

基于改进U-Net网络的甲状腺结节超声图像分割方法

王波^{①②} 李梦翔^① 刘侠^{*①②}

^①(哈尔滨理工大学自动化学院 哈尔滨 150080)

^②(黑龙江省复杂智能系统与集成重点实验室 哈尔滨 150080)

摘要: 针对甲状腺结节尺寸多变、超声图像中甲状腺结节边缘模糊导致难以分割的问题, 该文提出一种基于改进U-net网络的甲状腺结节超声图像分割方法。该方法首先将图片经过有残差结构和多尺度卷积结构的编码器路径进行降尺度特征提取; 然后, 利用带有注意力模块的跳跃长连接部分对特征张量进行边缘轮廓保持操作; 最后, 使用带有残差结构和多尺度卷积结构的解码器路径得到分割结果。实验结果表明, 该文所提方法的平均分割Dice值达到0.7822, 较传统U-Net方法具有更优的分割性能。

关键词: 图像分割; 甲状腺结节超声图像; 注意力机制; U-Net

中图分类号: TN911.73; TP391

文献标识码: A

文章编号: 1009-5896(2022)02-0514-09

DOI: 10.11999/JEIT210015

Ultrasound Image Segmentation Method of Thyroid Nodules Based on the Improved U-Net Network

WANG Bo^{①②} LI Mengxiang^① LIU Xia^{①②}

^①(School of Automation, Harbin University of Science and Technology, Harbin 150080, China)

^②(Heilongjiang Provincial Key Laboratory of Complex Intelligent System and Integration, Harbin 150080, China)

Abstract: An ultrasound image segmentation method of thyroid nodules based on the improved u-net network is proposed in this paper, in order to solve the problem of changeable size of thyroid nodules and difficulty in segmentation due to edge blur of thyroid nodules in the ultrasound image. Firstly, the image is downscaled to extract the features through an encoder path with a residual structure and a multi-scale convolution structure. Secondly, the long skip connection with an attention module is used to maintain the edge contour of characteristic tensor. Finally, the segmentation result is obtained by a decoder path with a residual structure and a multi-scale convolution structure. The experimental results show that with the method proposed in this paper, the average segmentation Dice value reaches 0.7822. It indicates that this method has better segmentation performance than the traditional U-Net method.

Key words: Image segmentation; Ultrasound image of thyroid nodule; Attention mechanism; U-Net

1 引言

甲状腺结节是常见的临床问题, 甲状腺癌发病率近年来呈上升趋势^[1]。超声检查因简单、方便易行、经济以及检查时间短等特点, 已成为临床上

首选的甲状腺结节检测手段。甲状腺结节的大小、形状、轮廓等特征是临床上甲状腺结节良恶性诊断中的重要依据。因此, 实现全自动、高精度的甲状腺结节超声图像分割具有较高的临床意义。

目前, 常见的甲状腺结节分割方法主要可分为基于活动轮廓模型的方法、基于区域的方法和基于深度学习的方法等3种类型。基于活动轮廓模型的方法^[2-4]利用超声图像中的轮廓信息对甲状腺结节进行分割, 这类方法需要预先设定初始轮廓, 而甲状腺结节, 尤其是恶性结节通常存在浸润性不规则边缘, 结节组织区域与周围环境难以区分, 分割效果并不理想。基于区域的方法^[5,6]利用区域间的强度统计特性(比如均值等)得到边界能量函数, 通过最小化边界能量函数生成结节的边缘轮廓, 这类方

收稿日期: 2021-01-05; 改回日期: 2021-03-31; 网络出版: 2021-04-16

*通信作者: 刘侠 liuxia@hrbust.edu.cn

基金项目: 国家自然科学基金(61172167), 哈尔滨理工大学“理工英才”计划科学研究项目(LGYC2018JC013), 黑龙江省青年科学基金项目(QC2017076)

Foundation Items: The National Natural Science Foundation of China (61172167), The Scientific Research Project of Talent Plan of Harbin University of Science and Technology (LGYC2018JC013), The Youth Science Foundation of Heilongjiang Province (QC2017076)

法适用于相同组织区域的灰度是均匀分布,而不同组织区域间的灰度分布是不均匀的图像。然而,在大多数甲状腺结节超声图像中,不同组织区域间的灰度差异并不明显,需要先验形状信息和位置信息来实现区域的准确分割。近年来,以U-Net^[7]为代表的深度学习模型在医学图像分割任务中取得了很好的效果,已成为该领域最大的研究热点。相较于前两种方法来说,基于深度学习的方法通过对图像中的像素或图像块进行分类,利用训练模型完成甲状腺结节图像的分割,无需人工干预。然而,现有的模型大都会等权重地分配特征张量上所有空间位置和通道上的信息而产生大量的计算冗余,导致模型训练速度变慢,分割精度变低,注意力机制的引入在一定程度上提升了模型的效率和性能。另外,这类方法除了模型训练非常耗时之外,还需要大量的训练样本和标签,才能获得理想的分割效果。

鉴于目前研究中存在的问题,本文提出了一种新的改进型U-Net网络模型用于甲状腺结节超声图像的自动分割。该模型以U-Net为主干网络,设计了一种结合残差结构和多尺度卷积模块用以提升对不同尺寸结节的分割精度。此外,为了保留甲状腺结节的边缘信息,在跳跃长连接过程中引入空间注意力和通道注意力机制。通过对比实验和注意力模块的可视化分析,证明本文所提的方法较传统U-Net具有更优的分割性能。

