

DOI:10.3969/j.issn.1005-202X.2026.04.008

医学影像物理

基于多尺度特征提取与多特征融合的甲状腺结节超声影像分割

丛培璐¹, 张冲¹, 贲恺¹, 赵爽², 赵文华¹, 马志庆¹

1. 山东中医药大学医学信息工程学院, 山东 济南 250355; 2. 山东中医药大学实验中心, 山东 济南 250355

【摘要】以传统的U型网络为架构,提出一种基于多尺度特征提取与多特征融合的甲状腺结节超声影像分割方法。首先,设计一种基于多个小卷积核叠加的特征提取策略。通过堆叠多个小尺寸的卷积核,模型能够在不同的感受野下捕捉图像中的细节特征和全局特征,从而实现多尺度特征的高效提取。其次,通过混合注意力机制,包括通道注意力与空间注意力,将不同阶段的特征图进行融合,以强化原有的跳跃连接。本文算法在甲状腺结节分割数据集TN3K和DDTI的95%豪斯多夫距离(HD95)分别为16.02和17.86 mm, F1分数分别为82.21%和75.74%,在所有对比算法中表现最佳。实验结果表明,该方法可以为临床医生提供辅助诊断。

【关键词】甲状腺结节; 超声; 图像分割; 多尺度特征提取; 注意力

【中图分类号】R318;R445

【文献标志码】A

【文章编号】1005-202X(2026)04-0473-07

Thyroid nodules segmentation in ultrasound image based on multi-scale feature extraction and multi-feature fusion

CONG Peilu¹, ZHANG Chong¹, YUN Kai¹, ZHAO Shuang², ZHAO Wenhua¹, MA Zhiqing¹

1. School of Medical Information Engineering, Shandong University of Traditional Chinese Medicine, Ji'nan 250355, China; 2. Experimental Center, Shandong University of Traditional Chinese Medicine, Ji'nan 250355, China

Abstract: Based on the classical U-Net, a novel method incorporating multi-scale feature extraction and multi-feature fusion for thyroid nodule segmentation in ultrasound image is proposed. Specifically, a feature extraction strategy based on stacked small-sized convolutional kernels is designed. By stacking multiple small-sized convolutional kernels, the model can capture both detailed and global features of images under different receptive fields, thereby achieving efficient multi-scale feature extraction. Through a hybrid attention mechanism which includes both channel and spatial attention, feature maps from different stages are effectively fused, thereby enhancing the original skip connections. The proposed algorithm achieves 95% Hausdorff distance (HD95) of 16.02 and 17.86 mm on the TN3K and DDTI thyroid nodule segmentation datasets, respectively, along with F1-scores of 82.21% and 75.74%, outperforming all other compared methods. Experimental results demonstrate that this approach can provide valuable assistance to clinicians in diagnostic practice.

Keywords: thyroid nodule; ultrasound; image segmentation; multi-scale feature extraction; attention

前言

甲状腺是重要的人体内分泌器官,通过分泌甲状腺激素来调节机体的代谢。甲状腺结节是由甲状腺组织细胞的异常导致的离散性病变,它是内分泌

系统中常见的一种疾病^[1]。根据结节的质地,甲状腺结节通常可分为实性结节、囊性结节以及混合性结节^[2]。一般来说,无论是实性结节还是囊性结节,都有发生癌变的可能。根据2017年国家癌症中心发布的最新资料显示,我国的甲状腺癌整体发病率占恶性肿瘤发病率第7位,并且这种趋势正在逐年增长^[3]。因此,对于患者而言,及早发现、及早诊断和及早治疗是至关重要的。电子计算机断层扫描、磁共振成像以及超声成像等常被用于甲状腺疾病的诊断^[4]。在这些成像方法中,超声成像因其对人体无创、相对便宜和实用等优点成为最常用的甲状腺结节检测方法^[5]。据统计,大约90%的甲状腺结节是良

【收稿日期】2025-12-16

【基金项目】山东省医药卫生科技项目(202425020411); 山东中医药大学科学研究基金(KYZK2024Q30)

【作者简介】丛培璐,硕士研究生,研究方向:医学图像处理与分析, E-mail: 15662643020@163.com

【通信作者】马志庆,教授,研究方向:医学图像处理与分析, E-mail: mazhq126@163.com

性的,因此恶性结节的诊断非常依赖于临床医生的经验^[6]。同时超声图像质量较低且容易受到噪声的干扰,因此经验不足的医生很可能造成误诊^[7]。因此,需要更多的计算机辅助诊断系统来帮助医生进行诊断。

