 mIoU、Dice 和 F_1 三个指标上都取得了最优值。但是所提模型较复杂,训练参数稍多,因此,后续本文将重点构建一个可部署到医疗设备中的轻量级高精度的图像分割网络。

参考文献:

- [1] PALUSKIEVICZ C M, CHANG D R, BLACKBURN K W, et al. Low-risk papillary thyroid cancer: Treatment de-escalation and cost implications[J]. Journal of Surgical Research, 2022, 275: 273-280.
- [2] CHEN Bo, FENG Mei, YAO Zhongyang, et al. Hypoxia promotes thyroid cancer progression through HIF1 α /FGF11 feedback loop[J]. Experimental Cell Research, 2022, 416(1): 113159.
- [3] 董芬,张彪,单广良. 中国甲状腺癌的流行现状和影响因素[J]. 中国癌症杂志, 2016, 26(1): 47-52.
DONG Fen, ZHANG Biao, SHAN Guangliang. Distribution and risk factors of thyroid cancer in China[J]. China Oncology, 2016, 26(1): 47-52.
- [4] VAN VELSEN E F S, LEUNG A M, KOREVAAR T I M. Diagnostic and treatment considerations for thyroid cancer in women of reproductive age and the perinatal period[J]. Endocrinology and Metabolism Clinics, 2022, 51(2): 403-416.
- [5] 王波,李梦翔,刘侠. 基于改进 U-Net 网络的甲状腺结节超声图像分割方法[J]. 电子与信息学报, 2022, 44(2): 514-522.
WANG Bo, LI Mengxiang, LIU Xia. Ultrasound image segmentation method of thyroid nodules based on the improved U-Net network[J]. Journal of Electronics & Information Technology, 2022, 44(2): 514-522.
- [6] PHUTTHARAK W, BOONROD A, KLUNGBOONKRONG V, et al. Interrater reliability of various thyroid imaging reporting and data system (TIRADS) classifications for differentiating benign from malignant thyroid nodules[J]. Asian Pacific Journal of Cancer Prevention: APJCP, 2019, 20(4): 1283.
- [7] 胡屹杉,秦品乐,曾建潮,等. 基于特征融合和动态多尺度空洞卷积的超声甲状腺分割网络[J]. 计算机应用, 2021, 41(3): 891-897.
HU Yishan, QIN Pinle, ZENG Jianchao, et al. Ultrasound thyroid segmentation network based on feature fusion and dynamic

表 7 不同权重的分割性能指标对比

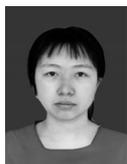
Table 7 Comparison of segmentation performance indicators with different weights

α	β	mIoU/%	Dice/%	F_1 /%
1.0	0.0	77.4	86.0	98.07
0.8	0.2	79.4	87.4	98.40
0.7	0.3	78.7	88.40	98.28
0.6	0.4	79.1	87.2	98.36
0.5	0.5	77.8	87.20	98.22
0.4	0.6	78.5	86.8	98.29
0.3	0.7	77.5	86.1	98.02
0.2	0.8	79.7	87.6	98.42
0.0	1.0	78.7	87.0	98.21

- multi-scale dilated convolution[J]. *Journal of Computer Applications*, 2021, 41(3): 891-897.
- [8] MAROULIS D E, SAVELONAS M A, IAKOVIDIS D K, et al. Variable background active contour model for computer-aided delineation of nodules in thyroid ultrasound images[J]. *IEEE Transactions on Information Technology in Biomedicine*, 2007, 11(5): 537-543.
- [9] CHAN T F, VESE L A. Active contours without edges[J]. *IEEE Transactions on Image Processing*, 2001, 10(2): 266-277.
- [10] SAVELONAS M A, IAKOVIDIS D K, LEGAKIS I, et al. Active contours guided by echogenicity and texture for delineation of thyroid nodules in ultrasound images[J]. *IEEE Transactions on Information Technology in Biomedicine*, 2008, 13(4): 519-527.
- [11] ZHAO Jie, ZHENG Wei, ZHANG Li, et al. Segmentation of ultrasound images of thyroid nodule for assisting fine needle aspiration cytology[J]. *Health Information Science and Systems*, 2013, 1(1): 1-12.
- [12] ALRUBAIDI W M H, PENG Bo, YANG Yan, et al. An interactive segmentation algorithm for thyroid nodules in ultrasound images[C]//*Proceedings of International Conference on Intelligent Computing*. Cham: Springer, 2016: 107-115.
- [13] WANG Lei, YANG Shujian, YANG Shan, et al. Automatic thyroid nodule recognition and diagnosis in ultrasound imaging with the YOLOv2 neural network[J]. *World Journal of Surgical Oncology*, 2019, 17(1): 12.
- [14] BUDA M, WILDMAN-TOBRINER B, CASTOR K, et al. Deep learning-based segmentation of nodules in thyroid ultrasound: Improving performance by utilizing markers present in the images[J]. *Ultrasound in Medicine & Biology*, 2020, 46(2): 415-421.
- [15] 卢宏涛, 罗沐昆. 基于深度学习的计算机视觉研究新进展[J]. *数据采集与处理*, 2022, 37(2): 247-278.
LU Hongtao, LUO Mukun. Survey on new progresses of deep learning based computer vision[J]. *Journal of Data Acquisition and Processing*, 2022, 37(2): 247-278.
- [16] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.]: IEEE, 2018: 7132-7141.
- [17] 刘安安, 李天宝, 王晓雯, 等. 基于深度学习的三维模型检索算法综述[J]. *数据采集与处理*, 2021, 36(1): 1-21.
LIU Anan, LI Tianbao, WANG Xiaowen, et al. Review of 3D model retrieval algorithms based on deep learning[J]. *Journal of Data Acquisition and Processing*, 2021, 36(1): 1-21.
- [18] MA Jinlian, WU Fa, JIANG Tian'an, et al. Ultrasound image-based thyroid nodule automatic segmentation using convolutional neural networks[J]. *International Journal of Computer Assisted Radiology and Surgery*, 2017, 12(11): 1895-1910.
- [19] 刘明坤, 张俊华, 李宗桂. 改进Mask R-CNN的甲状腺结节超声图像分割方法[J]. *计算机工程与应用*, 2022, 58(16): 219-225.
LIU Mingkun, ZHANG Junhua, LI Zongui. Improved mask R-CNN method for thyroid nodules segmentation in ultrasound images[J]. *Computer Engineering and Applications*, 2022, 58(16): 219-225.
- [20] LI Xuewei, WANG Shuaijie, WEI Xi, et al. Fully convolutional networks for ultrasound image segmentation of thyroid nodules[C]//*Proceedings of 2018 IEEE 20th International Conference on High Performance Computing and Communications*. [S.l.]: IEEE, 2018: 886-890.
- [21] ZHOU Shujun, WU Hong, GONG Jie, et al. Mark-guided segmentation of ultrasonic thyroid nodules using deep learning[C]//*Proceedings of the 2nd International Symposium on Image Computing and Digital Medicine*. New York: ACM, 2018: 21-26.
- [22] YING Xiang, YU Zhihui, YU Ruiguo, et al. Thyroid nodule segmentation in ultrasound images based on cascaded convolutional neural network[C]//*Proceedings of International Conference on Neural Information Processing*. Cham: Springer, 2018: 373-384.
- [23] 迟剑宁, 于晓升, 张艺菲. 融合深度网络和浅层纹理特征的甲状腺结节癌变超声图像诊断[J]. *中国图象图形学报*, 2018, 23(10): 1582-1593.
CHI Jianning, YU Xiaosheng, ZHANG Yifei. Thyroid nodule malignantrisk detection in ultrasound image by fusing deep and texture features[J]. *Journal of Image and Graphics*, 2018, 23(10): 1582-1593.
- [24] 裴昀. 医学影像分析中的注意力机制研究[D]. 长春: 吉林大学, 2022.
PEI Yun. Research on attention mechanism in medical image analysis[D]. Changchun: Jilin University, 2022.
- [25] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-10) [2022-06-13]. <https://doi.org/10.48550/arXiv.1409.1556>.