2 相关工作

针对甲状腺结节分割问题,部分研究者提出了基于活动轮廓模型的分割方法。文献[2]在无边主动轮廓(Active Contour Without Edges, ACWE)模型^[8]的基础上,提出了可变背景主动轮廓(Variable Background Active Contour, VBAC)模型,该模型在背景分布不均匀的甲状腺结节超声图像中分割效果比ACWE有所提升,但对非低回声结节的分割效果欠佳。文献[3]提出了联合回声反射性-纹理(Joint Echogenicity-Texture, JET)模型,该模型在VBAC模型的基础上加入了区域像素强度和纹理特征分布信息,对等回声的甲状腺结节超声图像分割效果较VBAC模型有所提升,但该模型并不能很好地区分大型血管和甲状腺结节结构。文献[4]提出了一种结合ACWE模型和区域扩张拟合能量(Region-Scalable Fitting energy, RSF)模型^[9]的甲状腺结节超声图像分割方法,获得了不错的分割结果,但需预先设置初始轮廓。这类方法虽然实现过程相对较为简单,但需要预先设定初始轮廓,且在每次分割中都需要长时间的迭代,效率较低。此外,由

于甲状腺结节之间的个体差异较大,这类方法对不同的病例图像分割结果存在较大的不确定性。

与此同时,研究者们也提出了基于区域的方法用于甲状腺结节超声图像的分割。文献[5]提出了一种基于归一化模型的甲状腺结节超声图像分割方法。该方法在归一化模型中加入同态滤波和各向异性扩散操作,在减少图像噪声的同时保留了重要的边缘细节,但缺乏通用性和普适性。文献[6]提出了一种基于径向梯度和方差缩减统计(Variance-Reduction Statistics, VRS)算法的超声甲状腺结节分割方法。该方法首先需要放射科医生手动标注结节的长短轴和中心点,然后使用VRS算法粗略估计结节径向线与结节边缘的交点,随后通过选择最近的相邻点进行连接来确定结节的形状,最后采用B样条方法提高了分割精度。但是该方法需要专业的放射科医生手动干预,效率较低。综合已有研究,基于区域的方法也需要丰富的先验信息才能获得较为精确的分割效果。

相较于上面的传统图像分割算法,深度学习算法在分割的准确率和算法的自动化程度上都有较大提升。文献[10]提出了一种加入了残差结构和注意力门机制的U-Net网络,该网络的分割效果较U-Net网络有所提升,但在对比度较低的甲状腺结节超声图像上效果较差,此外,当一张图像中出现多个结节时,该网络不能有效地分割出所有结节区域。文献[11]提出基于注意力机制的半监督神经网络对甲状腺结节超声图像进行分割,该网络通过使用弱注释的分类数据和少量完全注释的分割数据来完成甲状腺超声图像的分割,并取得了较好的分割结果,但该模型的泛化能力还有待提升。文献[12]提出了一种使用空间金字塔池化结构的模型,该模型将空间金字塔池化结构与编解码器路径相结合,使其能够更好地捕获图像中的上下文信息,虽然最后取得了较好的分割结果,但是该模型需要的训练时间较长。文献[13]在Mask-RCNN网络的基础上设计了一种具有多任务处理能力的网络框架,能够同时进行甲状腺结节检测、分割和分类的任务,但该模型对于小尺寸的结节分割效果较差。与上面两种传统的分割方法相比,基于深度学习的方法无论是在算法的自动化程度上还是在算法对不同病例图像的适应性上都有很大程度的提升。

近年来,注意力机制被引入深度学习中的图像分割领域。文献[14]在U-Net中加入了注意力机制,使用深层卷积提取的权重图来监督浅层卷积,将激活的部分限制于待分割的区域并减小背景的激活值以此来达到优化分割效果的目的。文献[15]提

出了一种通道和空间相结合的注意力模块,该模块的参数数量较少并且能够嵌入在主流网络中来有效提升分类和检测工作的准确率。文献[16]提出了一种用于语义分割的双重注意力网络,分别使用了空间和通道的注意力模块对同一通道内的上下文信息以及不同通道间的依赖关系进行提取,在语义分割任务中取得了不错的结果。文献[17]提出了一种带边界保持模块的医学超声图像分割模型,该模块利用边界关键点生成的权重图增强网络对于目标边界区域的关注度,能够让分割结果的形状和轮廓保持在专家金标准的附近。文献[18]认为语义分割任务应该分为像素预测和像素分组两个子任务来完成,同时提出了一种用于学习多尺度空间特征以及非局部特征的压缩和注意力(Squeeze-and-Attention, SA)模块,并将该模块添加在主流的分割模型上使其分割结果得到了较大优化。总的来说,注意力机制在图像分割领域起到了辅助作用,能够帮助网络模型达到更好的分割效果。

3 本文方法

针对甲状腺结节尺寸多变且边缘模糊导致图像难以分割的问题,本文提出了一种基于残差多尺度卷积和注意力机制的深度卷积神经网络模型。图1给出了该模型的整体框架。该模型以U-Net为主干

网络结构,包含编码路径、注意力模块和解码路径3个部分。首先对图像进行残差多尺度卷积和编码下采样,同时将每一层级输出的特征张量经过注意力模块复制到对应层级的解码器中,然后对下采样得到特征张量进行叠加和上采样,最终通过Dice损失函数实现甲状腺结节的掩膜分割。

3.1 残差多尺度卷积模块

在图1编解码器的每个层级中,特征张量都依次经过一个残差块和一个多尺度卷积块。图2给出了残差块结构图,其中 H, W, C 分别为特征张量的高、宽和通道的维度。该残差模块由两个卷积层组成,每个卷积层依次包含一个 3×3 卷积,一个批量归一化(Batch Normalization, BN)处理和一个修正线性单元(Rectified Linear Unit, ReLU)。输入特征张量经过两个卷积层后与自身相加,再经过一个ReLU激活函数,最终得到输出特征张量。

另外,本文分别设计了两种具有不同卷积核的多尺度卷积结构。图3给出了这两种卷积结构的框架图。其中,图3(a)为多尺度小核卷积,用于编码路径和解码路径中的上两层卷积操作,以获得更完整的局部几何细节信息;图3(b)为多尺度大核卷积,用于编解码中的下3层卷积,以生成更完整的高层语义信息。