近几年来,科技的发展带动越来越多的人工智能技术用于医学影像辅助诊断^[8-9]。由Ronneberger等^[10]提出的U-Net网络凭借其编码器-解码器的结构,并且采用跳跃连接的方式将低级特征信息与高级特征信息进行结合,在医学影像分割上取得优越的性能表现。但是,U-Net在处理边界情况复杂的目标时,会出现分割结果的边界较模糊的情况。Çiçek等^[11]提出的3D U-Net网络将U-Net用于三维的医学影像分割,扩展U-Net的应用范围。3D U-Net能够更好地捕捉三维目标的形态特征与空间关系,但同时也需要更多的计算资源。Chen等^[12]提出的TransUNet网络将Transformer^[13-14]引入到U-Net中,使得模型能够有效地捕捉全局上下文信息,解决卷积神经网络无法进行长距离依赖建模的问题,从而有助于更好地理解图像的语义信息和目标之间的关系。在甲状腺结节分割方面,Li等^[15]将全卷积网络应用于甲状腺结节的分割任务,该网络采用正则化技术以增强其泛化能力。通过在测试集上的评估,该方法表现出了比传统水平集方法和卷积神经网络更优越的性能,揭示全卷积网络在改善甲状腺结节诊断方面的巨大潜力。Gong等^[16]提出一个多任务学习网络,该网络能够同时分割甲状腺区域和结节区域。这样在分割结节时,可以加入有关甲状腺体重要的位置特征信息并且可以过滤掉与甲状腺结节不相关的部分腺体先验特征。Sun等^[17]提出的双路径分割网络TNSNet,该网络由区域路径和形状路径组成,通过利用形状路径在区域路径上进行软形状监

督,实现甲状腺结节的准确分割。但是,上述的甲状腺结节分割算法首先并未考虑到目标对象的尺度差异。其次,由于超声影像的质量普遍比较低且包含大量噪声,这些算法无法有效地结合各阶段的特征图以去噪,因此可能会出现误分割或无法识别出结节的情况。

针对上述问题,本文提出一种基于多尺度特征提取与多特征融合的分割方法。多尺度特征提取模块能够有效地从不同阶段的特征图中提取出不同尺度的特征信息。多特征融合模块则通过混合注意力融合这些不同尺度的特征信息。通过这种方式,可以过滤掉无关的噪声,使模型能够在低质量的超声影像中准确识别分割甲状腺结节区域,为临床甲状腺结节的诊断提供可靠的指导。

1 模型与方法

本文以传统的U-Net网络为框架,采用由多个小卷积核叠加而成的多尺度特征提取模块代替原本U-Net中每一阶段的普通卷积。为了实现更精细的特征信息,使用全局混合注意力特征融合的策略,替代原本的跳跃连接部分。在编码器阶段,图像在经多尺度特征提取模块后,使用最大池化进行下采样。输入的超声影像原图经过第1阶段的多尺度特征提取模块生成的特征图尺寸不变,只是改变通道数目。经过从第2到第5阶段多尺度特征提取模块以及下采样的处理后,图像所生成的特征图尺寸分别变为输入尺寸的1/2、1/4、1/8和1/16。在解码器阶段,下一阶段的特征图与上一阶段经过全局混合注意力特征融合的特征图拼接后,使用多尺度特征提取模块提取不同尺度的特征信息。整个网络结构的示意图如图1所示。

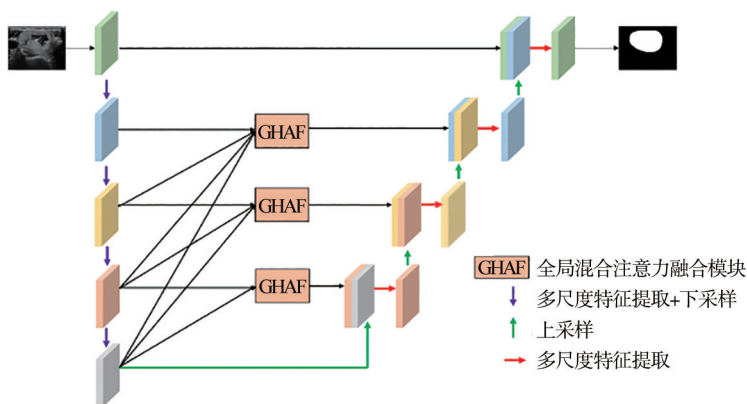


图1 网络结构图

Figure 1 Network structure

1.1 多尺度特征提取

在甲状腺结节的超声影像处理中,由于结节尺寸的大小会对影像特征产生重要影响,因此特征提取方法的选择对于精确诊断至关重要。然而,由于不同尺度的甲状腺结节超声影像中需要分割的结节尺度各异,使用固定尺寸的卷积核很难获取到不同尺度的特征信息。为了解决这个问题,本文采用多尺度残差连接模块来提取从不同尺度的甲状腺结节超声影像中得到的特征。