- [26] BA J L, KIROS J R, HINTON G E. Layer normalization[EB/OL]. (2016-07-21)[2022-06-13]. <https://doi.org/10.48550/arXiv.1607.06450>.
- [27] MAAS A L, HANNUN A Y, NG A Y. Rectifier nonlinearities improve neural network acoustic models[J]. Journal of Machine Learning Research, 2013, 30(1): 3.
- [28] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFS[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4): 834-848.
- [29] 周苏, 吴迪, 金杰. 基于卷积神经网络的车道线实例分割算法[J]. 激光与光电子学进展, 2021, 58(8): 381-388.
ZHOU Su, WU Di, JIN Jie. Lane instance segmentation algorithm based on convolutional neural network[J]. Laser & Optoelectronics Progress, 2021, 58(8): 381-388.
- [30] CHAURASIA A, CULURCIELLO E. Linknet: Exploiting encoder representations for efficient semantic segmentation[C]// Proceedings of 2017 IEEE Visual Communications and Image Processing (VCIP). [S.l.]: IEEE, 2017: 1-4.
- [31] JADON S. A survey of loss functions for semantic segmentation[C]// Proceedings of 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology. [S.l.]: IEEE, 2020: 1-7.
- [32] LI Xiaoya, SUN Xiaofei, MENG Yuxian, et al. Dice loss for data-imbalanced NLP tasks[EB/OL]. (2019-09-07)[2022-06-14]. <https://doi.org/10.48550/arXiv.1911.02855>.
- [33] XIONG Ruibin, YANG Yunchang, HE Di, et al. On layer normalization in the transformer architecture[C]// Proceedings of International Conference on Machine Learning. [S.l.]: PMLR, 2020: 10524-10533.
- [34] LOSHCHILOV I, HUTTER F. Decoupled weight decay regularization[EB/OL]. (2017-09-14)[2022-06-13]. <https://doi.org/10.48550/arXiv.1711.05101>.
- [35] YANG Lei, GU Yuge, HUO Benyan, et al. A shape-guided deep residual network for automated CT lung segmentation[J]. Knowledge-Based Systems, 2022, 250: 108981.
- [36] PERAZZI F, PONT-TUSET J, MCWILLIAMS B, et al. A benchmark dataset and evaluation methodology for video object segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2016: 724-732.
- [37] MOLCHANOV P, TYREE S, KARRAS T, et al. Pruning convolutional neural networks for resource efficient inference[EB/OL]. (2016-09-19)[2022-06-13]. <https://doi.org/10.48550/arXiv.1611.06440>.

作者简介:



张雅婷(1997-),女,硕士研究生,研究方向:深度学习医学图像处理等,E-mail: 1599830541@qq.com。



帅仁俊(1962-),通信作者,男,研究员,研究方向:深度学习、智能建筑等,E-mail: srjwhy@njtech.edu.cn。



黄道宏(1997-),男,硕士研究生,研究方向:深度学习医学图像处理。



赵宸(1995-),男,硕士研究生,研究方向:深度学习医学图像处理。



吴梦麟(1982-),男,副教授,博士,研究方向:医学图像处理、医疗信息的检索和挖掘。

(编辑:刘彦东)