对于多尺度小核卷积结构,如图3(a)所示,输

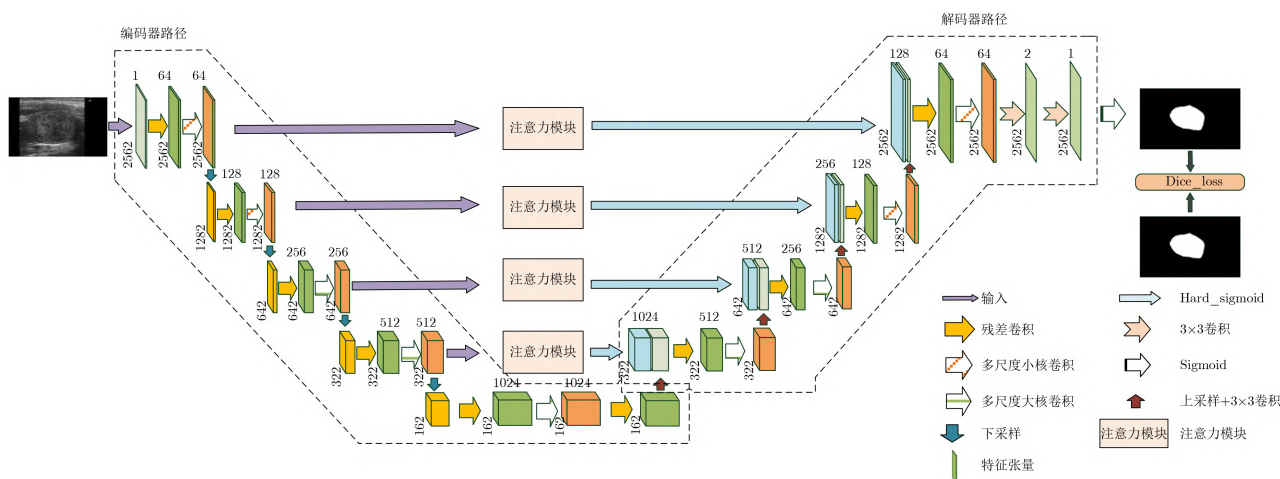


图1 基于残差多尺度卷积和注意力机制的深度卷积神经网络模型

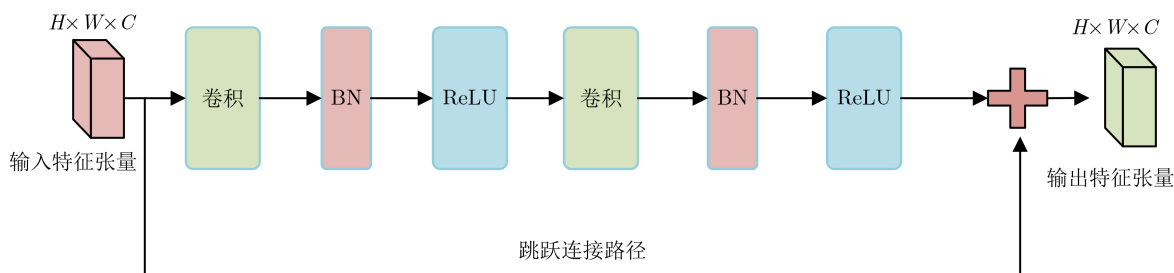


图2 本文中所使用的残差块

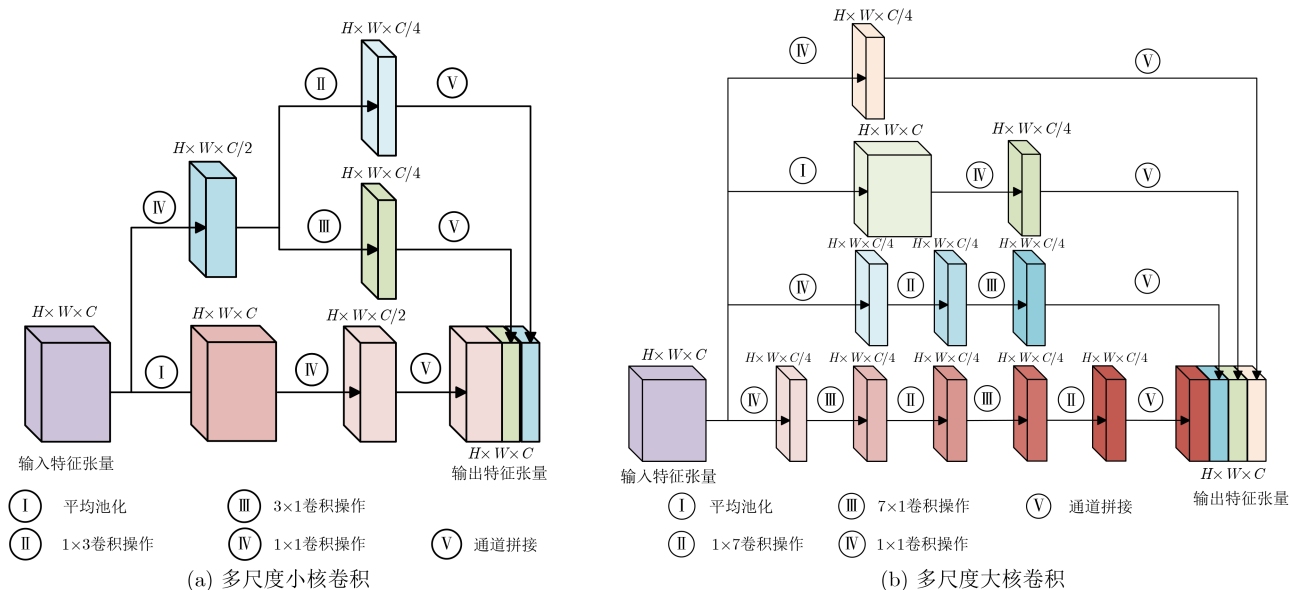


图3 多尺度卷积结构

入特征张量经过两个分支，在上方分支中输入特征张量先经过 1×1 的卷积，然后分别通过一个 1×3 卷积分支和另一个 3×1 卷积分支得到两个输出特征张量；在下方的分支中输入特征张量依次经过平均池化和 1×1 卷积得到输出特征张量。其中，每个卷积层后都依次包含一个BN处理和一个ReLU。上下两个分支的输出特征张量进行通道上的拼接得到多尺度小核卷积的最终输出特征张量。