多尺度残差连接模块由3个3×3卷积层、1个1×1卷积层、批量归一化层^[18]以及线性整流函数层(ReLU)^[19]组成。根据文献[20]所述,2个3×3的卷积操作实际上类似于1个5×5的卷积操作,3个3×3的卷积操作近似于1个7×7的卷积操作。使用小的卷积核替代大的卷积核既可以减少时间花费又可以降低计算消耗。来自上一层的特征图输入后,经过第1个3×3卷积,保存得到的特征图;再经过1个3×3的卷积层,这时已经经过2次3×3的卷积,相当于1次5×5的卷积,将两次卷积得到的特征图保存;最后经过第3个3×3的卷积层,这时相当于经过1个7×7的卷积操作,保存第3次卷积得到的特征图。将第1次、第2次和第3次的特征图连结起来,与把输入的特征图经过1×1卷积操作后得到的特征图进行残差连接后,就得到最终的输出特征图^[21]。多尺度残差连接模块示意图如图2所示。

1.2 全局混合注意力融合

在原始的U-Net中,特征信息从浅层到深层传递时,会失去一些全局的上下文信息。这是因为浅层

的特征图包含更多的细节和局部信息,而深层的特征图则更关注全局结构和上下文信息。然而,跳跃连接虽然能够结合本层的语义信息,但有时也会引入一些无关的噪声。本文提出全局混合注意力融合(Global Hybrid Attention Fusion, GHAF)模块来结合不同阶段的语义信息,以实现全局多特征的融合。以第3阶段的GHAF模块为例说明该模块的结构,模块的示意图如图3所示。

首先,将第4阶段和第5阶段的特征图经过3×3卷积映射到与第3阶段的特征图通道数目相同。其次,将生成的特征图用双线性插值法上采样到与第3阶段的特征图一样的大小并进行级联。最后,为了从不同层次的特征图中提取全局上下文信息,通过多尺度残差连接模块来提取不同尺度的特征信息。该过程可按式表示:

$$X = \text{MRConv}(\text{Concat}(\text{Conv}(X_1), \text{Up}(\text{Conv}(X_2)), \text{Up}(\text{Conv}(X_3)))) \quad (1)$$

其中,MRConv表示多尺度特征提取,Concat表示连接操作,Up表示上采样操作。

在混合注意力处理中,首先获得的特征图会经过通道注意力模块(Channel Attention Module, CAM)。为了强化关键性特征并削弱次要特征,CAM通过对每个通道分配不同的权重。具体而言,特征图的空间维度会分别应用自适应最大池化(Adaptive Max Pooling, AMP)和自适应平均池化(Adaptive Average Pooling, AAP)操作,以提取每个通道在整个特征图中最具代表性的特征。随后,针对每个池化后的特征图,采用卷积操作和ReLU激活函数以减少通道数。紧接着,另一卷积操作被用来恢复原始通道数。将这两个恢复的特征图相加后,通过Sigmoid激活函数来估算注意力权重。最终,这些权重与原始输入特征图通过Hadamard积结合,并与输入进行残差连接,得到通道注意力增强后的特征图。CAM操作过程可按式表示:

$$F = \sigma(\text{Conv}(\text{ReLU}(\text{Conv}(\text{AAP}(X)))) + \text{Conv}(\text{ReLU}(\text{Conv}(\text{AAP}(X)))) \otimes X + X \quad (2)$$