对于多尺度大核卷积结构，如图3(b)所示，采用inception-v4结构中的inception-B结构^[19]。其中，输入的特征张量将经过4个分支。在最上方的分支中输入特征张量经过 1×1 的卷积后得到输出特征张量。在上方第2个分支中输入特征张量依次经过平均池化以及 1×1 的卷积后得到输出特征张量。在上方第3个分支中输入特征张量依次经过 1×1 卷积、 1×7 卷积和 7×1 卷积后得到输出特征张量。在最下方分支中输入特征张量依次经过 1×1 卷积、 7×1 卷积、 1×7 卷积、 7×1 卷积和 1×7 卷积后得到输出特征张量。同样的，4个路径中的每个卷积层后都依次包含一个BN处理和一个ReLU。4个分支的

输出特征张量进行通道维度上的拼接得到多尺度大核卷积的最终输出特征张量。

3.2 注意力模块

为了解决U-Net模型中的跳跃长连接在传递关键细节特征时将无用信息或噪声同时传入解码层的问题，本文在网络模型中引入了注意力机制。图4给出了该注意力模块的网络结构图，其中 H, W, C 分别为特征张量的高、宽和通道的维度。该注意力机制包含空间注意力模块和通道注意力模块，输入特征张量经过空间注意力模块后的输出特征张量作为通道注意力模块的输入特征张量。

3.2.1 空间注意力模块

在空间注意力模块中，输入特征张量首先通过通道上的全局平均池化和全局最大池化后进行通道维度拼接，然后经过一个 7×7 卷积，并利用Sigmoid激活函数得到一个通道数为1的权重图，接着将权重图和原输入特征张量相乘，最后输出一个与原张量同型的特征张量。令 $\mathbf{F}_s \in \mathbb{R}^{H \times W \times C}$ 为空间注意力模块的输出特征张量，其中 H, W, C 分别为输入特征张量的高、宽和通道的维度，则有

$$\mathbf{F}_s = \mathbf{F}_1 \otimes \mathbf{F} = \begin{pmatrix} F_1^{1,1} & F_1^{1,2} & \dots & F_1^{1,W} \\ F_1^{2,1} & F_1^{2,2} & \dots & F_1^{2,W} \\ \vdots & \vdots & \ddots & \vdots \\ F_1^{H,1} & F_1^{H,2} & \dots & F_1^{H,W} \end{pmatrix} \otimes \begin{pmatrix} \mathbf{F}^{1,1} & \mathbf{F}^{1,2} & \dots & \mathbf{F}^{1,W} \\ \mathbf{F}^{2,1} & \mathbf{F}^{2,2} & \dots & \mathbf{F}^{2,W} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{F}^{H,1} & \mathbf{F}^{H,2} & \dots & \mathbf{F}^{H,W} \end{pmatrix}$$

$$= \begin{pmatrix} F_1^{1,1} \cdot \mathbf{F}^{1,1} & F_1^{1,2} \cdot \mathbf{F}^{1,2} & \dots & F_1^{1,W} \cdot \mathbf{F}^{1,W} \\ F_1^{2,1} \cdot \mathbf{F}^{2,1} & F_1^{2,2} \cdot \mathbf{F}^{2,2} & \dots & F_1^{2,W} \cdot \mathbf{F}^{2,W} \\ \vdots & \vdots & \ddots & \vdots \\ F_1^{H,1} \cdot \mathbf{F}^{H,1} & F_1^{H,2} \cdot \mathbf{F}^{H,2} & \dots & F_1^{H,W} \cdot \mathbf{F}^{H,W} \end{pmatrix} \quad (1)$$