其中, σ 表示Sigmoid激活函数, \otimes 表示Hadamard积。

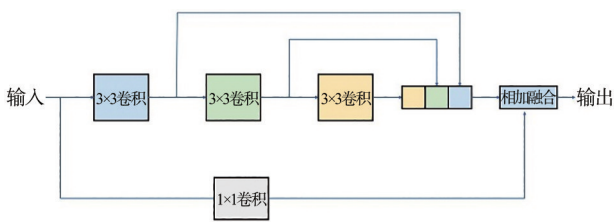
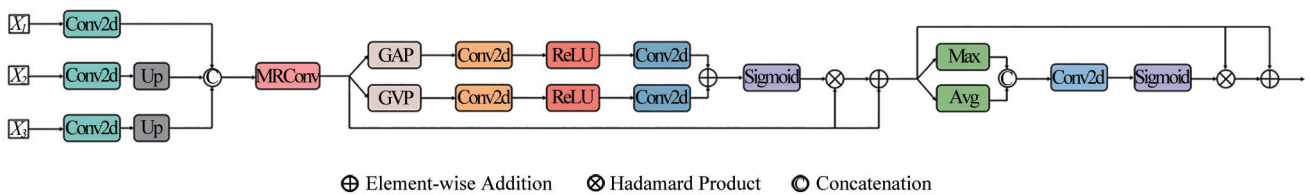


图2 多尺度残差连接模块

Figure 2 Multi-scale residual connection module



⊕ Element-wise Addition ⊗ Hadamard Product ● Concatenation

图3 第3阶段的GHAF模块示意图

Figure 3 Schematic diagram of the GHAF module of the 3rd stage

经过CAM处理的特征图随后会进入空间注意力模块(Spatial Attention Module, SAM)。SAM专注于输入图像的特定区域。首先,特征图在通道维度上通过最大值和平均值的聚合来强调局部特征。随后,采用大核卷积层来增强特征之间的局部上下文关联。接着,利用Sigmoid激活函数计算空间注意力权重。最后,这些权重同样通过Hadamard积与输入特征图结合,并进行残差连接,从而得到GHAF模块的最终输出特征图。GHAF操作过程可按下式表示:

$$F = \sigma(\text{Conv}([\text{Avgpool}(X); \text{Maxpool}(X)])) \otimes X + X \quad (3)$$

其中, σ 表示Sigmoid激活函数, \otimes 表示Hadamard积。

为了节省计算资源,本文设计的网络只包含3个GHAF模块。通过在编码器和解码器之间嵌入多个GHAF模块,可以将高级的特征信息引导至不同阶段,同时过滤掉无关的噪声,进一步增强模型的分割性能以及泛化性能。

2 实验与评价指标

2.1 实验配置

实验的操作系统为Ubuntu 20.04,编程语言Python 3.8,深度学习框架为PyTorch,显卡为NVIDIA Tesla A4000 16 G。实验优化器采用自适应矩估计(Adam)^[22],初始学习率为0.0001,批大小设为16,迭代次数200次,使用余弦退火学习速率调度器来动态调整学习率。

本文采用的损失函数是二元交叉熵损失(BCE)和戴斯相似系数(Dice)损失的结合,定义如下所示:

$$\text{BCE} = -y \cdot \log \hat{y} - (1 - y) \cdot \log(1 - \hat{y}) \quad (4)$$

$$\text{Dice} = 1 - \frac{2 \cdot |y \cap \hat{y}|}{|y| + |\hat{y}|} \quad (5)$$

$$L = \alpha \text{BCE}(\hat{y}, y) + \beta \text{Dice}(\hat{y}, y) \quad (6)$$

其中, \hat{y} 是预测结果的像素数, y 是真实目标的像素数。本文采取文献[23]的设置, α 设为0.5, β 设为1。

2.2 数据预处理

本文的数据采用由文献[24]和文献[25]创建的甲状腺结节超声影像数据集:TN3K和DDTI。TN3K数据集包括3493张图片,本实验划分2879张图片用于训练,614张图片用于测试。DDTI数据集包括637张图片,本实验划分511张图片用于训练,126张图片用于测试。所有图片都被统一成256×256的大小,并且使用水平和垂直翻转、随机旋转等数据增强方式。

2.3 评价指标

为了定量评估本文提出模型的性能,本文采用交并比(Intersection over Union, IoU),95%豪斯多夫

距离(95% Hausdorff Distance, HD95)以及F1分数作为模型的性能评估指标。其中,IoU通常用于评估模型预测值与标签真实值的重叠度,其取值为0~1,其值越接近1表示分割效果越好;HD95度量模型预测值和标签真实值之间的体积表面距离,其值越小表示分割效果越好。