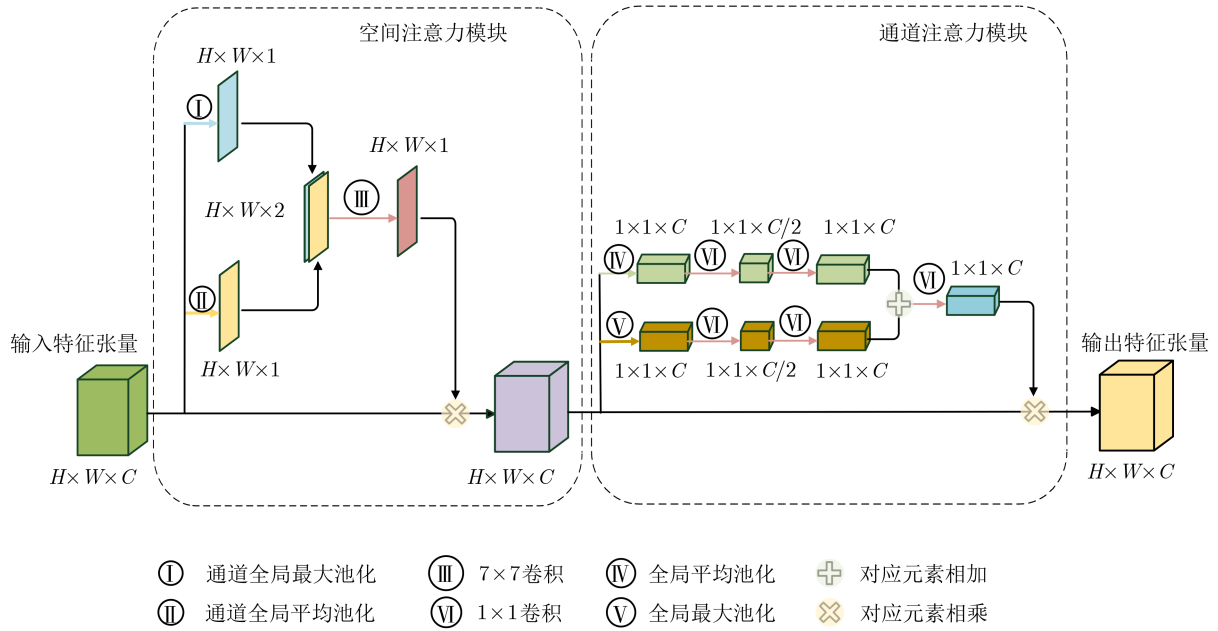


图4 注意力模块网络结构

其中, F_1 表示空间注意力模块的权重图, $F_1^{i,j}$ 代表权重图在位置 $P(i, j)$ 处的权重, $i = 1, 2, \dots, H$; $j = 1, 2, \dots, W$ 。 $F \in \mathbb{R}^{H \times W \times C}$ 表示空间注意力模块的输入, $F^{i,j} \in \mathbb{R}^{1 \times 1 \times C}$ 代表输入特征张量在位置 $P(i, j)$ 处的所有通道组成的向量, $i = 1, 2, \dots, H$; $j = 1, 2, \dots, W$ 。 \otimes 表示对应位置元素相乘。

3.2.2 通道注意力模块

在通道注意力模块中, 空间注意力模块的输出特征张量首先通过全局平均池化和全局最大池化后得到两个特征权重序列, 然后将这两个权重序列分别通过两次 1×1 卷积后相加, 并利用 Hard_sigmoid 激活函数得到一个权重序列, 接着将该权重序列与空间注意力模块的输出相乘, 最后输出一个与空间注意力模块输出特征张量同型的特征张量。经过通道注意力模块后得到的特征张量会被复制到解码路径的对应层级中。令 $F_c \in \mathbb{R}^{H \times W \times C}$ 为通道注意力模块的输出特征张量, 其中 H, W, C 分别为输出特征张量的高、宽和通道的维度, 则有

$$\begin{aligned} F_c &= F_2 \otimes F_s \\ &= [F_2^1 \ F_2^2 \ \dots \ F_2^C] \otimes [F_s^1 \ F_s^2 \ \dots \ F_s^C] \\ &= [F_2^1 \cdot F_s^1 \ F_2^2 \cdot F_s^2 \ \dots \ F_2^C \cdot F_s^C] \end{aligned} \quad (2)$$

其中, F_2 表示通道注意力模块得到的权重序列, F_2^i 表示第 i 个通道上的权重, $i = 1, 2, \dots, C$ 。 $F_s \in \mathbb{R}^{H \times W \times C}$ 表示空间注意力模块的输出特征张量, F_s^i 表示 F_s 的第 i 个通道, \otimes 表示对应位置元素相乘。

4 实验与分析

4.1 实验数据及参数设置

本文的实验数据由上海市胸科医院提供, 一共

有450例甲状腺结节超声图像的样本以及专家标注轮廓。原始超声图像大小为 500×300 , 去除隐私信息后, 将图片统一裁剪为 256×256 大小, 并按照 8:1:1 比例随机划分为训练集、验证集和测试集, 再利用旋转、尺度变换等操作对训练集数据进行10倍数据增强, 增强后训练集共有3960例数据。

实验中网络的具体参数为: 输入图像大小为 256×256 , Batch_size 大小为8, 初始学习率设置为0.001, 模型优化器为自适应矩估计(ADaptive Moment estimation, Adam)^[20], 使用Dice-loss作为损失函数, 最大迭代次数为150次。在训练过程中, 如果验证集上连续10轮损失值没有下降, 则学习率减半。实验的硬件环境为Intel (R) Xeon(R) CPU主频2.30 GHz, NVIDIA Tesla P-100显卡; 操作系统为Ubuntu 18.04, 编程语言为Python 3.7, 所有程序均在Tensorflow和Keras混合框架下实现。

4.2 评价指标

实验从分割结果准确性方面, 将本文所提方法同U-Net^[7], ADeepLabv3+^[21], BCDU-Net^[22], U-Net++^[23], AU-Net^[14]以及带有压缩和激励(Squeeze and Excitation, SE)模块^[24]的AU-Net进行比较。评价指标包括骰子相似系数(Dice Similarity Coefficient, DSC)、交并比(Intersection over Union, IoU)、豪斯多夫(Hausdorff)距离、过分割率(False Positive Rate, FPR)和欠分割率(False Negative Rate, FNR)。下面给出这些指标的具体定义。

给定两个区域 A 和 B , 则它们的骰子相似系数DSC和相似度IoU分别定义为

$$DSC = \frac{2|A \cap B|}{|A| + |B|} \quad (3)$$

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (4)$$

给定两个点集 A 和 B , 则它们之间的豪斯多夫距离 $H(A, B)$ 定义为

$$H(A, B) = \max[h(A, B), h(B, A)] \quad (5)$$

其中

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (6)$$

$$h(B, A) = \max_{b \in B} \min_{a \in A} \|b - a\| \quad (7)$$

FPR和FNR分别表示为

$$FPR = \frac{FP}{A \cup B} \quad (8)$$

$$FNR = \frac{FN}{A \cup B} \quad (9)$$

其中, FP和FN分别表示像素分类结果中假阳性和假阴性的数量。

4.3 分割结果与分析

图5给出了不同网络的分割结果, 其中图5(a)是原始数据, 图5(b)是专家标注的金标准, 图5(c)—图5(i)依次是U-Net, ADeeplabv3+, BCDU-Net, U-Net++, AU-Net, 带有SE模块的AU-Net和本文方法的分割结果, 图中线框标注了专家金标准、对比网络以及本文网络分割结果之间的部分差异。