3 实验与结果分析

3.1 模型分类性能研究

为了全面评估本文提出算法的性能,与FCN^[8]、Deeplabv3+^[26]、U-Net^[10]、ResUNet++^[27]、MultiResUNet^[28]、SGUNet^[29]、DC-UNet^[30]、TransUNet^[12]、MISSFormer^[31]、MA-UNet^[32]以及CPCANet^[33]进行对比分析,表1和表2分别展示各种算法在TN3K数据集和DDTI数据集上测试的分割结果。从表1和表2结果可知,本文算法在IoU、HD95和F1分数这3个评价指标上均表现出色,在TN3K数据集上分别达到73.34%、16.02 mm和82.21%,在DDTI数据集上分别达到64.18%、17.86 mm和75.74%。本文算法相较于U-Net在这3个评价指标上均有显著提升,U-Net仅包含普通的卷积操作和简单的跳跃连接,这突显本文采用的多尺度特征提取方法和全局混合注意力融合方法在提升算法表现的有效性。在衡量模型性能的IoU方面,本文提出的算法相较于对比算法中在TN3K数据集上取得最优表现的TransUNet提升0.78%,在DDTI数据集上比取得最优表现的DC-UNet提升1.84%。在F1分数方面,本文算法在TN3K数据集上比最优的Deeplabv3+提升1.06%,在DDTI数据集上比最优的Deeplabv3+提升1.73%。此外,在HD95指标上,本文算法在TN3K数据集上比最优的Deeplabv3+降低1.57 mm,在DDTI数据集上比最优的ResUNet++降低1.11 mm。与同样引入注意力机制的MA-UNet和CPCANet相比,本文提出的算法在全部3项评价指标上均取得更优的性能,这主要归因于所设计的多尺度注意力特征融合机制能够更有效地整合不同尺度的上下文信息,从而提升模型的整体表征能力。

甲状腺结节分割结果对比如图4所示,每一行表示同一张甲状腺结节超声影像通过不同模型分割所对应的分割结果,第1列为结节的超声原图,第2列为原图对应的真实手工标签图,第3列~第11列分别为FCN、Deeplabv3+、U-Net、ResUNet++、MultiResUNet、SGUNet、DC-UNet、TransUNet、MISSFormer、MA-UNet、CPCANet和本文算法的分割结果图。从图中可以看出,本文提出的算法获得

表1 甲状腺结节分割在TN3K数据集上的结果对比
Table 1 Comparison of thyroid nodule segmentation results on the TN3K dataset

方法	IoU/%	HD95/mm	F1/%
FCN	70.61	21.76	79.92
Deeplabv3+	71.98	17.59	81.15
U-Net	69.42	21.74	78.75
ResUNet++	71.76	18.99	80.84
MultiResUNet	61.37	25.76	72.28
SGUNet	70.86	22.66	80.09
DC-UNet	69.17	25.27	78.73
TransUNet	72.56	18.20	81.04
MISSFormer	65.27	20.47	75.23
MA-UNet	67.84	23.92	77.79
CPCANet	55.30	33.45	65.83
本文算法	73.34	16.02	82.21

表2 甲状腺结节分割在DDTI数据集上的结果对比
Table 2 Comparison of thyroid nodule segmentation results on the DDTI dataset

方法	IoU/%	HD95/mm	F1/%
FCN	58.60	21.28	71.18
Deeplabv3+	61.93	19.00	74.01
U-Net	60.46	22.54	72.60
ResUNet++	61.39	18.97	73.94
MultiResUNet	59.12	26.61	71.59
SGUNet	58.57	21.22	71.31
DC-UNet	62.34	20.91	72.98
TransUNet	59.88	22.36	71.10
MISSFormer	52.62	20.38	66.19
MA-UNet	51.50	29.31	64.73
CPCANet	53.75	23.89	67.10
本文算法	64.18	17.86	75.74

最好的分割结果。观察第1行和第3行,可以发现由于数据集的图像质量较低,这些算法会把非结节区域错误地分割为结节区域。这一现象影响分割的准确性,进而影响对结节的准确识别,而本文提出的算法则能够有效地解决上述问题。通过全局多特征融合方法,本文算法能够融合不同阶段的特征图,过滤掉无关的噪声,从而实现结节区域的准确分割。如

第2行和第4行所示,其余对比算法会出现不能识别多个结节的情况,原因在于未能提取到多尺度的特征信息。综上所述,通过充分利用多尺度的特征信息,以及采用全局多特征融合的方法,本文提出的算法在甲状腺结节分割的准确性上表现出色,明显优于其他对比算法。