从图5(c)、图5(d)、图5(g)、图5(h)可以看到, U-Net, ADeeplabv3+, AU-Net, 以及带有SE模块的AU-Net的分割结果中出现了许多过分割

现象, 会将一些距离专家标注轮廓较远处的区域划入预测结果中, 这种现象在结节尺寸较小处尤为明显。从图5(e)和图5(f)可以看到, BCDU-Net和U-Net++的分割结果也有类似的问题, 不同的是BCDU-Net对于小型结节的分割效果较好, 而U-Net++在结节尺寸较小处出现了一些欠分割现象。从图5(i)可以看出本文方法分割结果的边缘轮廓保持在专家标注的附近区域, 并且纵横比和形状都没有出现较大的误差, 而这两种特征也是后续医疗诊断工作中至关重要的特征。从结果来看, 本文提出的方法效果最好。

卷积神经网络在浅层会提取到一些边缘的细节特征, 随着网络深度的增加浅层边缘信息会有一定程度的丢失, 抽象的语义特征信息将会被提取。本文方法中将浅层的编码器提取到的特征通过注意力模块再连接到解码器的对应层级, 能够给那些浅层的特征中边缘细节信息以更高权重, 在保留边缘信息的同时也在一定程度上抑制了浅层特征中的无用信息或者噪声。此外, 甲状腺结节在尺寸上存在较大差异性, 而U-Net, BCDU-Net, U-Net++和AU-Net网络都仅使用了 3×3 卷积, 这种单一尺寸的卷积使得网络不能对不同尺寸的目标都获得合适的感受野。因此, 本文方法在编码器和解码器阶段采用了多尺度卷积模块, 将不同尺寸卷积的结果进行通道上的拼接, 使其表现在不同的通道上。而通道注意力模块在学习中选择合适尺寸的卷积层信息。

表1给出了本文模型与对比模型在测试集上的分割结果。

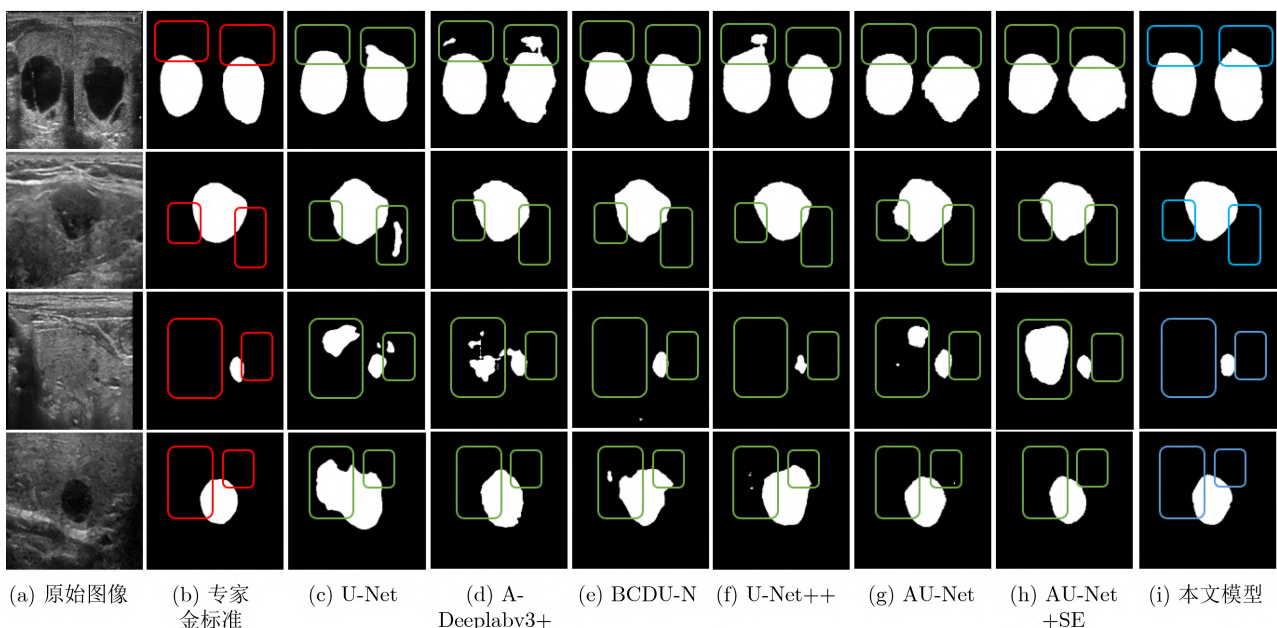


图5 不同网络分割结果对比

表 1 不同模型的量化分割结果

| 方法 | DSC | IoU | Hausdorff | FPR | FNR |
|--------------|----------------------|----------------------|-----------------------|----------------------|----------------------|
| U-Net | 0.6254±0.0719 | 0.4550±0.0709 | 35.1372±4.2380 | 0.0703±0.0216 | 0.1857±0.0532 |
| A-Deeplabv3+ | 0.6874±0.0241 | 0.5258±0.0280 | 26.7047±2.1334 | 0.0553±0.0123 | 0.2364±0.0244 |
| BCDU-Net | 0.7132±0.0086 | 0.5554±0.0104 | 22.8483±2.9490 | 0.0540±0.0245 | 0.2052±0.0416 |
| U-Net++ | 0.7177±0.0307 | 0.5597±0.0214 | 27.2110±2.3588 | 0.0407±0.0037 | 0.2556±0.0785 |
| AU-Net | 0.7413±0.0451 | 0.5889±0.0581 | 25.2377±1.1363 | 0.0426±0.0023 | 0.3024±0.0785 |
| AU-Net+SE | 0.7432±0.0231 | 0.5934±0.0289 | 24.3763±2.0682 | 0.0366±0.0022 | 0.2720±0.0301 |
| 本文方法 | 0.7822±0.0113 | 0.6423±0.0151 | 19.2769±1.2694 | 0.0306±0.0064 | 0.1991±0.0197 |