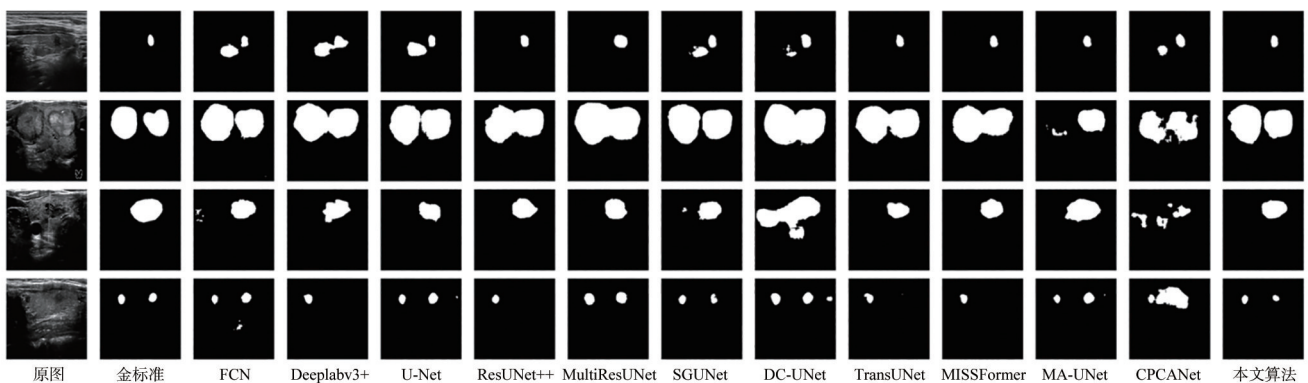


图4 不同网络的甲状腺结节分割结果对比

Figure 4 Comparison of thyroid nodule segmentation results of different networks

3.2 消融实验

为了验证全局混合注意力特征融合方法对实验结果的影响,本文分别在TN3K和DDTI数据集上进行实验。具体包括以下3部分:(1)不加入GHAF模块,使用跳跃连接;(2)加入GHAF模块但不使用混合注意力;(3)本文提出的GHAF模块。实验结果如表

3和表4所示,加入GHAF模块但不使用混合注意力后,IoU与F1分数在TN3K数据集上分别提高2.54%和2.00%,HD95下降2.66 mm;在DDTI数据集上分别提高3.11%和2.62%,HD95下降2.13 mm。此外,添加混合注意力后,与不使用混合注意力相比IoU与F1分数在TN3K数据集上分别提高1.13%和1.24%,

HD95 下降 2.47 mm; 在 DDTI 数据集上分别提高 0.78% 和 0.77%, HD95 下降 0.75 mm。这一数据表明, 全局混合注意力特征融合方法对于提高模型的分割性能起到至关重要的作用。

表3 GHAF 模块在 TN3K 数据集上的消融实验

Table 3 Ablation study of the GHAF module on the TN3K dataset

方法	IoU/%	HD95/mm	F1/%
Baseline	69.67	21.15	78.97
Baseline+GHAF_w/o_HA	72.21	18.49	80.97
Baseline+GHAF	73.34	16.02	82.21

表4 GHAF 模块在 DDTI 数据集上的消融实验

Table 4 Ablation study of the GHAF module on the DDTI dataset

方法	IoU/%	HD95/mm	F1/%
Baseline	60.29	20.74	72.35
Baseline+GHAF_w/o_HA	63.40	18.61	74.97
Baseline+GHAF	64.18	17.86	75.74

分割结果对比如图5所示, 每一行表示同一张甲状腺结节超声影像的分割结果, 第1列为结节的超声原图, 第2列为原图对应的真实手工标签图, 第3列为没有GHAF模块时的分割结果图, 第4列为GHAF模块未使用混合注意力时的分割结果图, 第5列为使用GHAF模块的分割结果图。从图5可知, 由于超声影像质量的限制, 在未使用GHAF模块之前, 模型在图像分割任务中呈现出误分割现象, 具体表现为对多个结节的识别不足以及分割结果的不完整。然而, 在引入GHAF模块但尚未应用混合注意力机制时, 上述问题得到一定程度的缓解, 但并未完全消除。当完整采用GHAF模块后, 模型在各项性能指标上均展现出最佳状态, 从而充分验证GHAF模块在提升算法整体性能方面的显著作用。

4 结论

本文提出一种结合多尺度特征提取与全局混合注意力特征融合的分割方法, 对甲状腺结节超声影像进行结节的分割。由多个小卷积核叠加实现的多尺度特征提取模块用于提取不同尺度的特征信息,

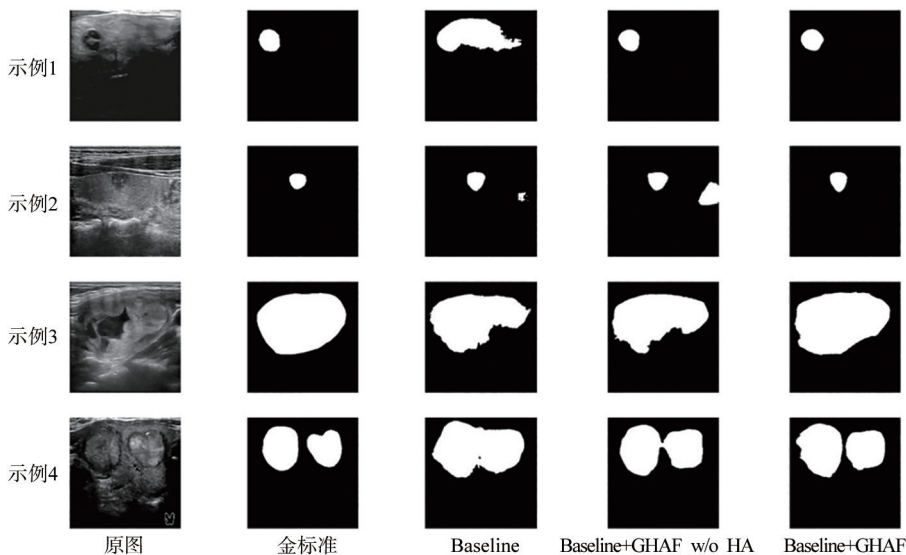


图5 消融实验分割结果对比图

Figure 5 Comparison of the segmentation results in the ablation study

从而提供更丰富的特征表示。全局混合注意力模块通过混合注意力来结合不同阶段的不同尺度的特征信息, 相较于原本 U-Net 网络中使用的跳跃连接, 全局混合注意力模块实现全局多特征的融合, 提高模型在处理不同尺度目标时的性能。本文算法与其他算法对比, 在 TN3K 和 DDTI 数据集上均展现出最优的表现。这表明本文提出的方法具有实际应用价值, 可以为临床医生提供可靠的参考依据, 有助于更加准确地诊断甲状腺结节相关疾病。

【参考文献】

[1] 中华医学会内分泌学分会, 中华医学会外科学分会甲状腺及代谢外科学组, 中国抗癌协会头颈肿瘤专业委员会, 等. 甲状腺结节和分化型甲状腺癌诊治指南(第二版)[J]. 中华内分泌代谢杂志, 2023, 39(3): 181-226.
Chinese Society of Endocrinology, Thyroid and Metabolism Surgery Group of the Chinese Society of Surgery, China Anti-Cancer Association, Chinese Association of Head and Neck Oncology, et al. Guidelines for the diagnosis and management of thyroid nodules and differentiated thyroid cancer (second edition)[J]. Chinese Journal of Endocrinology and Metabolism, 2023, 39(3): 181-226.

[2] 中华医学会超声医学分会浅表器官和血管学组, 中国甲状腺与乳腺超声人工智能联盟. 2020 甲状腺结节超声恶性危险分层中国指南: C-TIRADS[J]. 中华超声影像学杂志, 2021, 30(3): 185-200.