从表1可以看出,总的来说本文提出的方法效果最好,与改进前的U-Net网络相比,本文提出的模型得到的分割结果在DSC上提升了约16%。在IoU上提升了约19%。从豪斯多夫距离上来看,本文方法在分割结果的边缘轮廓精度上有较大提升,说明本文的注意力模块在边界保持方面起到了良好的作用。本文方法的FPR也有明显降低,最后的FNR表现不及U-Net,这是因为U-Net在分割结果上出现了较多的过分割现象,即分割结果的面积总是较大一些,因此不易出现欠分割现象。同样,与其余对比算法相比,本文的模型在各项指标上也均有较大领先。

4.4 注意力模块的可视化与分析

为了进一步说明注意力模块的作用,本文对两个注意力模块进行了可视化实验。图6给出了包含4个注意力模块的空间注意力权重图可视化结果。其中,图6(a)是原始图像,图6(b)是专家标注的金标准,图6(c)—图6(f)分别是图1中从上到下的4个注意力模块的权重图可视化后的结果。图中较亮的部分表示该地方权重接近于1,即该位置的空间信息对于分割结果有较大作用,较暗的部分表示该地方权重接近于0,即该位置的空间信息对于分割结果贡献较小。特别地,如果没有空间注意力模块,所有权重图应是全白的,表示所有位置的权重都为1即每个像素的空间位置信息在学习过程中都会被全部接收。

从图6(c)可以看到,结节的轮廓边缘被标记为较亮区域,但背景的许多部分也存在较亮区域。从

图6(d)可以看到,除结节内部区域以及部分背景外,其余位置亮度均较高。从图6(e)和图6(f)可以看到,随着网络层级的加深,权重图逐渐变得抽象,但是仍能看出在结节边缘部分较亮,而背景部分则相对较暗。

在浅层编码器路径提取到的特征张量中,结节边缘的轮廓信息被注意力模块赋予较大的权重,但是由于浅层网络的感受野较小,不能很好地感知整幅图像的全局信息,所以一些背景的空间信息同样也被赋予了较大的权重。而随着网络层次的加深,空间注意力模块的感受野会逐渐变大,因此能更好地区分前景和背景区域,从而对目标区域以外的空间位置信息会予以一定程度上的抑制。

图7给出了注意力模块输出特征张量中几个通道可视化后的结果。图7(a)是原始图像,图7(b)是专家标注的金标准,图7(c)—图7(e)是图1中最上层注意力模块输出特征张量选取比较有代表性的通道可视化后的结果,并在图的下方标注了该通道所对应的权重。

从图7(c)—图7(e)可以看出,该特征张量的通道6有较低的权重,而通道13和通道39有较高的权重。从图7可以看出通道6提取出的结果因其包含许多无用信息和噪声,所以难以分辨出有用的轮廓,而通道13和通道39则都在结节的边缘区域提取到了可以区分区域和边缘信息。也就是说,通道13和通道39代表了某种抽象的特征,并且这种特征可以将结节区域与背景区域进行区分。而这也解释了注意力模块为什么在学习过程中给予了这两个通道较高的权重。

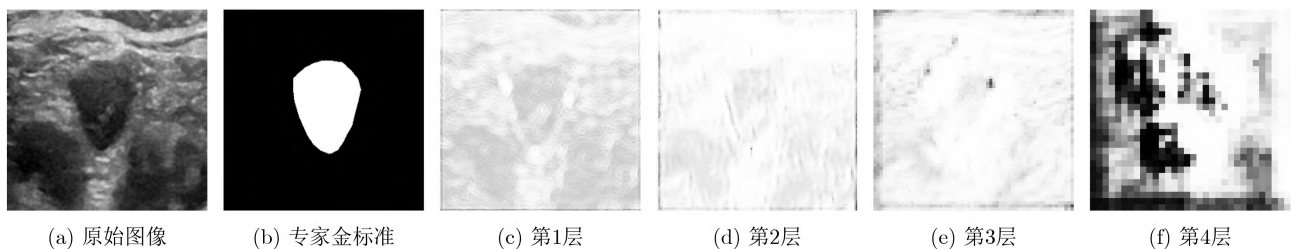


图 6 空间注意力权重图可视化后的结果

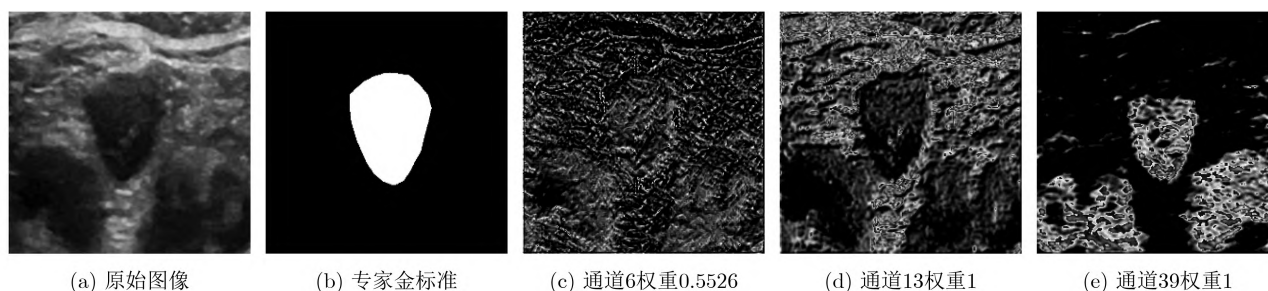


图7 通道注意力模块输出特征张量的不同通道可视化后的结果

5 结束语

本文提出了一种基于改进U-Net网络的甲状腺结节超声图像的分割方法,采用残差结构和多尺度卷积对U-Net网络的主干部分进行了优化,引入注意力机制进一步提升了网络的分割精度。通过实验证明本文所提出的方法较以往算法更加有效地提高了分割结果的准确率,对不同尺寸的结节都有较好的分割效果。如何减少网络结构的参数,以及将这种分割框架应用于其他医学图像分割任务是下一步的研究方向。

参 考 文 献

- [1] HAUGEN B R, ALEXANDER E K, BIBBLE K C, *et al.* 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer[J]. *Thyroid*, 2016, 26(1): 1–133. doi: 10.1089/thy.2015.0020.
- [2] 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. doi: 10.1109/TITB.2006.890018.
- [3] 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*, 2009, 13(4): 519–527. doi: 10.1109/TITB.2008.2007192.
- [4] 邵蒙恩, 严加勇, 崔峭峭, 等. 基于CV-RSF模型的甲状腺结节超声图像分割算法[J]. *生物医学工程研究*, 2019, 38(3): 336–340. doi: 10.19529/j.cnki.1672-6278.2019.03.15.
- [5] 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: 5. doi: 10.1186/2047-2501-1-5.
- [6] ALRUBAIDI W M H, PENG Bo, YANG Yan, *et al.* An interactive segmentation algorithm for thyroid nodules in ultrasound images[C]. The 12th International Conference on Intelligent Computing, Lanzhou, China, 2016: 107–115. doi: 10.1007/978-3-319-42297-8_11.
- [7] RONNEBERGER O, FISCHER P, and BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]. The 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 2015: 234–241. doi: 10.1007/978-3-319-24574-4_28.
- [8] CHAN T F and VESE L A. Active contours without edges[J]. *IEEE Transactions on Image Processing*, 2001, 10(2): 266–277. doi: 10.1109/83.902291.
- [9] LI Chunming, KAO C Y, GORE J C, *et al.* Minimization of region-scalable fitting energy for image segmentation[J]. *IEEE Transactions on Image Processing*, 2008, 17(10): 1940–1949. doi: 10.1109/TIP.2008.2002304.
- [10] DING Jianrui, HUANG Zichen, SHI Mengdie, *et al.* Automatic thyroid ultrasound image segmentation based on u-shaped network[C]. The 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, Suzhou, China, 2019: 1–5. doi: 10.1109/CISP-BMEI48845.2019.8966062.
- [11] WANG Jianrong, ZHANG Ruixuan, WEI Xi, *et al.* An attention-based semi-supervised neural network for thyroid nodules segmentation[C]. 2019 IEEE International Conference on Bioinformatics and Biomedicine, San Diego, USA, 2019: 871–876. doi: 10.1109/BIBM47256.2019.8983288.
- [12] WU Yating, SHEN Xuiliang, BU Feng, *et al.* Ultrasound image segmentation method for thyroid nodules using ASPP fusion features[J]. *IEEE Access*, 2020, 8: 172457–172466. doi: 10.1109/ACCESS.2020.3022249.
- [13] ABDOLALI F, KAPUR J, JAREMKO J L, *et al.* Automated thyroid nodule detection from ultrasound imaging using deep convolutional neural networks[J].

- Computers in Biology and Medicine*, 2020, 122: 103871. doi: 10.1016/j.combiomed.2020.103871.
- [14] OKTAY O, SCHLEMPER J, FOLGOC L L, *et al.* Attention U-Net: Learning where to look for the pancreas[J]. arXiv preprint arXiv: 1804.03999, 2018.
- [15] WOO S, PARK J, LEE J Y, *et al.* CBAM: Convolutional block attention module[C]. The 15th European Conference on Computer Vision, Munich, Germany, 2018: 3–19. doi: 10.1007/978-3-030-01234-2_1.
- [16] FU Jun, LIU Jing, TIAN Haijie, *et al.* Dual attention network for scene segmentation[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 2019: 3146–3154. doi: 10.1109/CVPR.2019.000326.
- [17] LEE H J, KIM J U, LEE S, *et al.* Structure boundary preserving segmentation for medical image with ambiguous boundary[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020: 4817–4826. doi: 10.1109/CVPR42600.2020.00487.
- [18] ZHONG Zilong, LIN Zhongqiu, BIDART R, *et al.* Squeeze-and-attention networks for semantic segmentation[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020: 13065–13074. doi: 10.1109/CVPR42600.2020.01308.
- [19] SZEGEDY C, IOFFE S, VANHOUCKE V, *et al.* Inception-v4, inception-ResNet and the impact of residual connections on learning[J]. arXiv preprint arXiv: 1602.07261, 2016.
- [20] KINGMA D P and BA J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv: 1412.6980, 2017.
- [21] AZAD R, ASADI-AGHBOLAGHI M, FATHY M, *et al.* Attention deeplabv3+: Multi-level context attention mechanism for skin lesion segmentation[C]. The European Conference on Computer Vision, Glasgow, UK, 2020. doi: 10.1007/978-3-030-66415-2_16.
- [22] AZAD R, ASADI-AGHBOLAGHI M, FATHY M, *et al.* Bi-directional ConvLSTM U-Net with densley connected convolutions[C]. 2019 IEEE/CVF International Conference on Computer Vision Workshop, Seoul, Korea (South), 2019: 406–415, doi: 10.1109/ICCVW.2019.00052.
- [23] ZHOU Zongwei, SIDDIQUEE M M R, TAJBAKHS N, *et al.* Unet++: A nested u-net architecture for medical image segmentation[C]. The 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, 2018: 3–11. doi: 10.1007/978-3-030-00889-5_1.
- [24] HU Jie, SHEN Li, and SUN Gang. Squeeze-and-excitation networks[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 7132–7141. doi: 10.1109/cvpr.2018.00745.
- 王 波: 男, 1982年生, 博士, 副教授, 研究方向为医学图像处理、模式识别、机器学习。
- 李梦翔: 男, 1996年生, 硕士生, 研究方向为医学图像分割、机器学习。
- 刘 侠: 男, 1975年生, 博士, 教授, 研究方向为医学图像处理、模式识别、机器学习。
- 责任编辑: 马秀强