- Superficial Organ and Vascular Ultrasound Group, Society of Ultrasound in Medicine, Chinese Medical Association, Chinese Artificial Intelligence Alliance for Thyroid and Breast Ultrasound. 2020 Chinese guidelines for ultrasound malignancy risk stratification of thyroid nodules: the C-TIRADS [J]. Chinese Journal of Ultrasonography, 2021, 30(3): 185-200.
- [3] Chen WQ, Zheng RS, Zhang SW, et al. Cancer incidence and mortality in China in 2013: an analysis based on urbanization level[J]. Chin J Cancer Res, 2017, 29(1): 1-10.
- [4] Chen JY, You HJ, Li K. A review of thyroid gland segmentation and thyroid nodule segmentation methods for medical ultrasound images [J]. Computer methods and programs in biomedicine, 2020, 185: 105329.
- [5] Tao Z, Dang H, Shi YT, et al. Local and context-attention adaptive LCA-Net for thyroid nodule segmentation in ultrasound images[J]. Sensors (Basel), 2022, 22(16): 5984.
- [6] Durante C, Grani G, Lamartina L, et al. The diagnosis and management of thyroid nodules: a review[J]. JAMA, 2018, 319(9): 914-924.
- [7] Chi JN, Li ZL, Sun ZY, et al. Hybrid transformer UNet for thyroid segmentation from ultrasound scans[J]. Computers in biology and medicine, 2023, 153: 106453.
- [8] 徐航, 随力, 张靖雯, 等. 卷积神经网络在医学图像分割中的研究进展[J]. 中国医学物理学杂志, 2019, 36(11): 1302-1306.
Xu H, Sui L, Zhang JW, et al. Progress on convolutional neural network in medical image segmentation [J]. Chinese Journal of Medical Physics, 2019, 36(11): 1302-1306.
- [9] Chen LC, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation [J]. arXiv preprint arXiv:1706.05587, 2017.
- [10] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.
- [11] Çiçek Ö, Abdulkadir A, Lienkamp SS, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2016. Cham: Springer International Publishing, 2016: 424-432.
- [12] Chen J, Lu Y, Yu Q, et al. Transunet: Transformers make strong encoders for medical image segmentation[J]. arXiv preprint arXiv: 2102.04306, 2021.
- [13] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017: 6000-6010.
- [14] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.
- [15] Li XW, Wang SJ, Wei X, et al. Fully convolutional networks for ultrasound image segmentation of thyroid nodules[C]//2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS). Piscataway, NJ, USA: IEEE, 2018: 886-890.
- [16] Gong HF, Chen JX, Chen GQ, et al. Thyroid region prior guided attention for ultrasound segmentation of thyroid nodules[J]. Comput Biol Med, 2023, 155: 106389.
- [17] Sun JW, Li CY, Lu ZD, et al. TNSNet: thyroid nodule segmentation in ultrasound imaging using soft shape supervision [J]. Comput Methods Programs Biomed, 2022, 215: 106600.
- [18] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd International Conference on International Conference on Machine Learning. Chia Laguna Resort, Sardinia, Italy: PMLR, 2015: 448-456.
- [19] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521 (7553): 436-444.
- [20] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 2818-2826.
- [21] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning [C]//Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI Press, 2017: 4278-4284.
- [22] Kingma DP, Ba J. Adam: a method for stochastic optimization [J]. arXiv preprint arXiv:1412.6980, 2014.
- [23] Valanarasu JM, Patel VM. UNeXt: MLP-based rapid medical image segmentation network[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2022. Cham: Springer Nature Switzerland, 2022: 23-33.
- [24] Gong HF, Chen GQ, Wang RR, et al. Multi-task learning for thyroid nodule segmentation with thyroid region prior[C]//2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). Piscataway, NJ, USA: IEEE, 2021: 257-261.
- [25] Pedraza L, Vargas C, Narváez F, et al. An open access thyroid ultrasound image database [C]//10th International Symposium on Medical Information Processing and Analysis. Bellingham, WA, USA: SPIE, 2015: 92870W.
- [26] Chen LC, Zhu YK, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 833-851.
- [27] Jha D, Smedsrud PH, Riegler MA, et al. ResUNet++: an advanced architecture for medical image segmentation [C]//2019 IEEE International Symposium on Multimedia (ISM). Piscataway, NJ, USA: IEEE, 2019: 225-2255.
- [28] Ibtihaz N, Rahman MS. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation [J]. Neural Netw, 2020, 121: 74-87.
- [29] Pan HT, Zhou Q, Latecki LJ. SGUNET: semantic guided UNET for thyroid nodule segmentation [C]//2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). Piscataway, NJ, USA: IEEE, 2021: 630-634.
- [30] Lou AG, Guan SY, Loew M. DC-UNet: rethinking the U-Net architecture with dual channel efficient CNN for medical image segmentation [C]//Medical Imaging 2021: Image Processing. Bellingham, WA, USA: SPIE, 2021: 115962T.
- [31] Huang X, Deng Z, Li D, et al. Missformer: an effective medical image segmentation transformer[J]. arXiv preprint arXiv:2109.07162, 2021.
- [32] Sun Y, Bi FK, Gao YT, et al. A multi-attention UNet for semantic segmentation in remote sensing images [J]. Symmetry, 2022, 14 (5): 906.
- [33] Huang HJ, Chen ZG, Zou Y, et al. Channel prior convolutional attention for medical image segmentation[J]. Comput Biol Med, 2024, 178: 108784.

(编辑:陈丽霞)